



Publication Year	2019
Acceptance in OA	2021-01-04T17:54:53Z
Title	The Gaia-ESO Survey: asymmetric expansion of the Lagoon Nebula cluster NGC 6530 from GES and Gaia DR2
Authors	Wright, Nicholas J., Jeffries, R. D., Jackson, R. J., Bayo, A., Bonito, Rosaria, DAMIANI, Francesco, Kalari, V., Lanzafame, Alessandro Carmelo, PANCINO, ELENA, Parker, R. J., PRISINZANO, Loredana, RANDICH, Maria Sofia, Vink, J. S., Alfaro, E. J., Bergemann, M., FRANCIOSINI, Elena, Gilmore, G., Gonneau, A., Hourihane, A., Jofré, P., Koposov, S. E., Lewis, J., MAGRINI, LAURA, MICELA, Giuseppina, MORBIDELLI, LORENZO, SACCO, GIUSEPPE GERMANO, Worley, C. C., ZAGGIA, Simone
Publisher's version (DOI)	10.1093/mnras/stz870
Handle	http://hdl.handle.net/20.500.12386/29465
Journal	MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY
Volume	486

The *Gaia*-ESO Survey: asymmetric expansion of the Lagoon Nebula cluster NGC 6530 from GES and *Gaia* DR2

Nicholas J. Wright,^{1★} R. D. Jeffries,¹ R. J. Jackson,¹ A. Bayo^{1b,2,3}, R. Bonito,⁴ F. Damiani,⁴ V. Kalari,⁵ A. C. Lanzafame,⁶ E. Pancino^{1b,7,8}, R. J. Parker,^{9†} L. Prisinzano,⁴ S. Randich,⁷ J. S. Vink,¹⁰ E. J. Alfaro,¹¹ M. Bergemann,¹² E. Franciosini,⁷ G. Gilmore,¹³ A. Gonneau,¹³ A. Hourihane,¹³ P. Jofré,¹⁴ S. E. Koposov,¹³ J. Lewis,¹³ L. Magrini,⁷ G. Micela,⁴ L. Morbidelli,⁷ G. G. Sacco,⁷ C. C. Worley¹³ and S. Zaggia¹⁵

¹*Astrophysics Group, Keele University, Keele ST5 5BG, UK*

²*Instituto de Física y Astronomía, Universidad de Valparaíso, Chile*

³*Núcleo Milenio Formación Planetaria – NPF, Universidad de Valparaíso, Av. Gran Bretaña 1111, Valparaíso, Chile*

⁴*INAF – Osservatorio Astronomico di Palermo, Piazza del Parlamento 1, I-90134 Palermo, Italy*

⁵*Departamento de Astronomía, Universidad de Chile, Casilla 36-D, Santiago, Chile*

⁶*Dipartimento di Fisica e Astronomia, Sezione Astrofisica, Università di Catania, via S. Sofia 78, I-95123 Catania, Italy*

⁷*INAF – Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, I-50125 Florence, Italy*

⁸*Space Science Data Center – Agenzia Spaziale Italiana, via del Politecnico, s.n.c., I-00133 Rome, Italy*

⁹*Department of Physics & Astronomy, University of Sheffield, Sheffield S3 7RH, UK*

¹⁰*Armagh Observatory, College Hill, Armagh BT61 9DG, Northern Ireland, UK*

¹¹*Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía s/n, Granada E-18008, Spain*

¹²*Max-Planck Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*

¹³*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*

¹⁴*Núcleo de Astronomía, Universidad Diego Portales, Av. Ejército 441, Santiago, Chile*

¹⁵*INAF – Osservatorio Astronomico di Padova, Vicolo della Osservatorio 5, I-35122 Padova, Italy*

Accepted 2019 March 20. Received 2019 February 22; in original form 2018 October 19

ABSTRACT

The combination of precise radial velocities from multi-object spectroscopy and highly accurate proper motions from *Gaia* DR2 opens up the possibility for detailed 3D kinematic studies of young star-forming regions and clusters. Here, we perform such an analysis by combining *Gaia*-ESO Survey spectroscopy with *Gaia* astrometry for ~ 900 members of the Lagoon Nebula cluster, NGC 6530. We measure the 3D velocity dispersion of the region to be $5.35^{+0.39}_{-0.34}$ km s^{−1}, which is large enough to suggest the region is gravitationally unbound. The velocity ellipsoid is anisotropic, implying that the region is not sufficiently dynamically evolved to achieve isotropy, though the central part of NGC 6530 does exhibit velocity isotropy that suggests sufficient mixing has occurred in this denser part. We find strong evidence that the stellar population is expanding, though this is preferentially occurring in the declination direction and there is very little evidence for expansion in the right ascension direction. This argues against a simple radial expansion pattern, as predicted by models of residual gas expulsion. We discuss these findings in the context of cluster formation, evolution, and disruption theories.

Key words: stars: formation – stars: kinematics and dynamics – open clusters and associations: individual: Lagoon Nebula, NGC 6530, M8.

1 INTRODUCTION

Stars form within turbulent and clumpy giant molecular clouds that result in a hierarchical and highly substructured spatial distribution of the youngest stars and protostars (e.g. Elmegreen 2002;

* E-mail: nick.nwright@gmail.com

† Royal Society Dorothy Hodgkin Fellow.

Gutermuth et al. 2008). The spatial distribution of very young stars, still in their natal star-forming regions, often retains this substructure (Larson 1995; Cartwright & Whitworth 2004; Gutermuth et al. 2008), though many centrally concentrated and smoothly distributed star clusters are also observed at this early time (e.g. Carpenter 2000; Pfalzner 2009). Once young stars have spatially decoupled from their birth environment the majority are found in unbound groups of some sort (Lada & Lada 2003), often in the form of associations (e.g. Brown, Dekker & de Zeeuw 1997; Wright et al. 2016).

The physical processes responsible for this evolution and the progression of stars from their formation sites to the Galactic field are not well constrained as there are many different processes that can play a role. The initial spatial and kinematic structure in which stars form can play a significant part in their future evolution (Parker et al. 2014; Parker & Wright 2016), as can dynamical interactions between stars and binaries (e.g. Marks & Kroupa 2012; Sills et al. 2018), and the evolving gravitational potentials of the stellar and gaseous parts of the system (e.g. Baumgardt & Kroupa 2007; Moeckel & Bate 2010). The latter issue has been argued to be critical to the survival or dispersal of young star clusters following residual gas expulsion (e.g. Tutukov 1978; Lada, Margulis & Dearborn 1984), though observational kinematic evidence has yet to verify this picture (Wright et al. 2016; Wright & Mamajek 2018) and there are questions over the efficiency and effectiveness of this process (Girichidis et al. 2012; Kruijssen et al. 2012; Dale, Ercolano & Bonnell 2015). The kinematics of young stars can provide important tests of these theories, allowing us to follow the motions of stars as stellar systems form, interact and dissolve.

We are currently undergoing a transformative improvement in kinematic data quantity and quality. This is driven by data from current radial velocity (RV) surveys (such as the 340-night *Gaia*-ESO Survey, GES; Gilmore et al. 2012; Randich, Gilmore & Gaia-ESO Consortium 2013) and astrometric data of unprecedented precision from *Gaia* (Gaia Collaboration 2016). Individually these datasets have provided insights into the structure and dynamics of young star forming regions (Jeffries et al. 2014; Rigliaco et al. 2016; Berlanas et al. 2019), the kinematics of runaway OB stars (Drew et al. 2018), and revealed the low-mass stellar content of dispersed OB associations and star-forming complexes (Zari et al. 2017, 2018; Armstrong, Wright & Jeffries 2018; Beccari et al. 2018). However, the real value comes from combining these data to achieve 3D kinematics, providing estimates of energies, angular momenta and dynamics, and avoiding the need to make crude isotropy assumptions. In this paper we combine GES RVs and *Gaia* proper motions (PMs) for hundreds of young stars in the Lagoon Nebula, using both spectroscopic and X-ray diagnostics of youth to compile a kinematically unbiased sample of members.

NGC 6530 is a young (1–2 Myr, Mayne et al. 2007; Bell et al. 2013) cluster projected against the centre of the Lagoon Nebula HII region (Messier 8), the illuminated part of a giant molecular cloud that extends several degrees (Lada et al. 1976; Sung, Chun & Bessell 2000), or ~ 50 pc at the estimated distance of 1326 pc (see Section 3.4). The main sources of ionization are the massive stars 9 Sgr [HD 164794, an O4V(f) star] and the massive binary HD 165052 (O6.5V+O7.5V), as well as a number of late O- and B-type stars (see e.g. Tothill et al. 2008, and Fig. 1). A rich population of relatively unreddened low-mass members has also been uncovered using optical and X-ray data (Sung et al. 2000; Damiani et al. 2004; Prisinzano et al. 2005; Kalari et al. 2015). The youngest stars in the region are associated with embedded populations in the Hourglass Nebula, which houses an ultracompact HII region, and the M8 E region around the massive protostar M8E-IR (Tothill et al. 2008).

The spatial arrangement of these separate populations is shown in Fig. 1. A recent review of the Lagoon Nebula can be found in Tothill et al. (2008).

In this paper we present the first 3D kinematic study of the Lagoon Nebula stellar population and its central cluster NGC 6530. In Section 2 we outline the observational data used and in Section 3 we discuss how a kinematically unbiased sample of members was compiled for our study. In Section 4 we present our results, including the calculation of the velocity dispersions, the search for velocity gradients, the assessment of the evidence for expansion and rotation, and a consideration of the kinematic outliers in the sample. In Section 5 we discuss our results and their implications for the understanding of young star clusters and in Section 6 we summarize our results.

2 OBSERVATIONAL DATA

In this section we describe the observational data used in this study, including GES spectroscopy, *Gaia* data release 2 (DR2) astrometry, and ancillary membership information from X-ray, infrared, and H α surveys.

2.1 *Gaia*-ESO Survey spectroscopy

The main source of data for this study is spectroscopy from GES observations of NGC 6530. The GES target selection strategy for clusters is carried out independently, but following homogeneous criteria, for each cluster observed during the survey. The targets for these observations were selected to cover an area of $55 \text{ arcmin} \times 30 \text{ arcmin}$ in the direction of NGC 6530, centred on $(\alpha, \delta) = (271.09, -24.33)$. The targets were chosen based on their location in the V versus $V - I$ colour-magnitude diagram (CMD) using photometry from Prisinzano et al. (2005) and the r' versus $r' - i'$ CMD from Drew et al. (2014). Targets were selected that fell between the 0.5 and 10 Myr Siess, Dufour & Forestini (2000) isochrones at a distance of 1.3 kpc and an extinction of $E(B - V) = 0.33$, and between V magnitudes of 12 and 19 (corresponding to pre-main sequence masses of $2.5\text{--}0.4 M_{\odot}$ at an age of 2 Myr). This led to a selection of 4066 target candidates that covered a wide area around the expected cluster pre-main sequence in the CMD (see Fig. 2), ensuring that target selection was not biased by the assumed cluster properties.

The targets were grouped according to their V magnitude and configured for multifibre spectroscopy across 27 observing blocks. These blocks were observed with the FLAMES¹ fibre-fed spectrograph at the VLT² over 17 nights in 2012 September and 2013 June to September, resulting in 1872 targets being observed. The targets that were observed are not preferentially brighter or fainter than those that were not observed, but were just those selected by the fibre allocation software. The only bias that this might lead to is the preference to select stars in non-crowded regions where it can be difficult to position multiple fibres close together. However, thanks to the large number of observing blocks observed this is unlikely to be a major bias. The observations were performed using the GIRAFFE intermediate-resolution spectrograph and the HR15N set-up, which gave spectra with a resolving power of 14 000–15 000 covering a wavelength range of 6444–6816 Å and the UVES high-resolution spectrograph with the U520 and U580 set-ups, which

¹Fibre Large Array Multi Element Spectrograph.

²Very Large Telescope.

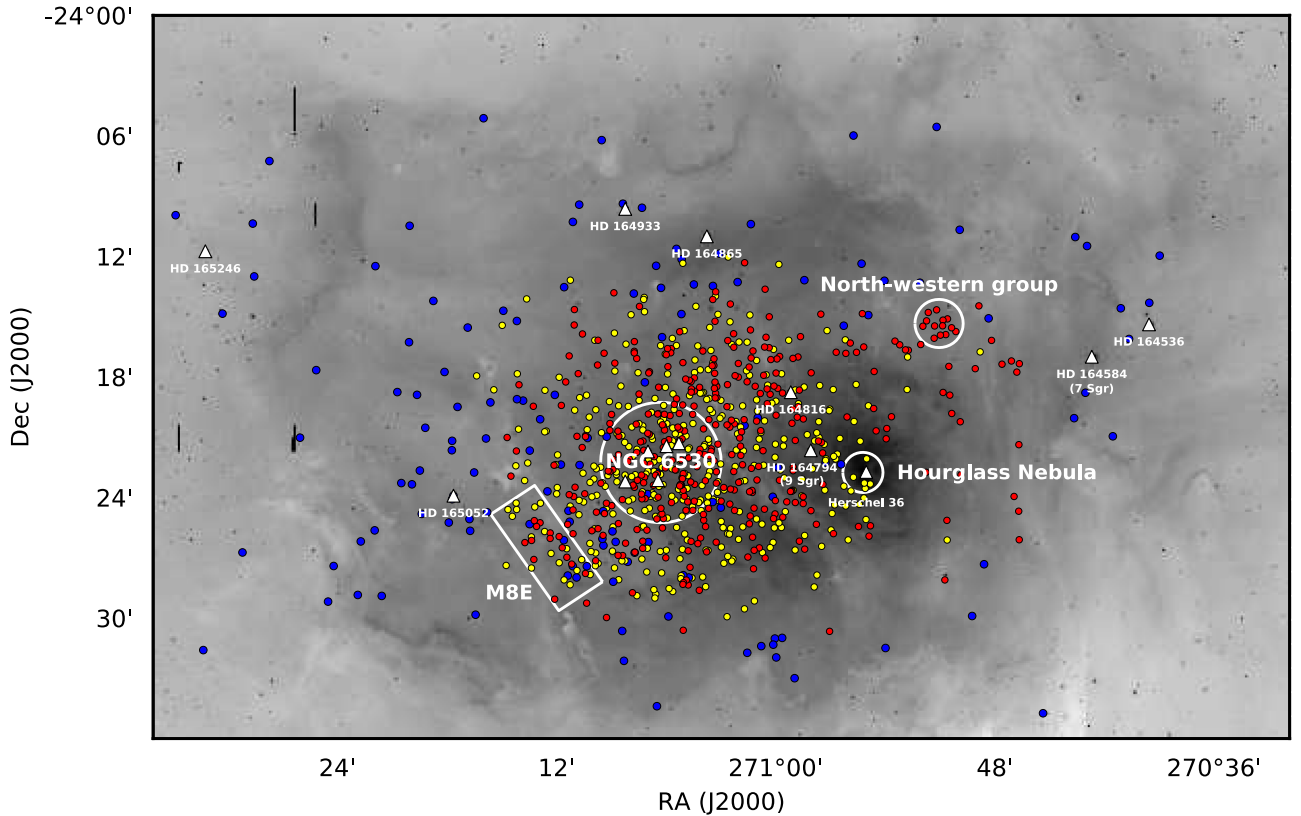


Figure 1. Map of the spatial distribution of our final sample of young stellar objects projected on to an inverted VPHAS + $H\alpha$ image. Sources are colour-coded by their origin: blue sources are selected based only on their spectroscopic GES information, red sources only on their MYSTIX membership probabilities, and yellow sources selected using both GES and MYSTIX information. The central NGC 6530 cluster, the Hourglass Nebula, the eastern ‘M8E’ group, and the north-western group of stars are indicated based on locations from Tothill et al. (2008) and Kuhn et al. (2014). The high-mass stars discussed in Section 3.3 are indicated with white triangles and labelled (HD 164906, HD 315031, CD-24 13829, CPD-24 6164, and CD-24 13837 fall within the NGC 6530 circle and are not labelled for clarity).

give a resolution of 47 000 and wavelength ranges of 4180–6212 and 4786–6830 Å, respectively. See Prisinzano et al. (2019) for more details.

GES data were reduced and analysed using common methodologies and software to produce a uniform set of spectra and stellar parameters, and are described in Jeffries et al. (2014) and Sacco et al. (2014) for the data reduction and Lanza et al. (2015) and Pancino et al. (2017) for the data analysis, calibration, and extraction of stellar parameters of stars in young clusters. In this paper we use data from the fifth internal data release of November 2015. In bright H II regions such as the Lagoon Nebula some stellar spectra can be contaminated with nebular emission lines that can affect the calculation of RVs. We therefore compared the RVs calculated from the GES pipeline (Jeffries et al. 2014) with those calculated from spectra with nebular lines masked (see Klutsch et al. in preparation) and retained those where the two methods agreed to within 3σ (75 per cent of GES RVs pass this test).

We also include in our analysis VLT/FLAMES observations of 228 stars in NGC 6530 contained in the ESO Archive. These observations, performed on 2003 May 27, were originally presented by Prisinzano et al. (2007), with targets selected from photometry presented by Prisinzano et al. (2005) and falling in the V versus $V - I$ CMD in the vicinity of X-ray selected members of the cluster (Damiani et al. 2004) and with V magnitudes between 14.0 and 18.2 mag. These data were re-reduced and analysed as part of GES following the same methodologies described above.

Combining GES and archival data and removing duplicates we have spectroscopy for 1957 objects. Out of these objects we exploit measured effective temperature for 1203 stars (61 per cent of the sample), lithium equivalent widths for 1545 stars (79 per cent), and RVs for 1326 stars (68 per cent).

2.2 Gaia DR2 astrometry

We use astrometric data from *Gaia* (Gaia Collaboration 2016) DR2 (Gaia Collaboration 2018), which contains parallaxes, ϖ , and PMs, μ_α^3 and μ_δ , calculated from the first 22 months of *Gaia* observations (2014–2016). This data achieve a very high level of astrometric precision for a sample of unprecedented size, though it is still calculated assuming single-star behaviour (Lindgren et al. 2018), meaning that close binaries (separation ≤ 100 mas) will be unresolved and the astrometry is for their photocentre. The astrometric parameters are aligned with the International Celestial Reference Frame to a precision of about 0.02 mas at epoch J2015.5 (see Lindgren et al. 2018, for details) and are absolute (they do not rely on an external reference frame and no significant residual offsets exist; Arenou et al. 2018).

³We follow standard practice and provide the PM in right ascension as $\mu_\alpha = \mu_{\alpha,0} \cos \delta$.

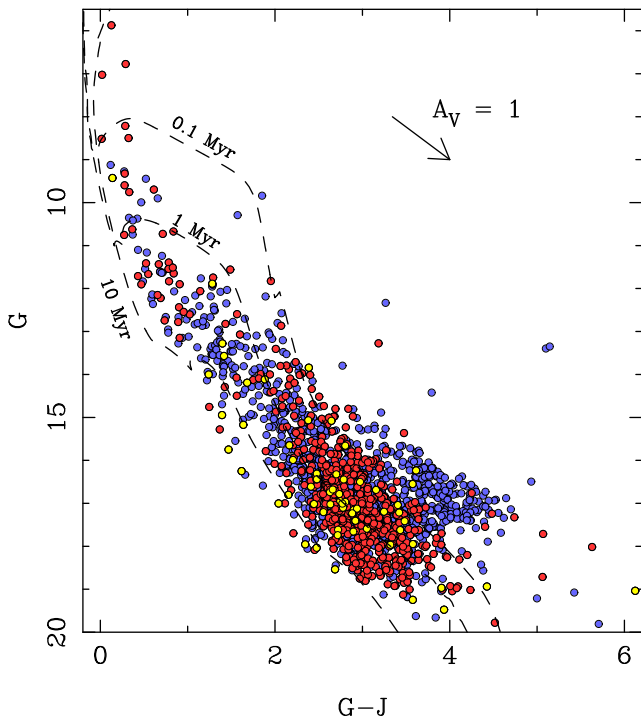


Figure 2. G versus $G - J$ CMD showing 1914 candidate members of NGC 6530 (those with both *Gaia* and 2MASS photometry from the 4066 target candidates considered for GES observations) in blue and overplotted in red and yellow are 683 of the 889 stars from our final list of members (those with both *Gaia* and 2MASS photometry). Red points are members of the Lagoon Nebula population in the main velocity distribution, while yellow points are PM kinematic outliers (see Section 4.1). Also shown are a 1 mag extinction vector and pre-main sequence stellar isochrones from Marigo et al. (2017), the latter reddened by $A_V = 1.1$ mag (Sung et al. 2000) and at a distance of 1326 pc (see Section 3.4).

DR2 is effectively complete in the range $G = 12$ to -17 mag, with a detection threshold of $G = 20.7$ mag. Very bright ($G < 7$ mag) and high proper-motion (> 0.6 arcsec yr $^{-1}$) stars suffer from incompleteness, though this should not affect us. In dense areas of the sky ($> 400\,000$ stars deg $^{-2}$) the effective magnitude limit of the survey can drop to ~ 18 mag and chance configurations can lead to confusion and mistakenly large parallaxes (Lindegren et al. 2018), though the stellar density of $220\,000$ stars deg $^{-2}$ towards NGC 6530 is not high enough to make this an issue. Spatial correlations in parallax and PM are believed to exist on small (< 1 deg) and intermediate (~ 20 deg) angular scales up to ± 0.1 mas for parallaxes and ± 0.1 mas yr $^{-1}$ for PMs (Lindegren et al. 2018; Luri et al. 2018). As a result of this, averaging parallaxes over small regions of the sky will not reduce the parallax uncertainty on the mean to below this (Luri et al. 2018). These systematic uncertainties can be estimated using tables of spatial covariances.⁴ Comparison with known quasars suggests the parallaxes have an overall negative bias of 0.03 ± 0.02 mas (i.e. 0.03 mas should be added to published values when calculating distances, Lindegren et al. 2018) and that the uncertainties may be underestimated by 8–12 per cent, or higher for bright stars and regions of higher density (Arenou et al. 2018). We do not use stellar effective temperatures from DR2 because they are derived under the assumption of no reddening (Andrae et al.

2018; Arenou et al. 2018), which is likely to be invalid for our targets.

Gaia DR2 data were downloaded in a circular region of radius 0.6 deg centred on NGC 6530, covering the existing GES observations, and including 249 351 sources. We filtered this source list based on two criteria outlined in the *Gaia* data release papers and technical notes, as described below, but did not apply other cuts based on absolute astrometric uncertainties or negative parallaxes, both of which have been shown to introduce significant biases (Arenou et al. 2018). First we required that all sources have more than eight visibility periods used for their astrometric solution, which removes sources with spurious PMs and parallaxes (Arenou et al. 2018). Then we applied the cut recommended in Lindegren (2018) using the ‘re-normalized unit weighted error’ (RUWE), requiring that $\text{RUWE} \leq 1.4$, and using the normalization factors provided by *Gaia*’s Data Processing and Analysis Consortium (DPAC). This removes sources with spurious astrometry and helps filter contamination from double stars, astrometric effects from binary stars and other contamination problems, and replaces all previously recommended astrometric filters (e.g. those outlined in Lindegren et al. 2018). We did not implement the photometric cut in equation C-2, since photometry is not the main use we will make of the *Gaia* data. Together the cuts remove 69 per cent of *Gaia* sources, the majority of which are faint, leaving 76 554 sources.

We experimented with an additional cut using the astrometric excess noise, which is a measure of the astrometric goodness of fit calculated as the extra noise per observation that has been included in the astrometric uncertainties to explain the scatter of residuals in the astrometric solution (Lindegren et al. 2018). An outright cut using this parameter has not been recommended when using *Gaia* DR2 data since this quantity is less discriminating than other cuts due to a bug in DR2 and because cuts based purely on astrometric uncertainties can lead to a biased sample (Luri et al. 2018). Despite this we repeated all astrometric analyses in this paper using a subset of our final sample made up only of sources with astrometric excess noise equal to zero. We found that our results and conclusions did not change and only that the uncertainty on many of our measured quantities actually increased, because of the significantly smaller sample used for this analysis.

The primary application of these data will be to use PMs to study the plane-of-the-sky kinematics alongside the line-of-sight kinematics from the RVs. We will briefly use the parallaxes to refine our membership and identify foreground and background contaminants, but will not exploit these data further. In the astrometric analysis that follows we always use the full (5×5) *Gaia* covariance matrix whenever propagating uncertainties, and consider all errors to follow a normal distribution (Arenou et al. 2018). In addition we increase the astrometric errors by factors of 1.093 (μ_α), 1.115 (μ_δ), and 1.081 (ϖ), in accordance with the extent to which the formal uncertainties appear to have been underestimated (Lindegren et al. 2018).

2.3 Supporting data

We complement the spectroscopic and astrometric data with supporting literature data to aid the identification of young stars across the Lagoon Nebula. This includes membership probabilities for 2427 sources from the Massive Young star-forming Complex Study in Infrared and X-rays (MYStIX; Feigelson et al. 2013) presented by Broos et al. (2013), and VST Photometric H α Survey (VPHAS; Drew et al. 2014) photometric H α equivalent widths for 235

⁴Available from <https://www.cosmos.esa.int/web/gaia/dr2-known-issues>.

sources from Kalari et al. (2015). Details of how these data were incorporated into our membership selection can be found in the next section. We also use photometry from the 2 Micron All Sky Survey (2MASS; Cutri et al. 2003). All catalogues are cross-matched with each other using a 1 arcsec matching radius. A subset of our candidate source list (those with both *Gaia* and 2MASS photometry) is shown in Fig. 2.

3 MEMBERSHIP SELECTION

In this section we outline the various indicators of stellar youth used to identify young stars in the direction of NGC 6530. To avoid any kinematic biases we have not used any sort of membership cut based on RVs or PMs.

3.1 Spectroscopic membership indicators

GES spectroscopy facilitates three separate indicators of stellar youth: the gravity index γ , the EW(Li), and the full width at zero intensity (FWZI) of the H α line. These three parameters have been used by multiple GES studies to define samples of young stars free of kinematic bias (e.g. Jeffries et al. 2014; Rigliaco et al. 2016; Sacco et al. 2017; Bravi et al. 2018), and we follow their methods here.

One of the main sources of contamination in a sample of candidate members selected from the CMD is background giants. These can be identified and removed based on their low surface gravity using the gravity index γ (Damiani et al. 2014).⁵ Fig. 3 shows γ as a function of the effective temperature for GES targets observed towards NGC 6530. Giant stars occupy the upper-right part of this plot, with $\gamma > 1$ and $T_{\text{eff}} < 5600$ K (Damiani et al. 2014). There are 1064 stars with γ and T_{eff} measurements in our sample, 233 of which meet these criteria and are classified as giants. A further 13 stars lack T_{eff} measurements and of these, 5 have $\gamma > 1$ and are also classified as giants. This leaves 839 candidate young stars.

The other major source of contamination is main-sequence stars in the foreground of NGC 6530. These sources can be separated from the young stars in NGC 6530 based on the presence of lithium in their atmospheres, since late-type stars deplete their lithium after approximately 10 (for M-type stars) to 100 Myr (for K-type stars) due to burning and subsequent mixing throughout the convection zone (Soderblom 2010). There are 827 non-giant stars with EW(Li) and T_{eff} measurements in our sample (field giants can show lithium enrichment and so both membership criteria are required), as shown in Fig. 3. Of these, 358 have EW(Li) above the threshold defined by Sacco et al. (2017) based on the distribution of EW(Li) values for stars in the 30–50 Myr cluster IC 2602 (Randich et al. 1997). A further eight candidate young stars lack T_{eff} measurements, but none of these have EW(Li) above the highest threshold value of 300 mÅ and so are not considered members.

The EW(Li) can be underestimated in stars with high-mass accretion rates if the continuum emission in excess produced by the accretion shock reduces the measured signal (e.g. Palla et al. 2005). From the 481 stars that fail the EW(Li) membership test, 82 have FWZI(H α) measures greater than 4 Å (Bonito et al. 2013;

Prisinzano et al. 2019) and 3 have photometric H α equivalent widths above the threshold defined by Barrado y Navascués & Martín (2003), and are therefore re-classified as members. This leaves 443 stars as high-confidence young stars based on their spectroscopic properties.

3.2 Non-spectroscopic membership indicators

Broos et al. (2013) estimate membership probabilities for stars towards NGC 6530 based on a combination of X-ray, near-IR, and mid-IR data. They employ a ‘Naive Bayes Classifier’ that estimates probabilities that each source is either a young star in the star forming region, a foreground or background field star, or an extragalactic object. These probabilities are estimated based on key observational quantities and likelihood models for each quantity and population. An object is assigned a membership class if the probability of the object being a member of that population is more than twice the probability of it being in any other population. The value of this dataset is not just the membership information, but also the published probabilities for each source that it is a member of a given population, which we can combine with other (e.g. spectroscopic) information to produce an improved membership probability (Section 3.4).

Broos et al. (2013) include 2427 sources in the direction of NGC 6530, of which 1828 are classified as young stars, 2 and 3 are foreground and background stars, respectively, 102 are extragalactic sources, and 492 are unclassified. Sources may be unclassified either because no individual population class has a probability more than twice as large as the next-largest probability, or if a source was initially classified as a member but lacks observational evidence for being a member other than the position-dependent prior. The total number of sources not classified as members (including unclassified sources) is consistent with the predicted number of contaminants from their simulations, which they argue supports their classification.

3.3 Known high-mass members

Previously known high-mass members of NGC 6530 and the surrounding Lagoon Nebula were gathered from the literature (e.g. Tothill et al. 2008, and others) and are listed in Table 1. Spectral types, RVs, and photometry were gathered from the literature, while parallaxes, PMs, and astrometric parameters were taken from *Gaia* DR2. Five of these stars were found to have poor DR2 astrometry (three lacking sufficient observing periods and two having high values of RUWE, see Section 2.2), and since these sources also lack literature RVs they were not included in our kinematic analysis. The star 7 Sgr (HD 164584), suspected to be a member of NGC 6530 (e.g. Damiani et al. 2017), has a large parallax that suggests it is a foreground star, and since it is not thought to be a binary (Eggleton & Tokovinin 2008) that might have uncertain astrometry, we do not include it in our analysis. The eight remaining OB stars were included in our sample.

3.4 Combining membership information

To identify a high-confidence sample of young stars in NGC 6530 and the surrounding Lagoon Nebula we combine the membership information discussed above and use *Gaia* DR2 parallaxes as an additional check on membership. PMs and RVs are not used to determine membership to prevent kinematic bias. After

⁵While pre-main sequence stars do have lower gravity than main-sequence stars it is not as low as for background giants and this method has successfully been used to separate young stars in Chamaeleon I from background stars (Sacco et al. 2017).

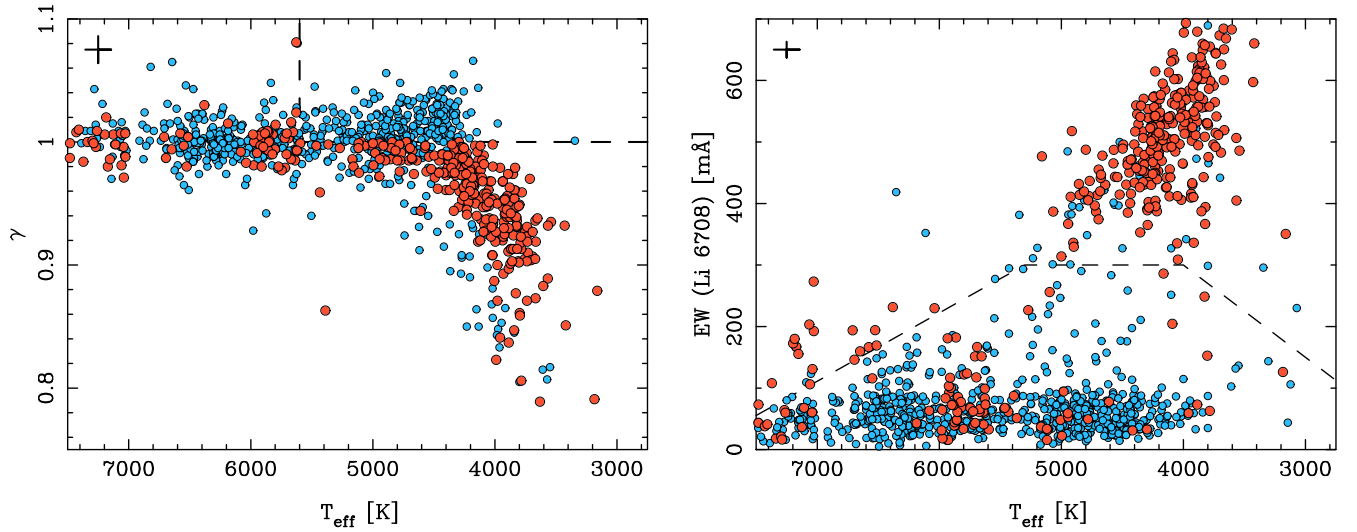


Figure 3. GES spectral quantities used for membership selection. The gravity index γ (left-hand panel) and the EW of the lithium 6708 Å line (right-hand panel) are shown as a function of effective temperature. In both panels red circles show sources from our final list of members and blue circles show non-members, as determined using both of these diagnostics and $H\alpha$ information. The dashed lines show the thresholds used to identify giants in the upper-right corner of the γ – T_{eff} plot and to identify young stars above the threshold in the EW(Li)– T_{eff} plot (stars can fall below this line and still be considered members if they exhibit $H\alpha$ emission). A typical error bar is shown in the top-left corner of each panel illustrating the typical uncertainties of 100 K in T_{eff} , 13 mÅ for EW(Li) and 0.011 for γ .

Table 1. Previously known high-mass members of the Lagoon Nebula. References are GOSS (Galactic O-Star Spectroscopic Survey, Sota et al. 2014), B89 (Boggs & Bohm-Vitense 1989), G77 (Garrison, Hiltner & Schild 1977), G89 (Gray & Garrison 1989), G14 (González, Veramendi & Cowley 2014), H65 (Hiltner, Morgan & Neff 1965), H88 (Houk & Smith-Moore 1988), and L06 (Levenhagen & Leister 2006). RVs taken from Conti, Leep & Lorre (1977), Pourbaix et al. (2004), and *Gaia* DR2. Parallaxes are from *Gaia* DR2.

Name(s)	Spectral type	Ref.	RV (km s ^{−1})	Parallax (mas)	Included in sample?
HD 164794 (9 Sgr)	O4V	GOSS	-0.8 ± 1.9	0.851 ± 0.095	Yes
HD 165052	O5.5V+O8V	GOSS	1.05 ± 0.31	0.784 ± 0.046	Yes
HD 164536	O7V+O7.5V	GOSS		0.792 ± 0.251	No, poor astrometry
Herschel 36	O7.5V	GOSS		0.902 ± 0.219	No, poor astrometry
HD 165246	O8V	GOSS		0.501 ± 0.113	No, poor astrometry
HD 164816	O9.5V+B0V	GOSS		0.844 ± 0.070	Yes
HD 164906	B0Ve	L06		0.810 ± 0.055	No, poor astrometry
HD 164865	B9Iab	G77		0.715 ± 0.052	Yes
HD 164584 (7 Sgr)	F2/3 II/III	G89	-5.62 ± 0.57	2.805 ± 0.152	No, large parallax
HD 315031	B0.5V+B1V	G14		0.690 ± 0.070	Yes
CD-24 13829	B1.5V	H65		1.000 ± 0.098	No, poor astrometry
CPD-24 6164	B1V	H65		0.841 ± 0.056	Yes
CD-24 13837	B1V	B89		0.858 ± 0.046	Yes
HD 164933	B1/2 I/II	H88		0.810 ± 0.066	Yes

cross-matching all the data we have 4080 candidate members, with the following information:

(i) 1525 sources only have spectroscopic information from GES or the literature, or photometric $H\alpha$ emission and so their membership is determined as outlined in Section 3.1, giving 214 members.

(ii) 1860 sources lack this data but have membership information from Broos et al. (2013), which we use as given (see Section 3.2), providing 1270 members.

(iii) 557 sources have both spectroscopic and Broos et al. (2013) membership information. In general these two sources show very good agreement, with the only difference being 25 sources identified as members by Broos et al. (2013) that have GES spectroscopy that suggests their surface gravities are more consistent with them being

background giants (based on the γ quantity, see Section 3.1), which we therefore list as non-members.⁶ From the 557 sources we find 523 to be members based on the combined information.

(iv) 138 sources have neither spectroscopy nor membership probabilities from Broos et al. (2013), and so are listed as non-members as they are currently unclassified (these are GES sources without spectroscopic measurements useful for membership determination).

⁶This also implies that ~ 5 percent of the stars classified by Broos et al. (2013) as members could be non-member background giants. Our parallax cut (below) will have removed the majority of these from our sample but it is worth noting for future studies that use the Broos et al. (2013) membership lists.

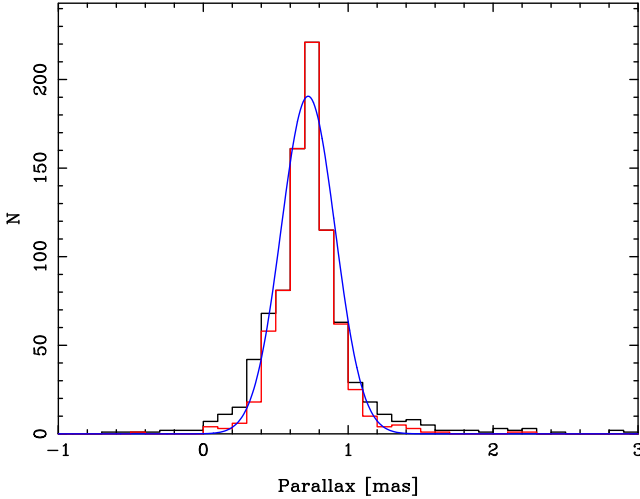


Figure 4. Parallax distribution for 896 candidate members of NGC 6530 with parallaxes, selected from GES and MYStIX candidate source lists (black histogram). The blue line shows a Gaussian fit to the distribution with centre 0.724 ± 0.006 mas and standard deviation 0.186 ± 0.024 mas. The red histogram shows the distribution of 761 members that pass our parallax cut (cut levels not shown as they depend on the individual parallax errors), as described in the text.

This process left us with 2014 candidate members. As a last step we implemented a parallax cut using data from *Gaia* DR2. Of the 2014 candidate members, 896 have astrometry from *Gaia*, as shown in Fig. 4, with a clear peak at 0.724 ± 0.006 mas. We estimate the additional systematic error to be 0.041 mas based on the spatial distribution of our sources and the table of spatial covariances provided for *Gaia* DR2. This corresponds to a distance of 1326^{+77}_{-69} pc (accounting for the -0.03 mas zero-point offset in the *Gaia* parallaxes). This is in good agreement with most previous estimates (e.g. Prisinzano et al. 2005), though it is significantly larger than the *Hipparcos*-based estimate of 600 pc from Loktin & Beshenov (2001, though this was based on only seven stars) and smaller than the photometric estimate of 1.8 kpc by Sung et al. (2000). The parallax distribution is reasonably modelled by a Gaussian with a standard deviation of 0.186 ± 0.024 mas. Accounting for and subtracting the contribution of the non-uniform parallax errors using the method of Ivezić et al. (2014), this leaves a parallax dispersion 0.084 mas. At a distance of 1326 pc this is equivalent to ~ 125 pc, though this is unlikely to represent the true physical dispersion, but instead may be due to underestimated or spatially correlated uncertainties. Based on this we exclude as members all stars with parallaxes that have $|\varpi - 0.724| > 2\sqrt{\sigma_{\varpi}^2 + 0.084^2}$, where ϖ and σ_{ϖ} are the measured parallax and its uncertainty. A cut at 2σ was chosen to balance the need to include the majority of members whilst rejecting a reasonable number of non-members. This cut excluded 135 sources from our membership list, leaving 761 members with parallaxes (shown in Fig. 4) and 1118 members without parallaxes (either too faint to be detected by *Gaia* or with poor astrometry).

Our final act is to remove all stars that lack both RVs and PMs (predominantly faint near-IR sources from Broos et al. 2013), leaving 889 sources for our kinematic study.

3.5 Final sample of members

In our final sample, 889 members are shown in Fig. 2 in a G versus $G - J$ CMD, which shows that the sample falls broadly between

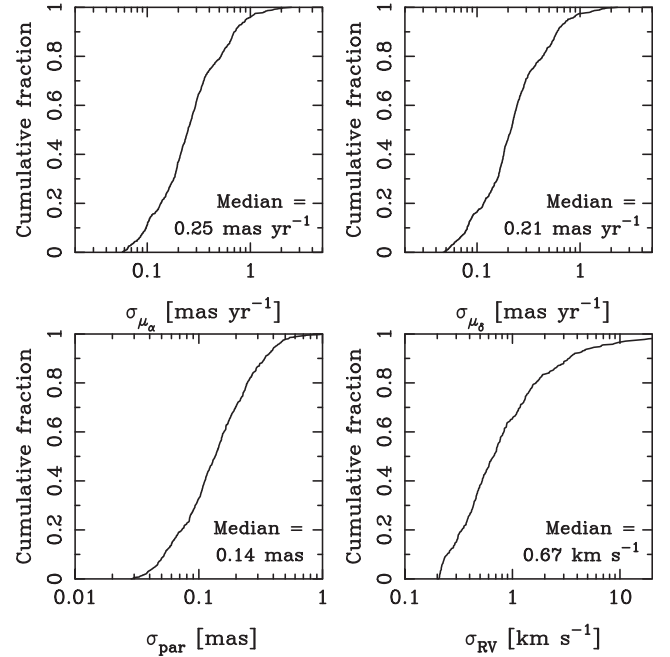


Figure 5. Cumulative uncertainty distributions for our sample of members for PMs (761 stars), parallax (761 stars), and RVs (404 stars).

the 0.1 and 10 Myr pre-main sequence stellar isochrones, as found by previous studies (e.g. Prisinzano et al. 2007), and as such there is no need to perform additional photometric filtering of the data. We calculated individual stellar masses for the stars in our sample by comparing the available photometry to Marigo et al. (2017) stellar isochrones, assuming a distance of 1326 pc and an extinction of $A_V = 1.1$ mag (Sung et al. 2000).⁷ The mass distribution of sources in our sample increases from ~ 20 to $\sim 0.7 M_{\odot}$ where it appears to turn over, with the lowest mass stars in our sample having masses of $0.1\text{--}0.2 M_{\odot}$. The majority of stars in our sample have masses of $0.4\text{--}1.0 M_{\odot}$.

Of the 889 members, 404 have RVs from GES or the literature and 761 have PMs from *Gaia* DR2, with 276 sources having both RVs and PMs providing 3D kinematics. Fig. 5 shows the cumulative uncertainty distributions for RV, PM, and parallax. The median RV uncertainty is 0.67 km s^{-1} , with some measurements with a precision as low as 0.25 km s^{-1} . The median PM uncertainties are 0.25 and 0.21 mas yr^{-1} in RA and Dec. (after adjusting for the underestimated PM uncertainties), respectively, which at a distance of 1326 pc are equivalent to 1.5 and 1.3 km s^{-1} .

Our full catalogue of members, including their photometry, astrometry, and spectroscopic parameters, is included in an online table made available on Vizier.

4 KINEMATIC ANALYSIS

Here we use the assembled kinematic data to study the dynamics of young stars in NGC 6530 and across the surrounding Lagoon Nebula. Fig. 6 gives an overview of the kinematics of young stars across the region. It is immediately clear that while both the RVs and PMs appear generally random within the RV range and PM

⁷Note that Prisinzano et al. (2019) find that most cluster members share a similar reddening so it is not unreasonable to assume a single extinction value for estimating stellar masses.

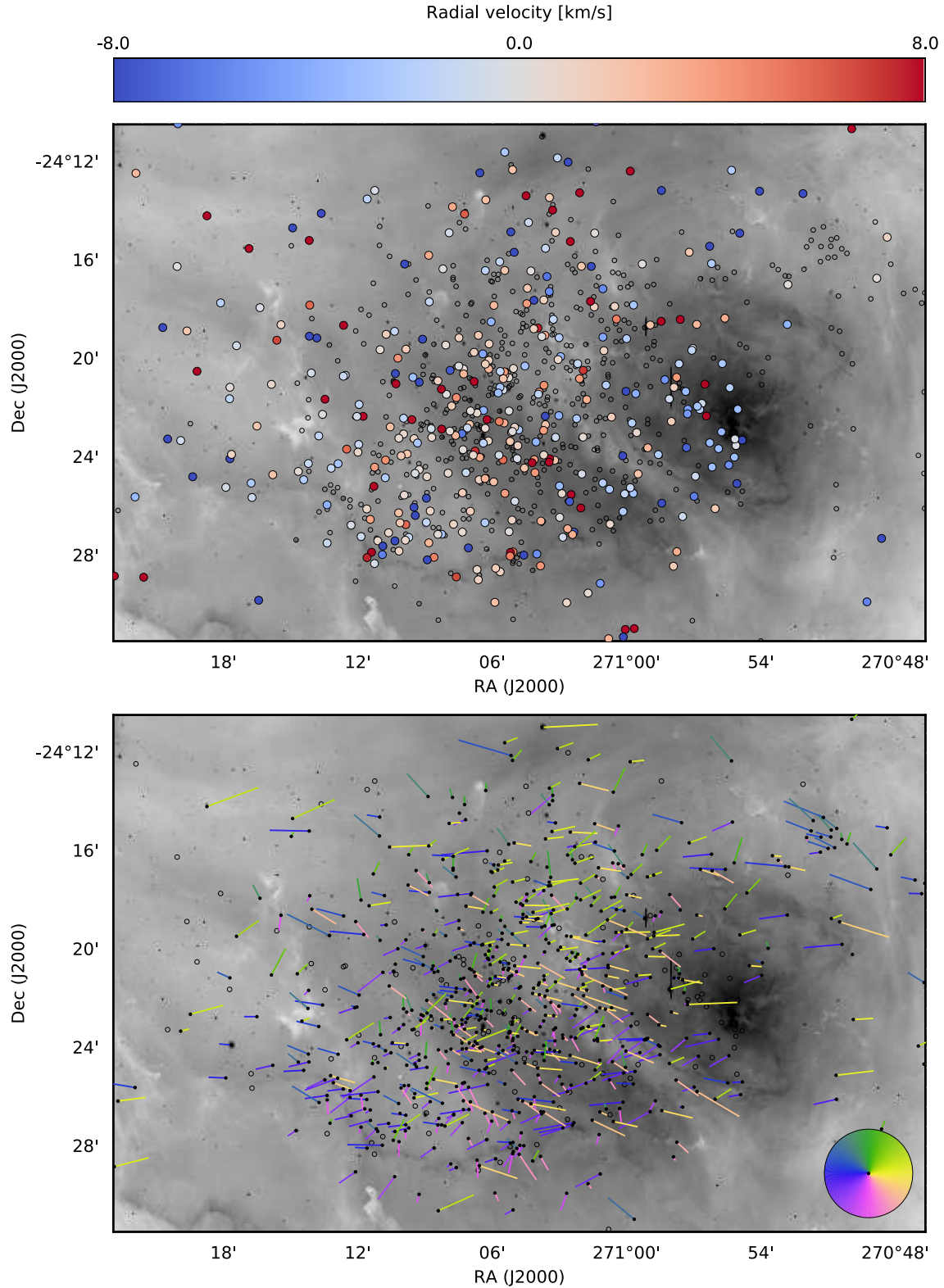


Figure 6. Kinematics of our sample of 889 young stars across the Lagoon Nebula projected on to an inverted VPHAS + H α image. *Top:* Spatial distribution of 404 sources colour-coded by their RV in the range $-8 < \text{RV}/\text{km s}^{-1} < 8$ (sources outside this range are coloured by the limiting values of the colour bar). The 485 sources without RVs are shown as empty circles. *Bottom:* Spatial distribution of 761 sources with vectors showing their PMs over 0.05 Myr, relative to the median PM of $(\mu_\alpha, \mu_\delta) = (1.27, -1.98) \text{ mas yr}^{-1}$. The PM vectors are colour-coded by their PM position angle, as indicated by the colour wheel in the lower right corner. The 128 sources without usable PMs are shown as empty circles.

reference frame shown, there is some evidence for substructure. This is evident on both small scales, where small and localized groups of stars have similar velocities to each other (for example the north-western group highlighted in Fig. 1), and on larger scales where hints of velocity gradients are evident particularly in the PM vector map. There is also evidence from Fig. 6 that the PMs have a preference for east–west motion, as evidenced by their colours. At first glance this might suggest evidence for expansion of the population in this direction, though as we will show later the expansion is primarily in the north–south direction (Section 4.3) and that the pattern seen here is primarily due to the larger velocity dispersion in the east–west direction (Section 4.4.2).

4.1 Kinematic outliers

A number of stars in the Lagoon Nebula have velocities that are significantly different from the general population, which we refer to as *kinematic outliers*. These stars may have recently been ejected from the region due to dynamical interactions or they may be contaminating non-members, but regardless, they usually do not probe the kinematics of the main population that we are interested in.

To identify these objects we calculate simple estimates of the central velocity and dispersion in each dimension using the median velocity and inter-quartile range (IQR), which is related to the standard deviation of a normal distribution as $\sigma = 0.741 \text{ IQR}$. When the velocity distributions do not follow normal distributions the IQR provides a useful, outlier-resistant method for estimating the velocity dispersion of the bulk of a population. From this we calculate central (median) velocities of $\text{RV}_0 = 0.37 \text{ km s}^{-1}$ and $(\mu_\alpha, \mu_\delta)_0 = (1.27, -1.98) \text{ mas yr}^{-1}$ and approximate velocity dispersions⁸ of $\sigma_{\text{RV}} = 2.89 \text{ km s}^{-1}$ and $(\sigma_{\mu_\alpha}, \sigma_{\mu_\delta}) = (0.705, 0.446) \text{ mas yr}^{-1}$. Stars that fall more than 3σ from the median velocity are identified as kinematic outliers.⁹ This gives 95 RV outliers and 66 PM outliers (35 and 47 in the two PM directions, and 16 being outliers in both PM directions), significantly more than would be expected from a normal distribution.

Since single-epoch RVs are heavily affected by binarity, the RV outliers may be dominated by true members of the Lagoon Nebula population that are in close binary systems. PMs are not affected by binarity to such a degree and therefore these outliers are expected to either not be part of the Lagoon Nebula population or to have recently been ejected from the region in some sort of dynamical interaction. The spatial distribution of the PM outliers shows no evidence for being clustered like the majority of our sources are, while their distribution in the CMD (Fig. 2) shows a slight preference for occupying the blue-edge of the distribution, suggesting that they might not be members of the Lagoon Nebula population. Based on this, and since both non-members and recently ejected sources will not help us resolve the dynamical state of the Lagoon Nebula, we have not included the 66 PM outliers in the kinematic analysis that follows, but have retained the RV outliers

⁸The velocity dispersion σ_{μ_δ} calculated here is smaller than that derived from a full model in Section 4.4.2, despite the latter accounting for the effects of uncertainties. This is because the IQR provides only a simplistic proxy to the dispersion, but this is sufficient to identify kinematic outliers.

⁹Note that when this PM cut is transferred to physical velocity space it will include a dependence on distance and causing it to vary if there is a large spread in distance. However, the resulting change in velocity thresholds is smaller than velocity uncertainties derived from the PMs, and thus not significant.

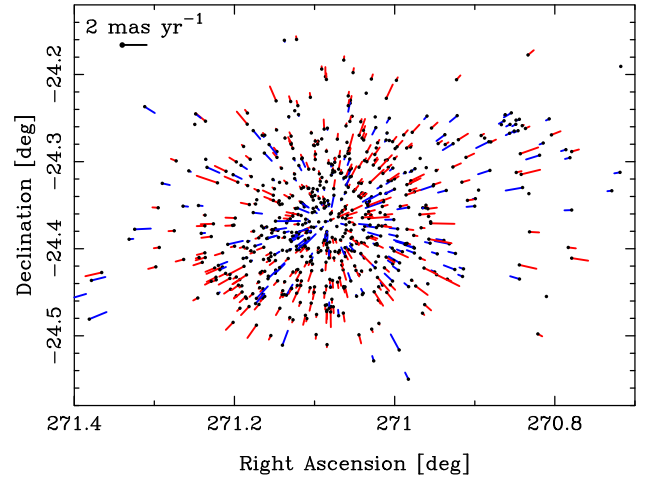


Figure 7. PM vector map showing the radial component of the PMs relative to the centre (median position) of our sample. The dots show the current positions of the stars while the vectors show the PM over 0.05 Myr colour-coded red if the stars are moving outwards and blue if they are moving inwards.

(this may introduce a bias to the RV distribution, but this is likely to be a very small effect).

4.2 Expansion and rotation

To study the large-scale dynamics of the Lagoon Nebula population we first search for evidence of expansion, contraction and rotation of the entire population using the method of Wright et al. (2016). To do this we separate the PMs of our sources into their radial and transverse components using the median positions of stars in the population, $(\alpha, \delta) = (271.080, -24.369) \text{ deg}$. Using the stellar masses calculated in Section 3.5 we find that the kinetic energy in the PMs is split approximately equally between the radial and transverse directions with the former constituting 51 ± 2 (statistical) ± 3 (positional) per cent of the energy.¹⁰

The transverse component of the PMs shows a preference for rotation with 65.2 ± 2.4 (stat.) ± 4.3 (pos.) per cent of the kinetic energy in the clockwise direction. The radial component of the PMs show a significant preference for expansion however, with 76.3 ± 1.7 (stat.) ± 6.7 (pos.) per cent of the kinetic energy in the form of expansion. This is in agreement with Kuhn et al. (2019) who found that the PM velocity vectors in their sample were primarily orientated away from the cluster centre. Fig. 7 shows the radial component of the PM vectors colour-coded as to whether they are moving towards (in blue) or away (in red) from the centre of the region. The preference for expansion is just discernible in this figure as a preference for vectors coloured red rather than blue (note that this figure highlights vector length or PM rather than kinetic energy).

4.3 Velocity gradients

To further explore the possible expansion of the Lagoon Nebula population we searched for velocity gradients in our kinematic data.

¹⁰The statistical uncertainty is calculated from a Monte Carlo experiment varying the masses and PMs of all sources by their uncertainties, while the positional uncertainty is calculated by varying the centre used by $\pm 0.1 \text{ deg}$ in each dimension.

Table 2. Velocity gradient fit results for all three kinematic axes as a function of RA and Declination (the line-of-sight spatial distribution within the region is not resolved so we do not fit against distance). The conversion to physical units was calculated using our parallax-derived distance of 1326 pc.

Velocity	Versus RA	Versus Dec.	Units
μ_α	0.00 ± 0.21 0.00 ± 0.057	-1.17 ± 0.33 -0.317 ± 0.090	$\text{mas yr}^{-1} \text{ deg}^{-1}$ $\text{km s}^{-1} \text{ pc}^{-1}$
μ_δ	-0.02 ± 0.14 -0.004 ± 0.038	2.74 ± 0.21 0.744 ± 0.059	$\text{mas yr}^{-1} \text{ deg}^{-1}$ $\text{km s}^{-1} \text{ pc}^{-1}$
RV	0.92 ± 8.3 0.041 ± 0.37	-19.0 ± 12.3 -0.82 ± 0.55	$\text{km s}^{-1} \text{ deg}^{-1}$ $\text{km s}^{-1} \text{ pc}^{-1}$

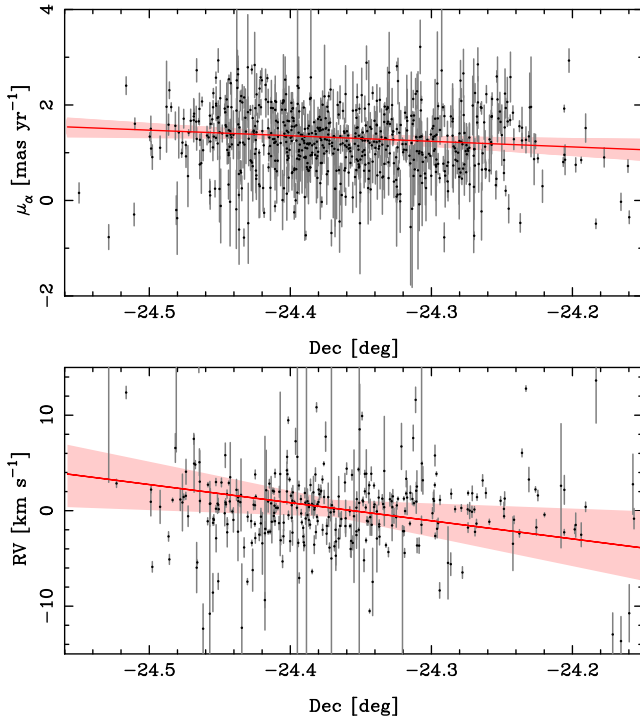


Figure 8. Velocity gradients that probe significant levels of rotation. The top panel shows μ_α plotted against declination and the bottom panel shows RV plotted against declination (error bars also shown). The red lines show the best-fitting velocity gradients of $-1.17 \pm 0.33 \text{ mas yr}^{-1} \text{ deg}^{-1}$ (μ_α) and $-19.0 \pm 12.3 \text{ km s}^{-1} \text{ deg}^{-1}$ (RV), with the light red bands showing the range of slopes within 1σ of the best-fitting value.

Velocity gradients have been observed in multiple star forming regions (e.g. Rigliaco et al. 2016; Sacco et al. 2017), as well as in dense gas tracers (André et al. 2007), and can be attributed to processes such as rotation or expansion. We searched for velocity gradients by fitting linear relationships of the form $v = Ax + B$ between the velocity, v , and spatial position, x , in each combination of dimensions. The gradient, A , and zero-point, B , were fitted by maximizing the likelihood function, using the Markov-Chain Monte Carlo (MCMC) ensemble sampler *emcee* to sample the posterior distribution. A third parameter (f) was introduced to represent the scatter in the relationship (see Hogg, Bovy & Lang 2010) and which was marginalized over to calculate the fit and uncertainties on the fitted velocity gradients.

The results of the velocity gradient fits are listed in Table 2 and can be divided into those that reveal expansion or contraction (a velocity gradient where the velocity is correlated with the position

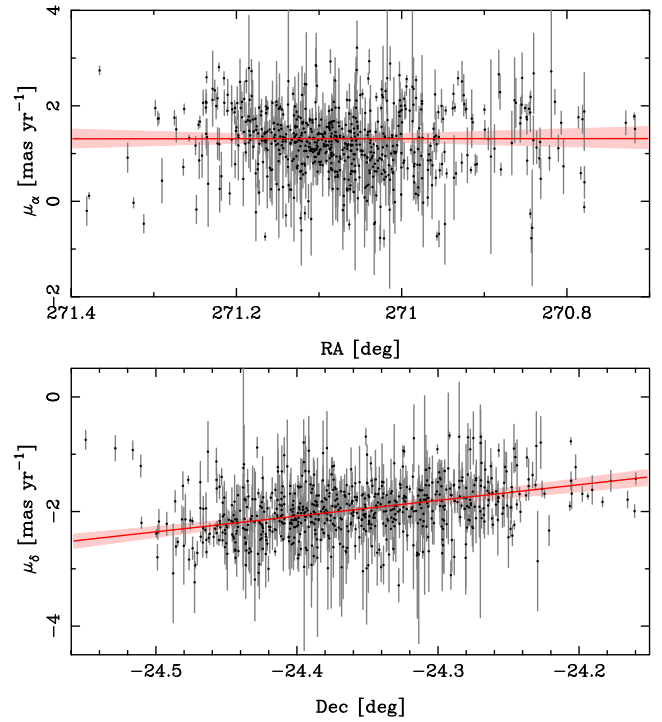


Figure 9. Velocity gradients that probe expansion and contraction. The top panel shows the PM in the right ascension direction plotted against the right ascension and the bottom panel shows the PM in the declination direction plotted against the declination (error bars also shown). The red lines show the best-fitting velocity gradients of 0.0 ± 0.21 and $2.74 \pm 0.21 \text{ mas yr}^{-1} \text{ deg}^{-1}$ in right ascensions and declination, respectively, with the light red bands showing the range of slopes within 1σ of the best-fitting value.

in the same dimension) and those that highlight some degree of rotation or residual angular momentum (where a velocity in one dimension is correlated with position in another dimension). Fig. 8 shows that there is evidence for some degree of rotation or residual angular momentum within the population, particularly in μ_α as a function of declination ($>3\sigma$ significance), with weak evidence for a correlation between RV and declination ($\sim 1.5\sigma$ significance). The correlation between μ_α and declination is consistent with the hint of clockwise rotation found in Section 4.2.

Fig. 9 shows the evidence for velocity gradients indicative of expansion or contraction. Notably there is very strong evidence of a clear correlation between μ_δ and declination with a slope of $2.74 \pm 0.21 \text{ mas yr}^{-1} \text{ deg}^{-1}$ (significance $>10\sigma$), but no evidence for a correlation between μ_α and RA, for which we fit a slope of $0.0 \pm 0.21 \text{ mas yr}^{-1} \text{ deg}^{-1}$. The velocity gradient in declination has a positive slope, meaning that stars on the northern (southern) side of the nebula have slightly higher (lower) PMs in declination, which implies expansion, in agreement with the results of Section 4.2. The lack of a velocity gradient in RA does not change if we limit ourselves to a sample of stars with *astrometric_excess_noise* = 0, for which we obtain a fit of $0.64 \pm 0.75 \text{ mas yr}^{-1} \text{ deg}^{-1}$.

Our results show an interesting disagreement with those presented by Kuhn et al. (2019). Those authors show a correlation between the median plane-of-sky expansion velocity and the radius from the cluster centre, with a slope of $0.6 \pm 0.2 \text{ km s}^{-1} \text{ pc}^{-1}$. They argue that this is evidence for a radially dependent expansion velocity showing a ‘Hubble flow’ like expansion pattern. However, we find that the correlation between velocity and position only exists in declination and not in RA. Kuhn et al. (2019) do claim to observe correlations between both μ_α and RA and between μ_δ

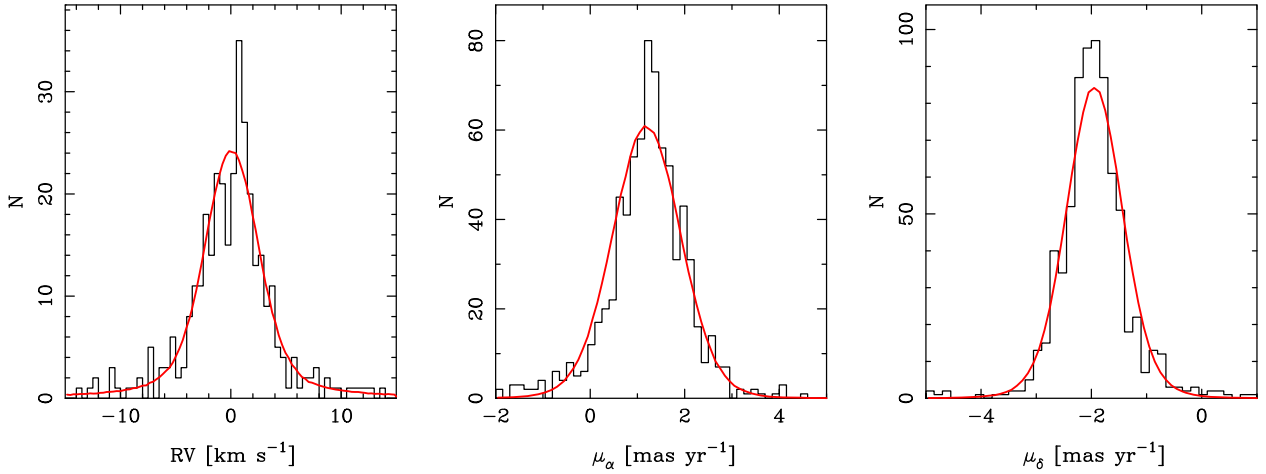


Figure 10. Velocity distributions (black histograms) for our sample of young stars in RVs and PMs compared to the best-fitting velocity dispersion model fit results (red distributions). The histograms show 337 out of 404 stars with RVs and 743 (μ_α) and 740 (μ_δ) out of 761 stars with PMs (the remainder, in all cases, fall outside of the plotted range). The model fit distributions were calculated by sampling the best-fitting forward models 1000 000 times and plotting the resulting probability distribution functions scaled to the observations.

and declination, though they do not present fits to them and the correlation in RA is ambiguous. The differences between our work and theirs may be due to their smaller sample size, their more compact area studied, the presence of contaminants (identified from GES spectroscopy) or the different quality cuts applied to the data (including their use of cuts not recommended in *Gaia* data release papers or by DPAC). We have shown that the expansion pattern in the Lagoon Nebula is not fully radially dependent and instead shows a distinct asymmetry. This argues against a simple explosive expansion pattern, which we will discuss further in Section 5.

4.4 Velocity distribution of the entire population

Fig. 10 shows the distribution of measured velocities in RV and PM for the stars in our sample. While at first glance the velocity distributions may appear to follow a broadly normal distribution, there is evidence for deviations from this in the form of central concentrations in velocity space in all three dimensions and hints of a double peak in the RV distribution. Shapiro–Wilk tests of normality conducted on the velocity distributions confirm this with all three dimensions rejecting the null hypothesis of a normal distribution at greater than 3σ . We suggest in Section 5 that these deviations from normality may be due to a core-halo structure with the central NGC 6530 cluster (which is both spatially and kinematically more compact, see Section 4.5.2) responsible for the core of the velocity distribution and the wider Lagoon Nebula population making up the halo.¹¹ In this section however we start by considering the velocity distribution of the entire population.

4.4.1 Forward modelling method

We calculate the velocity dispersions of our sample using Bayesian inference by modelling the distribution of observed velocities using

¹¹ A core-halo velocity distribution could also arise from the combination of a narrow, intrinsic velocity dispersion and a broad distribution of measurement uncertainties. However, the presence of kinematic subgroups within the halo distribution (Section 4.5) argues against this.

a simple parametrized model that we compare with the observations in a probabilistic way (see e.g. Wright et al. 2016; Wright & Mamajek 2018). The aim of this process is to determine which of the various sets of parameters, θ , best explain the observations, d . In Bayes’s theorem this is known as the posterior distribution, $P(\theta|d) = P(d|\theta)P(\theta)/P(d)$, where $P(d|\theta)$ is the likelihood model, $P(\theta)$ are the priors (which includes our a priori knowledge about the model parameters) and $P(d)$ is a normalizing constant. Bayesian inference is necessary for this because it allows the model predictions to be projected into observational space, where the measurement uncertainties are defined. This is particularly important when the observations have correlated uncertainties, as is the case here. Our forward model begins by constructing a population of $N = 10^5$ stars with 3D velocities randomly sampled from a trivariate Gaussian velocity distribution along the radial and two transverse directions (the large size of the modelled population reduces the noise in the modelled velocity distributions).

Unresolved binary systems can broaden the observed RV distribution due to the contribution that binary orbital motion makes to the measured velocity. To simulate this process we follow Odenkirchen et al. (2002) and Cottaar et al. (2012) by assuming that a fraction of our sample are in binary systems (we do not consider triple systems because their properties are poorly constrained and are typically hierarchical, meaning that the third star is usually on a wide, long-period orbit that does not introduce a large velocity offset). The fraction of binary stars in young star clusters is poorly constrained and so we set the binary fraction to be 46 percent, appropriate for solar-type field stars (Raghavan et al. 2010). The primary star masses were sampled from the observed stellar masses, while the secondary masses were selected from a power-law companion mass ratio distribution with index $\gamma = -0.5$ over the range of mass ratios $q = 0.1$ – 1.0 (Reggiani & Meyer 2011). The orbital periods were selected from a lognormal distribution with a mean period of 5.03 and a dispersion of 2.28 in $\log_{10}$ days (Raghavan et al. 2010). The eccentricities were selected from a flat distribution from $e = 0$ to a maximum that scales with the orbital period as proposed by Parker & Goodwin (2009). We then calculate instantaneous velocity offsets for the primary and secondary stars (relative to the centre of mass of the system) at a random phase within the binary’s orbit, and then

Table 3. Central velocities and velocity dispersions for the entire Lagoon Nebula population and the subsets of NGC 6530, M8E, and the north-western subgroup. Velocity dispersions are not calculated for the latter two as they are not sufficiently well sampled. The north-western subgroup lacks any sources with spectroscopy and so its RV is not known.

Sample	RV_0 (km s ⁻¹)	σ_{RV} (km s ⁻¹)	μ_{α_0} (mas yr ⁻¹)	σ_{μ_α} (mas yr ⁻¹)	μ_{δ_0} (mas yr ⁻¹)	σ_{μ_δ} (mas yr ⁻¹)
Lagoon Nebula (all members)	$0.08^{+0.34}_{-0.30}$	$2.12^{+0.24}_{-0.22}$	1.19 ± 0.05	$0.644^{+0.044}_{-0.045}$	-1.94 ± 0.04	$0.443^{+0.036}_{-0.038}$
NGC 6530 ($r < 2.5$ arcmin)	$0.82^{+0.34}_{-0.34}$	$1.80^{+0.40}_{-0.37}$	1.21 ± 0.04	$0.251^{+0.045}_{-0.043}$	-2.00 ± 0.03	$0.216^{+0.038}_{-0.036}$
M8E ($r < 2$ arcmin)	0.75 ± 0.72	–	1.78 ± 0.23	–	-2.14 ± 0.34	–
North-western group ($r < 1$ arcmin)	–	–	1.77 ± 0.20	–	-1.50 ± 0.17	–

use the luminosity-weighted average of the two as the photocentre RV_{12} which is then added to the modelled RV .

Finally we add measurement uncertainties for the RVs and PMs for each star, randomly sampling these from the observed uncertainty distributions and include the correlated PM uncertainties quoted in *Gaia* DR2 (random sampling of uncertainties is necessary because the modelled population is significantly larger than the observed population). This model has 6 free parameters, the velocity dispersion and central velocity in each dimension. Wide and uniform priors were used for each of these parameters covering 0–100 km s⁻¹ for the velocity dispersions and –100 to +100 km s⁻¹ for the central velocities.

To sample the posterior distribution function we use the affine-invariant MCMC ensemble sampler (Goodman & Weare 2010) EMCEE (Foreman-Mackey et al. 2013) and compared the model to the observations using an unbinned maximum-likelihood test, which is made efficient by the smooth velocity distributions modelled. For the MCMC sampler we used 1000 walkers and 2000 iterations, discarding the first half as a burn-in. The six parameters were found to have similar autocorrelation lengths, with the longest being σ_{μ_α} with a length of 104 iterations, resulting in ~ 20 independent samples per walker. The posterior distribution functions follow a normal distribution, and thus the median value was used as the best fit, with the 16th and 84th percentiles used for the 1σ uncertainties.

4.4.2 Results

The results are illustrated in Fig. 10. The best-fitting velocity dispersions are $\sigma_{RV} = 2.12^{+0.24}_{-0.22}$ km s⁻¹ and $(\sigma_{\mu_\alpha}, \sigma_{\mu_\delta}) = (0.644^{+0.044}_{-0.045}, 0.443^{+0.036}_{-0.038})$ mas yr⁻¹ (full fit details in Table 3), which at a distance of 1326 pc equate to $4.06^{+0.37}_{-0.36}$ and $2.79^{+0.29}_{-0.28}$ km s⁻¹ (uncertainties take into account full distance uncertainties). The three-velocity distributions are significantly different from each other, particularly the RV and μ_α dispersions, which imply anisotropy with a confidence of 2.3σ (based on the difference between the RV and μ_α dispersions). The binary fraction of young stars is very poorly constrained and thus the RV dispersion could be underestimated if we have overestimated the binary fraction. However, this effect is not large enough to increase the RV dispersion to that of the PMs – setting the binary fraction to 0 per cent results in a RV dispersion of $2.47^{+0.29}_{-0.24}$ km s⁻¹, still significantly smaller than the μ_α dispersion.

¹²This process compensates for the fact that for high-mass-ratio binaries some of the light from the secondary will contribute to the spectral features used to measure the RV, and thus the measured RV will be intermediate between that of the two stars. From simulations we found that this reduces the broadening of the RV distribution due to binaries by ~ 25 per cent.

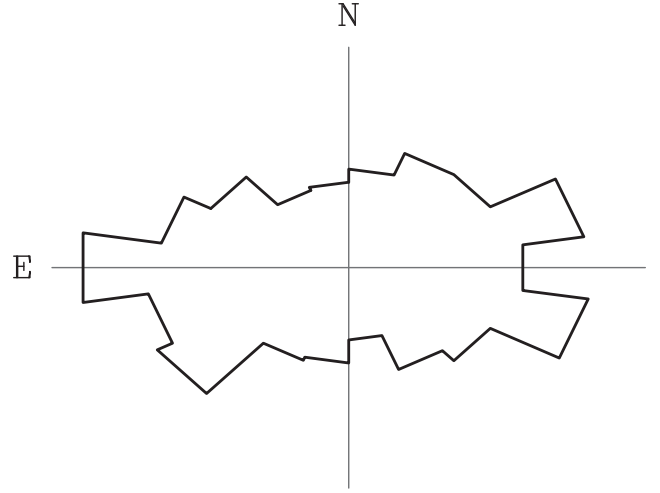


Figure 11. PM position angle distribution plotted as a circular histogram with north up and east to the left.

The PM velocity dispersions are also significantly different from each other and provide equally strong evidence for anisotropy (with a significance of 2.2σ between them). Fig. 11 shows the distribution of PM vector position angles on the plane of the sky, which shows that the PM velocity distribution is ellipsoidal, with a semimajor axis closely aligned with the right ascension axis (the larger velocity distribution in this direction goes a long way to explaining the preference for east–west motion as shown in Fig. 6). The velocity anisotropy implies the region is not sufficiently dynamically mixed to have developed an isotropic velocity dispersion. Such anisotropy has been observed in other regions, particularly OB associations, which are believed to be dynamically un-evolved (Wright et al. 2016; Wright & Mamajek 2018). In OB associations the anisotropy can often be attributed to kinematic substructure within the region, which may be the case in the Lagoon Nebula (see Section 4.5).

Our PM velocity dispersions are larger than those reported by Kuhn et al. (2019) of 2.7 ± 0.2 and 1.8 ± 0.2 km s⁻¹. This is because their sample covers a smaller area on the sky than ours, being limited to the area of the central *Chandra* observations included in the MYStIX survey, and so is dominated by the core of the velocity distribution that originates from the central NGC 6530 cluster (see Section 4.5.2). Our PM velocity dispersions are significantly smaller than the 0.85 ± 0.06 mas yr⁻¹ velocity dispersion measured by Chen, de Grijs & Zhao (2007)¹³ from photographic plates. This

¹³The authors only quote a 3D velocity dispersion of 1.48 ± 0.14 mas yr⁻¹, from which we derive an estimate of their measured 1D velocity dispersion.

could either be because the two samples have different membership or because the photographic plates may suffer from distortions.

Our combined three-dimensional velocity dispersion was calculated as $\sigma_{3D} = \sqrt{\sigma_{RV}^2 + \sigma_{\mu_\alpha}^2 + \sigma_{\mu_\delta}^2} = 5.35_{-0.34}^{+0.39}$ km s⁻¹.

4.4.3 Virial state

To measure the virial state of the Lagoon Nebula stellar population we use the virial equation, which in its three-dimensional form is given by

$$\sigma_{3D}^2 = \frac{GM_{vir}}{2r_{vir}}, \quad (1)$$

where σ_{3D} is the three-dimensional velocity dispersion, G is the gravitational constant, M_{vir} is the virial mass, and r_{vir} is the virial radius. We substitute the parameter $\eta = 6r_{vir}/r_{eff}$, where r_{eff} is the effective (or half-light) radius, and rearrange to give the virial mass as

$$M_{vir} = \eta \frac{\sigma_{3D}^2 r_{eff}}{3G}, \quad (2)$$

where the parameters η and r_{eff} are determined by fitting an Elson, Fall & Freeman (1987, hereafter EFF) surface brightness profile to the stellar distribution. We do this with our sample and find parameters of $\eta = 9.3 \pm 0.1$ and $r_{eff} = 4.83_{-0.48}^{+0.73}$ arcmin, which equates to $1.87_{-0.30}^{+0.41}$ pc at a distance of 1326 pc.

This gives a virial mass of $3.8_{-0.7}^{+1.0} \times 10^4 M_\odot$. The total stellar mass in the Lagoon Nebula has been estimated to be between 1000 (Prisinzano et al. 2005; Tothill et al. 2008) and 4000 $M_\odot$ (Kuhn et al. 2015). Based on our full sample of likely members, which we estimate to be complete in the range of 0.7–3.0 $M_\odot$, and extrapolating out using a Maschberger (2013) initial mass function, we estimate the total mass to be $\sim 2500 M_\odot$. The mass of the Lagoon Nebula H II region and parental molecular cloud in the immediate vicinity of the stellar population are not well constrained, though the mass of the molecular clouds in the wider area have been estimated to be $2\text{--}6 \times 10^4 M_\odot$ (Takeuchi et al. 2010). Such a mass of material would be sufficient to keep the stellar system in virial equilibrium, but since the stellar population is not embedded within these molecular clouds (as evidenced by the relatively low extinction towards most members) this material will not contribute to the virial state of the system. We therefore conclude that the stellar population within the Lagoon Nebula is gravitationally unbound and should continue to expand in the future.

4.5 Velocity distribution of subgroups within the Lagoon Nebula

The velocity dispersion calculations above consider stars across the Lagoon Nebula, including the central NGC 6530 cluster, a number of known subgroups, and a wider less-clustered population (see Fig. 1). If these substructures represent real clusters they may be kinematically distinct from the rest of the Lagoon Nebula population and could be gravitationally bound, even if the wider population is not. We therefore considered the kinematics of four groups of stars within the Lagoon Nebula complex, three that are commonly known (NGC 6530, M8E, and the Hourglass Nebula cluster) and one that is clear from the spatial distribution of X-ray sources and was first identified by Kuhn et al. (2014). We selected stars in each of these groups based on the regions shown in Fig. 1, which were themselves based on the central positions identified by Kuhn et al. (2014, with the exception of NGC 6530, see below). For each group we then

use a two-sample Anderson–Darling (AD) test and the tables of Scholz & Stephens (1987) to verify the null hypothesis that the stars in each group have kinematics drawn from the same distribution as the wider population.

4.5.1 M8E, the Hourglass Nebula, and the north-western subgroup

For the M8E subgroup we selected 38 stars within a radius of 2 arcmin. These stars have similar RVs and μ_δ velocities to the wider population, but have a preference for more eastward PMs relative to stars across the rest of the nebula (as evident from Fig. 6). AD tests confirm this, providing a p -value of 0.0004 for μ_α and allowing us to reject the null hypothesis. The M8E subgroup therefore appears kinematically distinct.

For the Hourglass Nebula we selected six stars within a radius of 1 arcmin, but found that their kinematics were similar to that of the main population. The AD tests do not provide significant evidence that the velocities of stars are not drawn from that of the wider distribution. Since the Hourglass Nebula cluster is highly embedded, our optically selected sample probably only includes sources projected against the cluster and therefore may not be representative of the Hourglass Nebula population.

We identify 14 stars as likely members of the north-western subgroup, all within a radius of 1 arcmin. The majority of these stars have PMs in the north-easterly direction (as seen in Fig. 6) that marks them out as kinematically distinct (though sparsely sampled). AD tests provide support for this, with p -values of 0.0031 (μ_α), and 0.0023 (μ_δ) implying that the null hypothesis can be rejected and that the group is kinematically distinct.

The central (median) velocities of these subgroups are listed in Table 3 (with the exception of the Hourglass Nebula population, which we believe has not been properly sampled), though we do not calculate velocity dispersions as the groups are too sparsely sampled.

4.5.2 The central NGC 6530 cluster and its virial state

The NGC 6530 cluster sits at the centre of the Lagoon Nebula and we selected a sample of 150 stars in a region with a radius of 150 arcsec centred on (RA, Dec.) = (271.105, −24.371) deg. The kinematics of these stars exhibit a narrower velocity distribution than that of the entire population (see Fig. 12), particularly in μ_α , and are offset towards slightly larger RVs. AD tests provide strong evidence for this with p -values of 0.036 (RV), 0.00084 (μ_α), and 0.019 (μ_δ) that allow us to reject the null hypothesis that the kinematics of these stars are drawn from the same distribution as the wider population. Shapiro–Wilk tests of normality conducted on the velocity distributions provide no evidence that the PM distributions are inconsistent with following normal distributions, while the RV distribution does deviate significantly from normality, most likely due to the presence of binaries. The three spatial subclusters within the NGC 6530 cluster that were identified by Kuhn et al. (2014) appear to be kinematically indistinct based on our data, with no evidence from the AD tests that they are not drawn from the same velocity distributions as each other, so we do not subdivide our sample any further. We fit the velocity dispersion of this subset using the same method as for the entire sample.

The results are shown in Fig. 12 and listed in Table 3. The best-fitting velocity dispersions are $\sigma_{RV} = 1.80_{-0.37}^{+0.40}$ km s⁻¹ and

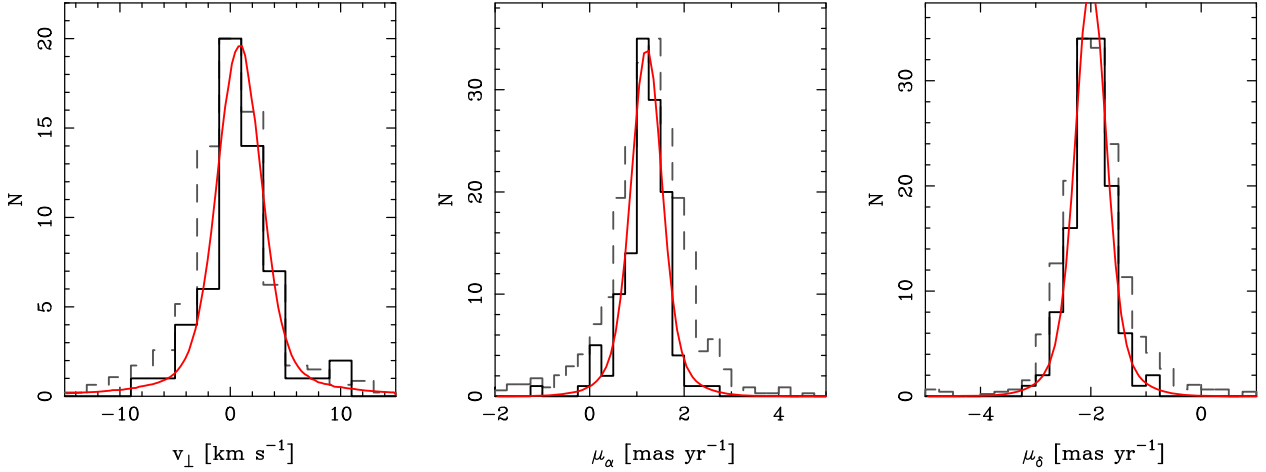


Figure 12. Velocity distributions (black histograms) for a sample of 150 young stars in the NGC 6530 cluster (within a radius of 150 arcsec) in RV and PMs. The histograms show 57 out of 60 stars with RVs (the remainder fall outside the plotted range) and all 126 stars with PMs. In red we show the best-fitting velocity dispersion model fits for the NGC 6530 cluster population. The model fit distributions were calculated by sampling the best-fitting forward model 1000 000 times and plotting the resulting probability distribution functions scaled to the observations. The velocity ranges plotted are the same as those in Fig. 10 and we also show for comparison the velocity distributions for all stars across the Lagoon Nebula (grey dashed histograms) as shown in that figure.

$(\sigma_{\mu_\alpha}, \sigma_{\mu_\delta}) = (0.251^{+0.045}_{-0.043}, 0.216^{+0.038}_{-0.036})$ mas yr⁻¹, which at a distance of 1326 pc equate to $1.58^{+0.31}_{-0.29}$ and $1.36^{+0.26}_{-0.24}$ km s⁻¹ (uncertainties take into account the distance uncertainties). Contrary to the whole sample these three velocity dispersions are in agreement with each other within the uncertainties (significance of anisotropy $< 1\sigma$), suggesting an isotropic velocity dispersion that implies the cluster has been sufficiently mixed to erase any primordial anisotropy. The 3D velocity dispersion is $2.75^{+0.39}_{-0.28}$ km s⁻¹.

We fit an EBF profile in the same manner described earlier and find an effective radius of $r_{\text{eff}} = 1.83^{+0.42}_{-0.35}$ arcmin or $0.71^{+0.20}_{-0.17}$ pc at a distance of 1326 pc, and $\eta = 9.3 \pm 0.1$. Combined with the 3D velocity dispersion this equates to a virial mass of $3800^{+1100}_{-900} M_\odot$. Extrapolating the stellar population to a full initial mass function we estimate the stellar mass of the central NGC 6530 cluster to be $480 \pm 100 M_\odot$, which is an order of magnitude lower than the virial mass. Therefore, despite the central cluster having an isotropic velocity dispersion that would imply it is dynamically mixed, the cluster is not in virial equilibrium and should disperse in the future.

5 DISCUSSION

Here we discuss our results and their implications for understanding young star clusters. Our results can briefly be summarized as follows.

(i) There is strong evidence for expansion of the Lagoon Nebula stellar population, with 76 per cent of the kinetic energy in the radial component of the PMs in the direction of expansion. However this expansion is preferentially in the declination direction, as evidenced by the strong velocity gradient between μ_δ and declination, and there is no evidence for a systematic expansion pattern in the right ascension direction.

(ii) The 3D velocity distributions are non-Gaussian, with a strong central peak due to the NGC 6530 cluster, and the velocity dispersions are anisotropic, with the velocity dispersion in the right ascension direction being significantly larger than that in the declination direction. The 3D velocity dispersion is $5.35^{+0.39}_{-0.34}$ km s⁻¹, giving a virial mass of $\sim 38\,000 M_\odot$, an order of magnitude larger

than the known stellar mass, implying that the region is globally unbound.

(iii) The central cluster NGC 6530 has a 3D velocity dispersion of $2.75^{+0.39}_{-0.28}$ km s⁻¹, which gives a virial mass of $\sim 3800 M_\odot$, also about an order of magnitude larger than the known stellar mass and implying the cluster is unbound. The velocity dispersion is consistent with isotropy and the PM distributions follow a normal distribution, suggesting the cluster is dynamically mixed.

The broad picture that the kinematics give us is that the stellar population is anisotropic, gravitationally unbound and expanding, though this expansion is not isotropic and (at least within the plane of the sky) is preferentially occurring in the declination direction (though since the region is gravitationally unbound it should expand in the future in both directions). We discuss the two main results here.

5.1 Current and past dynamical states

The non-normal velocity distributions show a central peak that appears to be due to the kinematically distinct NGC 6530 cluster in the centre of the Lagoon Nebula, and potentially other kinematic substructures that are less-well sampled. This core-halo velocity pattern, where the core cluster exhibits a narrower velocity distribution than the wider halo population, has been observed in other young regions (e.g. Jeffries et al. 2014). In this scenario the denser NGC 6530 cluster would have formed within a larger, lower density and unbound region that we now see as the Lagoon Nebula stellar population. The dispersal of the residual gas left over from star formation may have affected the dynamics of both groups of stars, particularly if either was gravitationally bound.

The velocity dispersions of both the entire stellar population and the central NGC 6530 cluster are sufficiently large that they appear to be gravitationally unbound at the moment. It is not clear whether the entire region or the NGC 6530 cluster was ever gravitationally bound in the past. Gravitationally bound stellar groups tend to be well mixed, as the necessary high stellar density for virial equilibrium leads to a short dynamical time-scale. The central NGC 6530 cluster does have an isotropic and Gaussian velocity

dispersions that suggests it is well mixed and therefore may have been a bound cluster in the past.

For the wider Lagoon Nebula population the non-Gaussian velocity distributions and clear anisotropy suggest that the system has not been sufficiently mixed and therefore was not a bound cluster in the past. The semimajor axis of the anisotropic velocity distribution on the plane of the sky is orientated almost east–west, roughly aligned with the spatial elongation of the nebula itself, and, to a lesser extent, the stellar system. This may be an indication that the velocity distribution and anisotropy are primordial and have not been well mixed.

Some N -body simulations predict that velocity anisotropy can arise in star clusters following violent relaxation in a tidal field (Vesperini et al. 2014). However the main signal of this anisotropy is predicted to occur at radii of between 2 and 5 half-mass radii, significantly beyond the area covered by our current sample (which extends to approximately 2 and 3 half-mass radii in declination and right ascension, respectively), and that the inner regions should be well mixed after such events. These simulations also model the initial distribution of stars as simple spherical and centrally concentrated systems, whereas young star forming regions are considerably more substructured (e.g. Cartwright & Whitworth 2004; Schmeja 2011) and as such further work that considers more realistic initial conditions may be required.

5.2 Expansion of the Lagoon Nebula population

The cause of the observed expansion across the Lagoon Nebula is unclear. The most commonly cited mechanism for the disruption and expansion of a bound cluster is residual gas expulsion, whereby feedback from young stars disperses the gas left over from star formation leaving the remaining stellar part of the system in a supervirial state and prone to dissolution (e.g. Tutukov 1978; Lada et al. 1984). The stars in the Lagoon Nebula are very young, but are no longer embedded in their parental molecular cloud and Damiani et al. (2017) identify a shell of ionized gas expanding towards us away from the region that suggests feedback and therefore residual gas expulsion could be a possible cause of expansion. However, early work on residual gas expulsion predicted a symmetric and radial expansion of the unbound cluster, which is inconsistent with the asymmetric expansion observed here. More recent theoretical studies actually suggest that gas expulsion does not always lead to the stars becoming unbound (e.g. Kruijssen et al. 2012; Parker & Dale 2013), which also argues against this being the cause of the observed expansion pattern.

Preferential expansion in one direction could be caused by tidal disruption of the cluster by a nearby giant molecular cloud, an idea which has been explored in detail by Elmegreen & Hunter (2010) and Kruijssen et al. (2011). Takeuchi et al. (2010) mapped out the molecular clouds in the vicinity of the Lagoon Nebula, identifying massive (each $\sim 2\text{--}6 \times 10^4 M_\odot$) clouds centred on the Hourglass Nebula and M8E, as well as clouds to the south-west and north of the main stellar population. The location of molecular clouds approximately to the north and south of the Lagoon Nebula population might therefore support the role of tidal heating in driving the expansion of the stellar population.

An alternative explanation for the asymmetric expansion pattern is that it is a relic of the formation process (either of the stars or the cluster itself) and that the stellar population has not been sufficiently mixed to erase it. If the region formed from the collision between two molecular clouds in such a way that the remnant cluster was gravitationally unbound and retained a kinematic imprint of the

formation process then this could lead to an apparent expansion pattern along the collision axis. Another possibility is that the region formed from the merger of multiple sub-clusters along a certain axis and is now expanding along that axis following violent relaxation (e.g. Parker & Wright 2016).

The main difference between these scenarios and those of asymmetric residual gas expulsion or tidal heating is whether the expansion pattern is due to a cluster-unbinding mechanism that can produce this asymmetry or whether the kinematics are a remnant of an asymmetry that was present before the star and cluster formation process began (and possibly played some role in the formation). The latter theory does require that the cluster has not been sufficiently well mixed to have erased this primordial asymmetry, but the non-normal and anisotropic velocity dispersions do suggest that such mixing has not yet occurred. Distinguishing between these ideas will either require improved simulations of these process to compare to existing observations or observations of the gas-phase kinematics to search for evidence of a collision.

6 SUMMARY

We have performed a 3D kinematic study of the young stellar population in the Lagoon Nebula and centred around the young cluster NGC 6530. This was performed using a combination of *Gaia*-ESO Survey spectroscopy that provides RVs and *Gaia* DR2 astrometry that provides PMs, as well as various indicators of youth used to identify the stellar population. This led to the following findings.

- (i) Using Bayesian inference and forwarding modelling we have calculated the 3D velocity dispersion of the Lagoon Nebula stellar population to be $5.35^{+0.39}_{-0.34} \text{ km s}^{-1}$, which gives a virial mass more than an order of magnitude larger than the known stellar mass. This implies that the system is gravitationally unbound.
- (ii) The velocity dispersion is anisotropic and the velocity distributions are not well represented by Gaussians, implying that the region is not fully mixed on large scales. The PM velocity dispersion in the right ascension direction is significantly larger than in the declination direction, with the PM velocity ellipsoid aligned east–west.
- (iii) There is moderate evidence for rotation and strong evidence that the region is expanding, with 76 per cent of the kinetic energy in the PM radial component in the direction of expansion. However, the expansion is not symmetric and is almost entirely in the declination direction.
- (iv) The central NGC 6530 cluster has a more compact velocity distribution with a dispersion of $2.75^{+0.39}_{-0.28} \text{ km s}^{-1}$. This also implies a virial mass that is much larger than the known stellar mass, though there is evidence the cluster might have been bound in the past.

To conclude, the stellar population in the Lagoon Nebula is gravitationally unbound and shows clear evidence for expansion. However, the expansion pattern is not consistent with simple residual gas expulsion models that predict a radial expansion pattern. This suggests either a more complex gas expulsion process, tidal disruption from nearby giant molecular clouds, or that the stellar population was never gravitationally bound and that the expansion pattern observed is a remnant of the primordial gas or stellar kinematics from before the cluster formed. The combination of a larger velocity dispersion in the right ascension direction, but a stronger expansion pattern in the declination direction is intriguing and suggests the region is not (and likely never has been) dynamically relaxed and mixed. These results add to the debate

over the impact and effectiveness of residual gas expulsion on the dynamics of young star clusters and stellar systems (e.g. Kruijssen et al. 2012; Parker & Dale 2013; Wright et al. 2014; Ward & Kruijssen 2018; Wright & Mamajek 2018). It is clear that the simple picture of embedded star clusters in virial equilibrium being disrupted by residual gas expulsion cannot reproduce the complex observational picture coming from recent kinematic studies. It remains to be seen whether more complex simulations of the effects of gas expulsion on embedded star clusters can match the observations, or whether all such systems start their lives from more complex initial conditions.

ACKNOWLEDGEMENTS

NJW acknowledges an STFC Ernest Rutherford Fellowship (grant number ST/M005569/1). RJP acknowledges support from the Royal Society in the form of a Dorothy Hodgkin Fellowship. AB acknowledges support from ICM (Iniciativa Científica Milenio) via the Núcleo Milenio de Formación Planetaria. EJA acknowledges support from the Spanish Government Ministerio de Ciencia, Innovación y Universidades through grant AYA2016-75 931-C2-1 and from the State Agency for Research of the Spanish MCIU through the ‘Center of Excellence Severo Ochoa’ award for the Instituto de Astrofísica de Andalucía (SEV-2017-0709). This research has made use of NASA’s Astrophysics Data System and the Simbad and VizieR databases, operated at CDS, Strasbourg. The authors would like to thank Jim Dale and Alexis Klutsch for comments and discussion on this work.

This work is based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory under program ID 188.B-3002. These data products have been processed by the Cambridge Astronomy Survey Unit (CASU) at the Institute of Astronomy, University of Cambridge, and by the FLAMES-UVES reduction team at INAF-Osservatorio Astrofisico di Arcetri. These data have been obtained from the Gaia-ESO Survey Data Archive, prepared and hosted by the Wide Field Astronomy Unit, Institute for Astronomy, University of Edinburgh, which is funded by the UK Science and Technology Facilities Council. This work also made use of results from the European Space Agency (ESA) space mission *Gaia*. *Gaia* data are being processed by the *Gaia* DPAC. Funding for the DPAC is provided by national institutions, in particular the institutions participating in the Gaia MultiLateral Agreement (MLA). The *Gaia* mission website is <https://www.cosmos.esa.int/gaia>. The *Gaia* archive website is <https://archives.esac.esa.int/gaia>. This paper also makes use of public survey data (programme 177.D-3023, the VST Photometric H α Survey of the Southern Galactic Plane and Bulge) obtained at the European Southern Observatory.

REFERENCES

Andrae R. et al., 2018, *A&A*, 616, A8
 André P., Belloche A., Motte F., Peretto N., 2007, *A&A*, 472, 519
 Arenou F. et al., 2018, *A&A*, 616, A17
 Armstrong J. J., Wright N. J., Jeffries R. D., 2018, *MNRAS*, 480, L121
 Barrado y Navascués D., Martín E. L., 2003, *AJ*, 126, 2997
 Baumgardt H., Kroupa P., 2007, *MNRAS*, 380, 1589
 Beccari G., Boffin H. M. J., Jerabkova T., Wright N. J., Kalari V. M., Carraro G., De Marchi G., de Wit W.-J., 2018, *MNRAS*, 481, L11
 Bell C. P. M., Naylor T., Mayne N. J., Jeffries R. D., Littlefair S. P., 2013, *MNRAS*, 434, 806
 Berlanas S. R., Wright N. J., Herrero A., Drew J. E., Lennon D. J., 2019, *MNRAS*, 484, 1838
 Boggs D., Bohm-Vitense E., 1989, *ApJ*, 339, 209

Bonito R., Prisinzano L., Guarcello M. G., Micela G., 2013, *A&A*, 556, A108
 Bravi L. et al., 2018, *A&A*, 615, A37
 Broos P. S. et al., 2013, *ApJS*, 209, 32
 Brown A. G. A., Dekker G., de Zeeuw P. T., 1997, *MNRAS*, 285, 479
 Carpenter J. M., 2000, *AJ*, 120, 3139
 Cartwright A., Whitworth A. P., 2004, *MNRAS*, 348, 589
 Chen L., de Grijs R., Zhao J. L., 2007, *AJ*, 134, 1368
 Conti P. S., Leep E. M., Lorre J. J., 1977, *ApJ*, 214, 759
 Cottaar M., Meyer M. R., Andersen M., Espinoza P., 2012, *A&A*, 539, A5
 Cutri R. M. et al., 2003, 2MASS All Sky Catalog of Point Sources. The IRSA 2MASS All-Sky Point Source Catalog. Available at: <http://irsa.iac.caltech.edu/applications/Gator/>
 Dale J. E., Ercolano B., Bonnell I. A., 2015, *MNRAS*, 451, 987
 Damiani F., Flaccomio E., Micela G., Sciortino S., Harnden F. R., Jr., Murray S. S., 2004, *ApJ*, 608, 781
 Damiani F. et al., 2014, *A&A*, 566, A50
 Damiani F. et al., 2017, *A&A*, 604, A135
 Drew J. E. et al., 2014, *MNRAS*, 440, 2036
 Drew J. E., Herrero A., Mohr-Smith M., Monguió M., Wright N. J., Kupfer T., Napiwotzki R., 2018, *MNRAS*, 480, 2109
 Eggleton P. P., Tokovinin A. A., 2008, *MNRAS*, 389, 869
 Elmegreen B. G., 2002, *ApJ*, 564, 773
 Elmegreen B. G., Hunter D. A., 2010, *ApJ*, 712, 604
 Elson R. A. W., Fall S. M., Freeman K. C., 1987, *ApJ*, 323, 54 (EFF)
 Feigelson E. D. et al., 2013, *ApJS*, 209, 26
 Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
 Gaia Collaboration, 2016, *A&A*, 595, A1
 Gaia Collaboration, 2018, *A&A*, 616, A1
 Garrison R. F., Hiltner W. A., Schild R. E., 1977, *ApJS*, 35, 111
 Gilmore G. et al., 2012, *The Messenger*, 147, 25
 Girichidis P., Federrath C., Banerjee R., Klessen R. S., 2012, *MNRAS*, 420, 613
 González J. F., Veramendi M. E., Cowley C. R., 2014, *MNRAS*, 443, 1523
 Goodman J., Weare J., 2010, *Comm. App. Math. Comp. Sci.*, 5, 65
 Gray R. O., Garrison R. F., 1989, *ApJS*, 69, 301
 Gutermuth R. A. et al., 2008, *ApJ*, 674, 336
 Hiltner W. A., Morgan W. W., Neff J. S., 1965, *ApJ*, 141, 183
 Hogg D. W., Bovy J., Lang D., 2010, preprint ([arXiv:1008.4686](https://arxiv.org/abs/1008.4686))
 Houk N., Smith-Moore M., 1988, *Michigan Catalogue of Two-dimensional Spectral Types for the HD Stars*, Vol. 4. Department of Astronomy, University of Michigan
 Ivezić Ž., Connolly A. J., VanderPlas J. T., Gray A., 2014, *Statistics, Data Mining, and Machine Learning in Astronomy*. Princeton Univ. Press, Princeton, NJ
 Jeffries R. D. et al., 2014, *A&A*, 563, A94
 Kalari V. M. et al., 2015, *MNRAS*, 453, 1026
 Kruijssen J. M. D., Pelupessy F. I., Lamers H. J. G. L. M., Portegies Zwart S. F., Icke V., 2011, *MNRAS*, 414, 1339
 Kruijssen J. M. D., Maschberger T., Moekkel N., Clarke C. J., Bastian N., Bonnell I. A., 2012, *MNRAS*, 419, 841
 Kuhn M. A. et al., 2014, *ApJ*, 787, 107
 Kuhn M. A., Feigelson E. D., Getman K. V., Sills A., Bate M. R., Borissova J., 2015, *ApJ*, 812, 131
 Kuhn M. A., Hillenbrand L. A., Sills A., Feigelson E. D., Getman K. V., 2019, *ApJ*, 870, 32
 Lada C. J., Lada E. A., 2003, *ARA&A*, 41, 57
 Lada C. J., Gull T. R., Gottlieb C. A., Gottlieb E. W., 1976, *ApJ*, 203, 159
 Lada C. J., Margulis M., Dearborn D., 1984, *ApJ*, 285, 141
 Lanzafame A. C. et al., 2015, *A&A*, 576, A80
 Larson R. B., 1995, *MNRAS*, 272, 213
 Levenhagen R. S., Leister N. V., 2006, *MNRAS*, 371, 252
 Lindegren L., 2018, Re-normalising the astrometric chi-square in Gaia DR2. Technical note, GAIA-C3-TB-LU-LL-124-01, Gaia Data Processing and Analysis Consortium
 Lindegren L. et al., 2018, *A&A*, 616, A2

- Loktin A. V., Beshenov G. V., 2001, *Astron. Lett.*, 27, 386
- Luri X. et al., 2018, *A&A*, 616, A9
- Marigo P. et al., 2017, *ApJ*, 835, 77
- Marks M., Kroupa P., 2012, *A&A*, 543, A8
- Maschberger T., 2013, *MNRAS*, 429, 1725
- Mayne N. J., Naylor T., Littlefair S. P., Saunders E. S., Jeffries R. D., 2007, *MNRAS*, 375, 1220
- Moeckel N., Bate M. R., 2010, *MNRAS*, 404, 721
- Odenkirchen M., Grebel E. K., Dehnen W., Rix H.-W., Cudworth K. M., 2002, *AJ*, 124, 1497
- Palla F., Randich S., Flaccomio E., Pallavicini R., 2005, *ApJ*, 626, L49
- Pancino E. et al., 2017, *A&A*, 598, A5
- Parker R. J., Dale J. E., 2013, *MNRAS*, 432, 986
- Parker R. J., Goodwin S. P., 2009, *MNRAS*, 397, 1041
- 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
- Pfalzner S., 2009, *A&A*, 498, L37
- Pourbaix D. et al., 2004, *A&A*, 424, 727
- Prisinzano L., Damiani F., Micela G., Sciortino S., 2005, *A&A*, 430, 941
- Prisinzano L., Damiani F., Micela G., Pillitteri I., 2007, *A&A*, 462, 123
- Prisinzano L. et al., 2019, *A&A*, 623, A159
- Raghavan D. et al., 2010, *ApJS*, 190, 1
- Randich S., Aharpour N., Pallavicini R., Prosser C. F., Stauffer J. R., 1997, *A&A*, 323, 86
- Randich S., Gilmore G., Gaia-ESO Consortium, 2013, *The Messenger*, 154, 47
- Reggiani M. M., Meyer M. R., 2011, *ApJ*, 738, 60
- Rigliaco E. et al., 2016, *A&A*, 588, A123
- Sacco G. G. et al., 2014, *A&A*, 565, A113
- Sacco G. G. et al., 2017, *A&A*, 601, A97
- Schmeja S., 2011, *Astron. Nachr.*, 332, 172
- Scholz F.-W., Stephens M. A., 1987, *J. Am. Stat. Assoc.*, 82, 918
- Siess L., Dufour E., Forestini M., 2000, *A&A*, 358, 593
- Sills A., Rieder S., Scora J., McCloskey J., Jaffa S., 2018, *MNRAS*, 477, 1903
- Soderblom D. R., 2010, *ARA&A*, 48, 581
- Sota A., Maíz Apellániz J., Morrell N. I., Barbá R. H., Walborn N. R., Gamen R. C., Arias J. I., Alfaro E. J., 2014, *ApJS*, 211, 10
- Sung H., Chun M.-Y., Bessell M. S., 2000, *AJ*, 120, 333
- Takeuchi T. et al., 2010, *PASJ*, 62, 557
- Tothill N. F. H., Gagné M., Stecklum B., Kenworthy M. A., 2008, in Reipurth B., ed. *The Lagoon Nebula and its Vicinity*. Astronomical Society of the Pacific. p. 533
- Tutukov A. V., 1978, *A&A*, 70, 57
- Vesperini E., Varri A. L., McMillan S. L. W., Zepf S. E., 2014, *MNRAS*, 443, L79
- Ward J. L., Kruijssen J. M. D., 2018, *MNRAS*, 475, 5659
- Wright N. J., Mamajek E. E., 2018, *MNRAS*, 476, 381
- Wright N. J., Parker R. J., Goodwin S. P., Drake J. J., 2014, *MNRAS*, 438, 639
- Wright N. J., Bouy H., Drew J. E., Sarro L. M., Bertin E., Cuillandre J.-C., Barrado D., 2016, *MNRAS*, 460, 2593
- Zari E., Brown A. G. A., de Bruijne J., Manara C. F., de Zeeuw P. T., 2017, *A&A*, 608, A148
- Zari E., Hashemi H., Brown A. G. A., Jardine K., de Zeeuw P. T., 2018, *A&A*, 620, A172

This paper has been typeset from a $\text{\LaTeX}$ file prepared by the author.