Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **485,** 2457–2485 (2019)
Advance Access publication 2019 February 17

doi:10.1093/mnras/stz471

# 'The Brick' is not a *brick*: a comprehensive study of the structure and dynamics of the central molecular zone cloud G0.253+0.016

J. D. Henshaw [1★] A. Ginsburg,[2] T. J. Haworth [3] S. N. Longmore,[4] J. M. D. Kruijssen [5] E. A. C. Mills,[6] V. Sokolov,[7] D. L. Walker,[8,9] A. T. Barnes,[10] Y. Contreras,[11] J. Bally,[12] C. Battersby,[13] H. Beuther,[1] N. Butterfield,[14] J. E. Dale,[15] T. Henning,[1] J. M. Jackson,[16] J. Kauffmann,[17] T. Pillai,[18,19] S. Ragan [20] M. Riener[1] and Q. Zhang[21]

*Affiliations are listed at the end of the paper*

Accepted 2019 February 4. Received 2019 January 11; in original form 2018 November 29

## ABSTRACT

In this paper we provide a comprehensive description of the internal dynamics of G0.253+0.016 (a.k.a. 'the Brick'); one of the most massive and dense molecular clouds in the Galaxy to lack signatures of widespread star formation. As a potential host to a future generation of high-mass stars, understanding largely quiescent molecular clouds like G0.253+0.016 is of critical importance. In this paper, we reanalyse Atacama Large Millimeter Array cycle 0 HNCO $J = 4(0, 4) - 3(0, 3)$ data at 3 mm, using two new pieces of software that we make available to the community. First, SCOUSEPY, a Python implementation of the spectral line fitting algorithm SCOUSE. Secondly, ACORNS (Agglomerative Clustering for ORganising Nested Structures), a hierarchical $n$-dimensional clustering algorithm designed for use with discrete spectroscopic data. Together, these tools provide an unbiased measurement of the line-of-sight velocity dispersion in this cloud, $\sigma_{v_{\rm los},1{\rm D}} = 4.4 \pm 2.1$ km s$^{-1}$, which is somewhat larger than predicted by velocity dispersion-size relations for the central molecular zone (CMZ). The dispersion of centroid velocities in the plane of the sky are comparable, yielding $\sigma_{v_{\rm los},1{\rm D}}/\sigma_{v_{\rm pos},1{\rm D}} \sim 1.2 \pm 0.3$. This isotropy may indicate that the line-of-sight extent of the cloud is approximately equivalent to that in the plane of the sky. Combining our kinematic decomposition with radiative transfer modelling, we conclude that G0.253+0.016 is not a single, coherent, and centrally condensed molecular cloud; 'the Brick' is not a *brick*. Instead, G0.253+0.016 is a dynamically complex and hierarchically structured molecular cloud whose morphology is consistent with the influence of the orbital dynamics and shear in the CMZ.

**Key words:** turbulence – stars: formation – ISM: clouds – ISM: kinematics and dynamics – ISM: structure – galaxy: centre.

## 1 INTRODUCTION

The lifecycles of molecular clouds and stars are inextricably linked. Molecular cloud evolution drives the formation of the stellar populations that light the Universe and, in turn, feedback from these stars drives the dispersal of the gas clouds from which they are born. It is a self-regulating process that helps to control the evolution of galaxies through cosmic time.

Developing a complete understanding of molecular cloud evolution requires detailed studies that probe a vast range of physical conditions. While nearby molecular clouds (i.e. those within ~500 pc of the Earth) have been studied in extensive detail over the past decades (see e.g. André et al. 2014 and references therein), only now, with facilities such as the Atacama Large Millimeter Array (ALMA), are we able to target the more extreme ends of this parameter space over an equivalent spatial dynamic range.

### 1.1 Star formation in the Milky Way's central molecular zone

The central molecular zone (hereafter, CMZ) of the Milky Way (i.e. the central ~500 pc) contains some of the Galaxy's densest and most massive molecular clouds and star clusters, offering an important window into molecular cloud evolution under extreme physical conditions. The interstellar medium (ISM) conditions found in the

★ E-mail: jonathan.d.henshaw@gmail.com

CMZ differ substantially from those found in the Galactic disc. Molecular gas densities (Guesten & Henkel 1983; Bally et al. 1987; Longmore et al. 2013a; Rathborne et al. 2014a; Mills et al. 2018), pressures (Oka et al. 2001; Rathborne et al. 2014b; Walker et al. 2018), temperatures (Huettemeister et al. 1993; Ao et al. 2013; Mills & Morris 2013; Ott et al. 2014; Ginsburg et al. 2016; Krieger et al. 2017), and velocity dispersions (Bally et al. 1988; Shetty et al. 2012; Henshaw et al. 2016a; Kauffmann et al. 2017a) of CMZ clouds, as well as the cosmic-ray ionization rate (Oka et al. 2005; Yusef-Zadeh et al. 2007) and the interstellar radiation field (Clark et al. 2013), can be factors-of-several to orders of magnitude greater than those found in solar-neighbourhood clouds when compared on the same spatial scale. Although the conditions found in the CMZ are therefore often considered to be extreme in the context of the Milky Way, Kruijssen & Longmore (2013) argue that they are comparable to those found in high-redshift galaxies (e.g. Swinbank et al. 2012) at the time of peak cosmic star formation rate (SFR; around $z \sim 2$–3; Madau & Dickinson 2014). Consequently, understanding stellar mass assembly in the CMZ may help to provide a representative view of the conditions necessary for star formation at its cosmic peak.

One currently open question regarding star formation in the CMZ is that despite harbouring a vast reservoir of dense ($\gtrsim 10^3$ cm$^{-3}$) gas ($\sim$a few $10^7$ M$_\odot$ or roughly $\sim$5 per cent of the total molecular gas content of the Milky Way; e.g. Dahmen et al. 1998), the estimated SFR is just $\lesssim 0.09$ M$_\odot$ yr$^{-1}$ (Longmore et al. 2013a; Koepferl et al. 2015; Barnes et al. 2017). This SFR is approximately one order of magnitude below that expected from the observed linear relationship between the SFR and the gas mass above a surface density of $\Sigma_{\rm gas} = 116\,{\rm M}_\odot{\rm pc}^{-2}$ (Lada, Lombardi & Alves 2010; Lada et al. 2012), despite almost all of the molecular gas in the CMZ lying above this threshold (Longmore et al. 2013a; Barnes et al. 2017). This low SFR cannot be explained by incomplete statistical sampling of independent star-forming regions (Kruijssen & Longmore 2014). Instead, the current underproduction of stars in the CMZ appears to be genuine.

Numerous possible explanations for this discrepancy were discussed by Kruijssen et al. (2014). The authors hypothesized that the low SFR in the CMZ may be due to the high turbulent gas pressure, which would result in an elevated critical density threshold for star formation.[1] This led Kruijssen et al. (2014) to suggest that star formation in the CMZ may be episodic, entering a starburst phase every 10–20 Myr. In this episodic picture, turbulent gas flows towards the Milky Way's CMZ along the Galactic bar providing the fuel for new generations of star formation (as demonstrated in simulations; e.g. Emsellem et al. 2015; Krumholz & Kruijssen 2015; Sormani et al. 2018). The key point is that this process takes *time*: time to build up sufficient gas mass such that gravity can overcome the high turbulent pressure and star formation can proceed at a normal rate (Krumholz & Kruijssen 2015; Krumholz, Kruijssen & Crocker 2017). Previous starburst activity is evident throughout the CMZ. A large population of 24 μm point sources at negative Galactic longitudes (e.g. Hinz et al. 2009) and the young massive clusters known as the Arches and Quintuplet (Figer et al. 1999; Longmore et al. 2014) may add support to the notion of episodicity.

Of course, the CMZ is not in a period of complete dormancy. In fact, it hosts some remarkable star-forming complexes, namely Sgr A, Sgr B1, Sgr B2, and Sgr C (Guesten & Downes 1983; Goss et al. 1985; Mehringer et al. 1992; Mehringer, Palmer & Goss 1993; Yusef-Zadeh et al. 2009; Kendrew et al. 2013; Ginsburg et al. 2018). Where star formation is underway, there is evidence to suggest that it is closely coupled to the orbital dynamics of the gas. Longmore et al. (2013b), studying the subset of CMZ clouds known as the 'dust ridge' (Lis et al. 1994) noted an increase in star formation activity as a function of increasing Galactic longitude along the dust ridge and argued that these clouds may share a common formation timeline. Longmore et al. (2013b) further postulated that star formation may have been triggered by the tidal compression experienced by the clouds as they pass close ($\sim$60 pc; Kruijssen, Dale & Longmore 2015) to the minimum of the global Galactic gravitational potential located at the position of the central supermassive black hole, Sgr A*. The link between the orbital dynamics of the gas and star formation in the dust ridge molecular clouds is supported by trends in observed star formation activity (Immer et al. 2012; Barnes et al. 2017; Ginsburg et al. 2018; Walker et al. 2018) and, less directly, in increasing gas temperatures with increasing Galactic longitude (Ginsburg et al. 2016; Krieger et al. 2017). However, the notion of an evolutionary sequence has also been disputed (see e.g. Kauffmann et al. 2017b; Simpson 2018).

Henshaw et al. (2016b) extended the Longmore et al. (2013b) hypothesis following the discovery of several quiescent molecular clouds situated upstream from (but connected in position–position–velocity or PPV space to) the dust ridge clouds (Henshaw et al. 2016a). Having possibly formed via gravitational instabilities, this portion of the CMZ possibly represents a physically continuous sequence of molecular clouds that we can follow from their formation and on-going assembly through to their subsequent collapse and emergent star formation in the dust ridge.

Theoretically, this picture is supported by recent hydrodynamical simulations of molecular clouds orbiting the Galactic centre. These simulations demonstrate that many of the observed physical features of CMZ clouds are plausibly controlled by the background gravitational potential and their passage through the orbit's pericentre (Kruijssen et al. 2019). However, it is worth noting that the effect of the potential is dominant here, with the triggering of star formation due to pericentre passages expected to be rare (occurring in only $\sim$10–30 per cent of accretion events into the inner CMZ; Jeffreson et al. 2018). Although there are numerous models with differing perspectives on the 3D structure and orbital configuration of the CMZ (e.g. Sofue 1995; Sawada et al. 2004; Molinari et al. 2011; Kruijssen et al. 2015; Ridley et al. 2017),[2] as well as some disagreement on the physical mechanisms driving the flow of material along the Galactic bar and into the CMZ (e.g. Krumholz & Kruijssen 2015; Sormani et al. 2018), there is general agreement that Galactic dynamics plays an important role in the regulation of star formation in this environment (e.g. Kruijssen et al. 2014; Sormani & Barnes 2019).

The aforementioned prominent features are displayed in Fig. 1, where we show a three-colour image of the CMZ generated from *Spitzer* GLIMPSE wavebands. Here, blue is 3.6 μm, green is 5.8 μm, and red is 8.0 μm emission. The group of molecular

[1]The SFR of a molecular cloud is determined in turbulent theories of star formation by computing the gas mass fraction above an effective critical density threshold, $\rho_{\rm crit}$. These theories assume that clouds are supersonically turbulent, and that star-forming cores arise as self-gravitating density fluctuations in the turbulent flow. In the models of Krumholz & McKee (2005) and Padoan & Nordlund (2011), $\rho_{\rm crit} \propto \mathcal{M}_{\rm 3D}^2$, where $\mathcal{M}_{\rm 3D}$ is the turbulent Mach number, leading to an elevated critical density for star formation with increasing turbulent pressure. Although, as summarized by Federrath & Klessen (2012), note that Hennebelle & Chabrier (2011) instead predict $\rho_{\rm crit} \propto \mathcal{M}_{\rm 3D}^{-2}$.

[2]See Henshaw et al. (2016a) for a summary of how some of these geometries can either be ruled out or further constrained by observations.

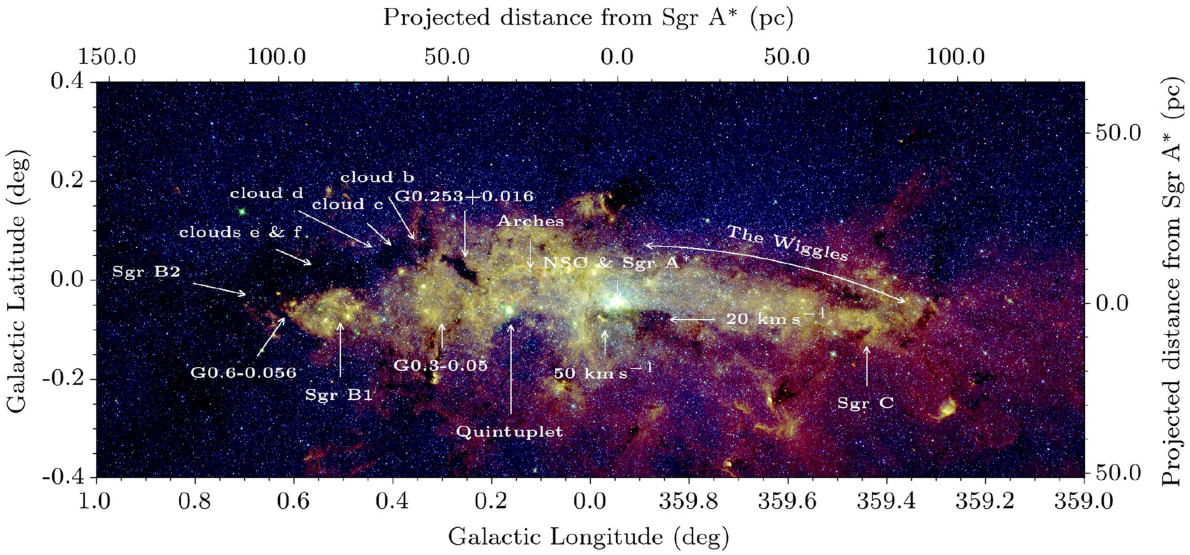


**Figure 1.** G0.253+0.016 in context. A three-colour composite image of the CMZ, highlighting some of the most prominent features. All data are from *Spitzer* GLIMPSE (Churchwell et al. 2009). Blue is 3.6 μm, green is 5.8 μm, and red is 8.0 μm emission. We highlight the dust ridge clouds (G0.253+0.016, clouds 'b', 'c', 'd', 'e & f', and Sgr B2), additional well-known and studied molecular clouds (Sgr C, and the 20 and 50 km s$^{-1}$ clouds), star forming complexes (Sgr B1, G0.6–0.056, and G0.3–0.05), young massive clusters (the Arches and Quintuplet), the location of the velocity oscillations identified in Henshaw et al. (2016a); Henshaw, Longmore & Kruijssen (2016b), and finally, the location of the nuclear star cluster and Sgr A*. G0.253+0.016 can be clearly identified as a strong extinction feature against the bright mid-IR emission arising from the Galactic centre.

clouds collectively known as the dust ridge are those stretching from G0.253+0.016 to Sgr B2.

### 1.2 G0.253+0.016: the prototypical infrared dark cloud

A key proving ground for understanding star formation in the CMZ is the molecular cloud G0.253+0.016 (also, GCM0.253+0.016, G0.216+0.016, M0.25+0.01, M0.25+0.11, or 'The Brick'). G0.253+0.016 is the first cloud in the dust ridge sequence. With a mass of $\sim 10^5$ M$_\odot$ and a radius of just ∼2–3 pc, G0.253+0.016 is one of the densest and most massive molecular clouds within the Galaxy (Lis et al. 1994; Longmore et al. 2012; Kauffmann, Pillai & Zhang 2013; Rathborne et al. 2015). Paradoxically, however, G0.253+0.016 shows very few signatures of active star formation (Mills et al. 2015) and appears mostly in absorption at 8 μm (see Fig. 1). The only direct (and published) evidence for star formation in the cloud comes from an $H_2O$ maser identified by Lis et al. (1994).[3] This makes G0.253+0.016 one of the only $\gtrsim 10^5$ M$_\odot$ molecular clouds in the Galaxy, identified thus far, that does not display the signatures of advanced star formation (Ginsburg et al. 2012; Tackenberg et al. 2012; Urquhart et al. 2014; Longmore et al. 2017). The star formation potential of the cloud is therefore far from certain. Despite G0.253+0.016 having sufficient mass to form an arches-like cluster, it is not clear if we are observing a cloud on the verge of collapse (Longmore et al. 2012; Rathborne et al. 2014a,b, 2015) or if instead the internal turbulent pressure and dynamic surrounding environment will hinder this evolution towards star formation (Kauffmann et al. 2013, 2017a).

Establishing the role of environment on the evolution of G0.253+0.016 is vital if we are to understand its fate. Recently, Federrath et al. (2016) performed an investigation into the physical and dynamical state of the cloud, speculating that shearing motions on large scales may be responsible for the dearth of star formation. The authors discuss this in the context of turbulent star formation theory. Simulations indicate that solenoidal motions (i.e. those with a high degree of vorticity) are capable of suppressing the SFR of a molecular cloud by approximately one order of magnitude in comparison to fully compressive modes (Federrath & Klessen 2012). Combining estimates of the turbulent velocity dispersion and the magnetic field strength, Federrath et al. (2016) conclude that turbulence within the cloud is dominated by solenoidal modes which is the result of the shear on large scales. Highlighting the potential importance of the orbital dynamics, Kruijssen et al. (2015) argue that G0.253+0.016's recent pericentre passage may be the source of the shear. This argument was supported by recent hydrodynamical simulations of molecular clouds following the Kruijssen et al. (2015) orbit, which show that the observed velocity gradient across G0.253+0.016 (e.g. Rathborne et al. 2015) is consistent with shear-induced counter-rotation (Kruijssen et al. 2019).

In this paper, we aim to perform a detailed investigation into the structure and kinematics of G0.253+0.016, which have thus far often been analysed using moment analysis (Higuchi et al. 2014; Johnston et al. 2014; Rathborne et al. 2015; Federrath et al. 2016, although see Kauffmann et al. 2013). Henshaw et al. (2016a) demonstrated that moment analysis' insensitivity to complex line-of-sight density and velocity structure can result in critical information being missed. We therefore revisit the analysis of the kinematics of G0.253+0.016 with the view of categorizing and understanding its internal dynamics. In Section 2 we describe the data used throughout this paper. In Sections 3 and 4 we present our results. In Section 5 we make detailed comparison to previous

[3]Note that there have been claims of ongoing star formation based on more indirect measures. Lis et al. (2001) estimate the internal luminosity of G0.253+0.016 to be of the order of $\sim 2.7 \times 10^5$ L$_\odot$, which they claim is approximately equivalent to that of four B0 zero-age main-sequence stars. Moreover, the presence of emission from warm dust towards the edge of the cloud has been interpreted as being caused by heating from embedded protostars (Marsh et al. 2016). However, these indirect tracers of star formation activity are yet to be supported by independent lines of evidence.

results in the literature. In Section 6 we summarize our new view of the structure of G0.253+0.016 before drawing our conclusions in Section 7.

## 2 DATA

This paper makes use of the ALMA Early Science Cycle 0 Band 3 observations of G0.253+0.016 originally presented in Rathborne et al. (2014b, 2015). The ALMA 12 m observations cover the full 3 arcmin × 1 arcmin extent of the cloud using a 13-point mosaic. The correlator was configured to use four spectral windows in dual-polarization mode centred at 87.2, 89.1, 99.1, and 101.1 GHz, each with 1875 MHz bandwidth and 488 kHz (1.4–1.7 km s$^{-1}$) channel spacing. Because the data were Hanning smoothed by default by the ALMA correlator in Cycle 0, the spectral resolution of the data is 3.4 km s$^{-1}$ (Rathborne et al. 2015). The spatial resolution of the observations is 1.7 arcsec. This corresponds to a physical spatial resolution of ∼0.07 pc assuming a distance to the Galactic centre of 8.34 ± 0.16 kpc (Reid et al. 2014), which we adopt throughout this work, assuming that G0.253+0.016 is at an equivalent distance.

The ALMA data set provided data cubes for 17 different molecular species. Rathborne et al. (2015) studied each of these in detail, making a statistical comparison with the available continuum data (these data were combined with single-dish data provided by the *Herschel* Space Observatory). Measuring the 2-D cross-correlation coefficients, the authors were able to look for similarities between the molecular species and the dust continuum (used here as a proxy for density). The strongest correlations were found between $NH_2CHO$, HNCO, $CH_3CHO$. Out of these species we select the HNCO 4(0,4)–3(0,3) transition (rest freq. ≈87.925 GHz) as our primary tracer of the kinematics since it is bright and extended. HNCO is often spatially extended towards galactic centres (e.g. Dahmen et al. 1997; Meier & Turner 2005; Jones et al. 2012) and has proved fruitful for tracing the gas kinematics on both large (∼pc; Henshaw et al. 2016a) and small (∼0.1 pc; Federrath et al. 2016) scales. The ALMA data were combined with single-dish data available from the Millimetre Astronomy Legacy Team 90 GHz Survey (MALT90; Foster et al. 2011; Jackson et al. 2013) obtained with the Mopra 22 m telescope. For further information regarding the data reduction and image processing we refer the reader to Rathborne et al. (2015).

## 3 A GLOBAL LOOK AT THE KINEMATICS OF G0.253+0.016

### 3.1 SCOUSEPY decomposition of the ALMA HNCO data

Our kinematic decomposition of the ALMA HNCO data is performed using a newly developed Python implementation of the Semi-Automated multi-COmponent Universal Spectral-line fitting Engine (SCOUSE), first presented in Henshaw et al. (2016a).[4] SCOUSEPY is a semi-automated routine used to fit large quantities of complex spectroscopic data in an efficient and systematic way. The procedure followed by SCOUSEPY is discussed in detail by Henshaw et al. (2016a), but we highlight the key points here.

Briefly, the SCOUSEPY fitting procedure can be broken down into several stages. SCOUSEPY first identifies the spatial region over which it will perform the fitting. This can be tailored by the user to target localized regions (in both position and velocity), or to target data above a specified noise threshold. The philosophy behind this step is to minimize workload. For example, although the G0.253+0.016 HNCO data contains $>3 \times 10^5$ pixels, we masked all spectra whose peak flux is below 0.03 mJy beam$^{-1}$. The unmasked region is (approximately) comparable to that studied by Federrath et al. (2016), who employed an $H_2$ column density threshold for their study of $5 \times 10^{22}$ cm$^{-2}$.

SCOUSEPY then breaks up the map into small areas, referred to as Spectral Averaging Areas (SAAs), and extracts a spatially averaged spectrum from each. In the new Python implementation, the user has the option to refine the size of the SAA depending on the local complexity of the line profiles. To gauge the complexity of a spectrum a very simplistic metric is used. We compute the difference in velocity between the intensity-weighed average velocity (i.e. moment 1; $v_1$) to the velocity of the channel containing the peak emission in the spectrum ($v_{\rm peak}$). The idea is that for a simple, singly peaked, symmetric line profile, the difference between these two quantities $\Delta v_{\rm m} \equiv |v_1 - v_{\rm peak}| \sim 0$. Alternatively, $\Delta v_{\rm m}$ will be $>0$ for a highly asymmetric line profile. This is demonstrated in Fig. A1 located in Appendix A. The map is then divided up into different sized SAAs, where the smallest areas contain spectra with a high degree of complexity.

The refinement of the SAA size leads to higher quality fits overall, particularly for large and complex data sets, because of the greater accuracy of the input guesses supplied to the automated fitting procedure. Moreover, having many overlapping SAAs (of potentially different sizes) provides a variety of models to any given pixel, enabling SCOUSEPY to make an informed choice about which is the best-fitting solution.

The spatially averaged spectra extracted from each SAA are then manually fitted by the user. Fitting is performed interactively using PYSPECKIT,[5] whose extensible framework facilitates the modelling of a variety of line profiles (including Gaussian, Voigt, and Lorentzian profiles, as well as hyperfine structure fitting). Specifically, for the ALMA HNCO data, we assume that the spectra can be decomposed into individual Gaussians. This assumption is reasonable given the lack of line wings in the spectral profiles as well as the likelihood that the HNCO emission is optically thin (we quantify this statement further in Section 5.1.2).

Best-fitting solutions to the SAAs are then supplied to the fully automated fitting procedure that targets all of the individual spectra contained within each region. This process is controlled by a number of tolerance levels. For a full description of the tolerances see Henshaw et al. (2016a). In summary, we fixed the following tolerance criteria during our search: (i) all detected components must have a flux density that is greater than three times the local noise value ($T_1 = 3.0$; Henshaw et al. 2016a); (ii) each Gaussian component must have a full-width-at-half-maximum (FWHM) line-width of at least one channel ($T_2 = 1.0$);[6] (iii) for two Gaussian components to be considered distinguishable, they must be separated by at least half of the FWHM of the narrowest of

[4] SCOUSEPY is publicly available for download here: https://github.com/jdhenshaw/scousepy. Alternatively, the original IDL implementation can be downloaded here: https://github.com/jdhenshaw/scouse.

[5] PYSPECKIT can be downloaded here: https://github.com/pyspeckit/pyspeckit.

[6] It should be noted that this leads to the detection of unresolved velocity components. Often these components are necessary for a good fit to the remaining spectral components, and so we choose to fit them. However, as we will discuss later, these components are removed for the clustering analysis (see Section 4).

the two ($T_5 = 0.5$). The remaining two tolerance levels ($T_3$ and $T_4$) restrict the degree to which the parameters describing the velocity components can deviate from their closest matching counterparts in the SAA spectrum. We set both of these tolerance levels to 3.0. As in Henshaw et al. (2016a), the final best-fitting solution for each pixel is that which has the smallest value of the (corrected) Akaike information criterion (Akaike 1974).

The statistical information regarding the SCOUSEPY fitting of the G0.253+0.016 can be found in Table A1 which can be found in Appendix A. To summarize, a total of 2355 SAAs were manually fitted. This resulted in best-fitting solutions to 133 065 out of a total 315 219 pixels (note the total here includes those pixels that were masked during stage 1 of the fitting process), and a total of 457 264 velocity components. Multiple component fits are required to describe the spectral line profiles over a significant (∼96 per cent) portion of the map. These large values indicate the complexity of the velocity structure.

### 3.2 Centroid velocities: ubiquitous velocity oscillations, cloud substructure, and velocity gradients

The result of the fitting procedure is displayed in Fig. 2. This image is a 3D PPV diagram highlighting the distribution of HNCO gas throughout G0.253+0.016. Each data point represents the $\{l,\ b,\ v\}$ coordinates of an individual Gaussian component extracted by SCOUSEPY. The colour (light to dark) of each data point encodes the peak flux density of each spectral component.

The velocity structure of the cloud is clearly complex. The most striking features of Fig. 2 are the vertical velocity oscillations appearing in the gas distribution appearing across a range of spatial scales. These oscillatory gradients are reminiscent of those first identified on larger scales in Henshaw et al. (2016a,b) and suggest that such gradients are a common feature of the ISM in the CMZ. However, unlike those analysed in detail by Henshaw et al. (2016b), which display a characteristic amplitude (∼3.7 ± 0.1 km s$^{-1}$) and wavelength (∼22.5 ± 0.1 pc), the G0.253+0.016 oscillations appear to be more stochastic. This will be explored further in a future publication (Henshaw et al., in preparation).

Further, one notices two large scale, dominant features that appear to merge (caution: in PPV space) towards the southern portion of the cloud. The first appears at a velocity of ∼35–50 km s$^{-1}$. The second shows a distinct velocity gradient increasing in velocity from ∼0 km s$^{-1}$ in the north and appears to merge in PPV space[7] with the first feature at a velocity of ∼30 km s$^{-1}$ towards the south of the cloud. Many studies have described the prominent velocity gradient observed across G0.253+0.016 (e.g. Higuchi et al. 2014; Johnston et al. 2014; Rathborne et al. 2015). Most recently, it has been cited as evidence for the rotation induced by the orbital dynamics of the CMZ (Federrath et al. 2016), which was argued from a theoretical perspective by Kruijssen et al. (2015), and further quantified using hydrodynamical simulations (Kruijssen et al. 2019). In this picture, as a cloud makes its closest approach to the bottom of the Galactic gravitational potential well, the side of the cloud closest to the central potential accelerates with respect to the far-side, inducing shear, and causing the cloud to counter-rotate with respect to its orbital motion.

We can estimate the velocity gradient across G0.253+0.016 using the intensity-weighted velocity field provided by the first-order moment

$$v_{\rm m1} = \frac{\sum_{n=i}^{N} S_v(v_i) v_i}{\sum_{n=i}^{N} S_v(v_i)}, \tag{1}$$

where $S_v(v_i)$ is the flux density at a velocity channel $v_i$ and $N$ is the number of channels. Following Federrath et al. (2016), we compute this over a velocity range of 0–45 km s$^{-1}$ and clip all data below $3\sigma_{\rm rms}$. The velocity gradient is estimated as a fit to all $\{l,\ b,\ v\}$ data points assuming that the velocity field is well approximated by a first-degree bivariate polynomial (e.g. Goodman et al. 1993; Henshaw et al. 2016a)

$$v = v_0 + \mathscr{G}_{v_l} \Delta l + \mathscr{G}_{v_b} \Delta b. \tag{2}$$

Here, $v_0$ is the systemic velocity of the mapped region, $\Delta l$ and $\Delta b$ are the offset Galactic longitude and latitude values (expressed in radians), and $\mathscr{G}_{v_l}$ and $\mathscr{G}_{v_b}$ are free-parameters in the least-squares fit and refer to the magnitudes of the velocity gradients in the $l$ and $b$ directions, respectively (in km s$^{-1}$ rad$^{-1}$). The magnitude of the velocity gradient ($\mathscr{G}$), and its direction ($\Theta_{\mathscr{G}}$), are then estimated using:

$$\mathscr{G} \equiv |\mathscr{G}_{v_{l,b}}| = \frac{(\mathscr{G}_{v_l} + \mathscr{G}_{v_b})^{1/2}}{D}, \tag{3}$$

and

$$\Theta_{\mathscr{G}} \equiv \tan^{-1}\left(\frac{\mathscr{G}_{v_l}}{\mathscr{G}_{v_b}}\right), \tag{4}$$

whereby $D$ is the distance to the cloud in pc (see Section 2). For the velocity gradient, $\mathscr{G}_{v_{\rm m1}}$, we find 4.0 km s$^{-1}$ pc$^{-1}$ ($\mathscr{G}_{v_{\rm m1}}$ = 9.7 km s$^{-1}$ arcmin$^{-1}$).

The computed velocity gradient is consistent with that reported by Federrath et al. (2016), $\mathscr{G}_{v_{\rm m1}} = 3.9$ km s$^{-1}$pc$^{-1}$ ($\mathscr{G}_{v_{\rm m1}}$ = 9.5 km s$^{-1}$arcmin$^{-1}$), where the slight difference is most likely due to the slight difference in the intensity-threshold used.[8] Moreover, this is similar to, albeit slightly larger than the value derived from single-dish MALT90 data $\mathscr{G}_{v_{\rm m1}} = 3.1$ km s$^{-1}$pc$^{-1}$ ($\mathscr{G}_{v_{\rm m1}} = 7.43$ km s$^{-1}$arcmin$^{-1}$; Rathborne et al. 2014a).[9] Despite this general agreement with other observational work, each of these derived-gradients is considerably smaller than the 20 km s$^{-1}$pc$^{-1}$ value quoted by Higuchi et al. (2014), who compute the gradient using the full range of velocities that is spatially coincident with G0.253+0.016 (see below and Fig. 3). However, as discussed in Henshaw et al. (2016a), the southern portion of G0.253+0.016 spatially overlaps with portions of the CMZ gas stream at velocities of ∼70 km s$^{-1}$. This gas, according to our best understanding of the 3D geometry of the gas distribution in the CMZ, is physically unassociated with the cloud. Finally, the velocity gradient derived from the intensity-weighted velocity field is also similar to, but larger than that extracted from simulations of molecular clouds following the Kruijssen et al. (2015) orbit, $\mathscr{G}_v = 2.4$ km s$^{-1}$pc$^{-1}$ ($\mathscr{G}_v = 5.9$ km s$^{-1}$arcmin$^{-1}$; Kruijssen et al. 2019), where this gradient is driven by shear.

[7]We stress that this does not necessarily indicate a merger of structure in physical space.

[8]Here we have simply used an intensity threshold cut in HNCO whereas Federrath et al. (2016) make a cut based on the continuum-derived column density (see Section 3.1). Differences between our results derived from moment analysis and those of Federrath et al. (2016) will therefore propagate throughout any comparisons made in this work. However, we note that the differences are inconsequentially small.

[9]This value actually differs from that reported by Rathborne et al. (2014a), which has been corrected due to a conversion error (see also Kruijssen et al. 2019).

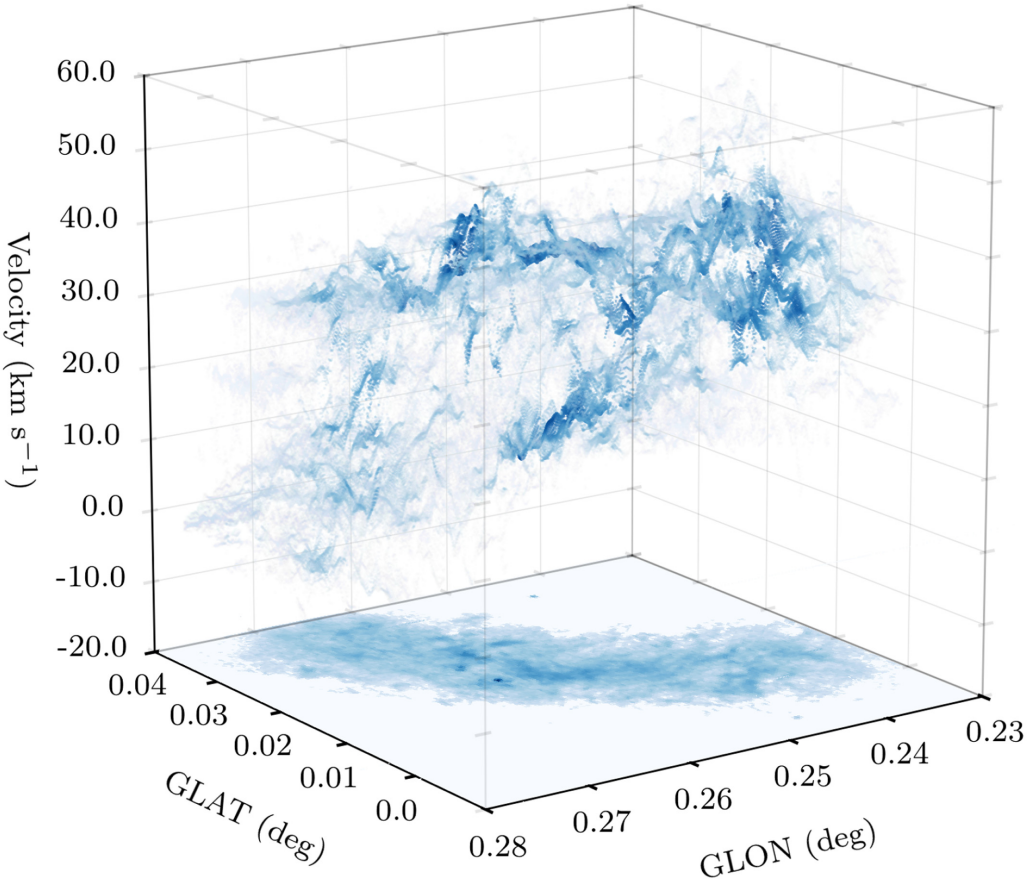


**Figure 2.** A PPV image of G0.253+0.016. Each data point denotes the location and centroid velocity of a Gaussian component identified in HNCO emission in the ALMA data and extracted using SCOUSEPY. The colour of each data point (from light to dark) is proportional to the peak intensity of the corresponding spectral component. Note that only the data between −20.0 and 60.0 km s$^{-1}$ are shown. Emission outside of this velocity range, although spatially coincident with G0.253+0.016 may not be associated with the cloud itself. The full extent of the data can be seen in Fig. 3. At the base of the plot, we plot the 3 mm dust continuum emission first presented in Rathborne et al. (2015). This figure highlights the kinematic complexity of G0.253+0.016.

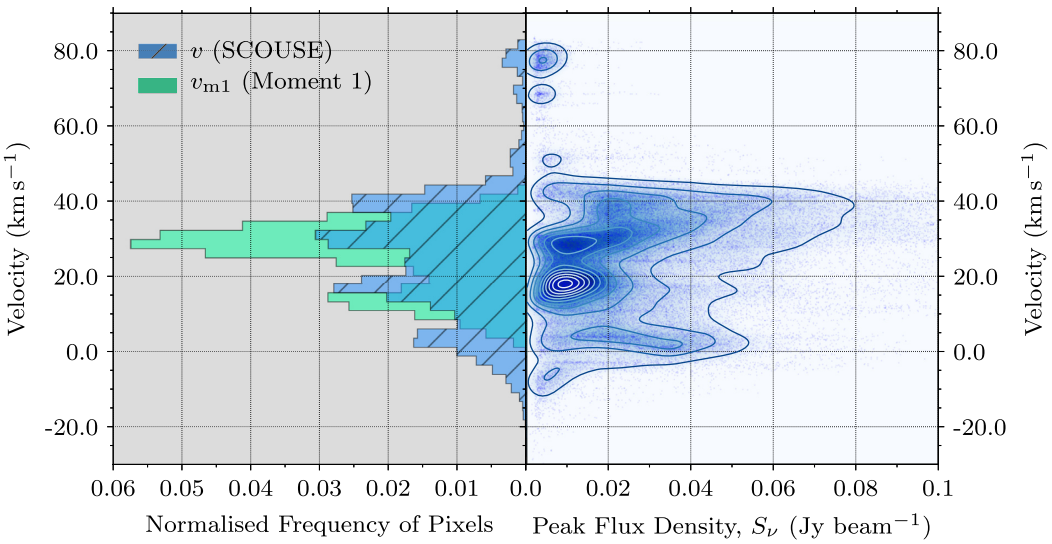


**Figure 3.** Left-hand panel: A histogram of the centroid velocities extracted using SCOUSEPY ($v$, blue) compared with the intensity-weighted average velocities extracted using moment analysis ($v_{m1}$, green). The latter histogram shows a double-peaked profile that has previously been interpreted as a signature of cloud rotation. Right-hand panel: SCOUSEPY centroid velocities, $v$, as a function of peak flux density, $S_\nu$. The contours reflect the point density. Note that we have truncated the $x$-axis in order to show the main structure in the $S_\nu - v$ plane (peak flux densities actually go up to ∼0.4 Jy beam$^{-1}$). The SCOUSEPY decomposition displays significantly more structure over a broader distribution of velocities than that derived from moment analysis.

In the same way as described above, we can also compute the velocity gradient using the information available from our SCOUSEPY decomposition. Here, we find $\mathscr{G} = 4.3\,\mathrm{km\,s^{-1}\,pc^{-1}}$ ($10.5\,\mathrm{km\,s^{-1}arcmin^{-1}}$) in a direction $\Theta_{\mathscr{G}} = -151.7^{\circ}$ east of north. This value is consistent with, albeit larger than, the other observational derivations (see above). This discrepancy is likely a result of the fact that we are utilizing all of our SCOUSEPY measured velocities and ignoring the complex structure presented in Fig. 2. We will revisit this topic in Section 4.

To emphasize the difference between moment analysis and spectral decomposition, we plot in the left-hand panel Fig. 3 a histogram of the SCOUSEPY data in blue and the first-order moment in green. The SCOUSEPY velocity data covers the range $-35.0\,\mathrm{km\,s^{-1}} < v < 87.7\,\mathrm{km\,s^{-1}}$ (note that these extremes may themselves not be associated with G0.253+0.016) and has a mean of $\langle v \rangle \sim 24.96 \pm 0.03\mathrm{km\,s^{-1}}$ (median $= 25.78\,\mathrm{km\,s^{-1}}$; standard deviation $= 16.8\,\mathrm{km\,s^{-1}}$), where the uncertainty here refers to the standard error of the mean. In the right-hand panel we plot the centroid velocity of all the identified velocity components as a function of their peak flux density (we overplot the point density as contours). Both panels of Fig. 3 illustrate that the SCOUSEPY data can be split into four (possibly five) main features. In the histogram there are peaks at ∼3, ∼16, ∼31, and ∼40 $\mathrm{km\,s^{-1}}$ (and a smaller peak at ∼80 $\mathrm{km\,s^{-1}}$), each of which is clearly evident in the $S_v - v_{\mathrm{LSR}}$ plane in the right-hand panel. Some of the multiplicity observed in both panels of Fig. 3 may be a result of the velocity gradients observed across the dominant features seen in Fig. 2 (in the same way that the double-peaked feature in the moment 1 histogram seen in green in the left-hand panel has been interpreted as a signature of rotation; Federrath et al. 2016).

The above analysis demonstrates that although intensity-weighted average quantities may encode important information about the bulk gas dynamics throughout G0.253+0.016, Figs 2 and 3 clearly show that these quantities miss significant detail in the structure and kinematics of the cloud. Therefore, while the kinematics may be interpreted as displaying the hallmarks of rotation, our SCOUSEPY decomposition indicates that a single-component model (i.e. a singular, coherent, and rotating cloud), may be too simplistic in describing the complexity of G0.253+0.016's dynamics and that complex line-of-sight structure is present (we will discuss our interpretation of the cloud structure further in Section 5).

### 3.3 Velocity dispersions and estimated (line-of-sight) turbulent Mach numbers

In the left-hand panel of Fig. 4 we show the distribution of 1D line-of-sight velocity dispersions, $\sigma_{v_{\mathrm{los,1D}}}$, measured with SCOUSEPY. Dispersions range between $0.8\,\mathrm{km\,s^{-1}} < \sigma_{v_{\mathrm{los,1D}}} < 23.1\,\mathrm{km\,s^{-1}}$ (the 25th and 75th percentile are 2.9 and 5.6 $\mathrm{km\,s^{-1}}$, respectively), with a mean value $\langle \sigma_{v_{\mathrm{los,1D}}} \rangle = 4.4\,\mathrm{km\,s^{-1}}$ (median $= 4.0\,\mathrm{km\,s^{-1}}$), and a standard deviation of $2.1\,\mathrm{km\,s^{-1}}$. The standard error of the mean is of the order of $\sim 10^{-3}\,\mathrm{km\,s^{-1}}$.[10] The distribution is skewed towards higher values (with a skewness of ∼1). The skew can also be seen in the right-hand panel, where we show the velocity dispersion as a function of the peak flux density. In the left-hand panel of Fig. 4 we also show the velocity dispersion as derived using moment analysis for comparison. The second-order moment is given by

$$\sigma_{v_{\mathrm{m2}}} = \left[ \frac{\sum_{n=i}^{N} S_v(v_i)(v_i - v_{\mathrm{m1}})^2}{\sum_{n=i}^{N} S_v(v_i)} \right]^{1/2}, \quad (5)$$

where $S_v(v_i)$ is the flux density at a velocity $v_i$ and $v_{\mathrm{m1}}$ refers to the first-order moment (equation 1). The second-order moment values are distributed about a mean value of $\sigma_{v_{\mathrm{m2,1D}}} = 11.1\,\mathrm{km\,s^{-1}}$ and have standard deviation of $4.3\,\mathrm{km\,s^{-1}}$. This latter value is consistent with that derived by Rathborne et al. (2014a; see their fig. 11). The mean velocity dispersion is more than a factor of 2 greater than that extracted using SCOUSEPY which is due to the presence of multiple velocity components in the data.

Our SCOUSEPY-measured mean velocity dispersion differs significantly from the value reported by Federrath et al. (2016). However, Federrath et al. (2016) performed a fundamentally different measurement. These authors instead used the standard deviation of centroid velocities. This represents a measurement of the dispersion of line-of-sight velocities across the plane of the sky (which we label $\sigma_{v_{\mathrm{m1,1D}}}$; as their value is derived from moment analysis) rather than along the line of sight, as is measured (directly) with SCOUSEPY. Consequently, their measured value of $\sigma_{v_{\mathrm{m1,1D}}}$ is a factor of ∼2 larger than our SCOUSEPY-derived mean dispersion, $\sigma_{v_{\mathrm{los,1D}}}$. Repeating their analysis using the first-order moment we find $\sigma_{v_{\mathrm{m1,1D}}} = 9.1\,\mathrm{km\,s^{-1}}$, which is close to the value quoted in Federrath et al. (2016), $\sigma_{v_{\mathrm{m1,1D}}} = 8.8\,\mathrm{km\,s^{-1}}$. By comparison, if we take the standard deviation of all centroid velocity measurements made by SCOUSEPY we find $16.8\,\mathrm{km\,s^{-1}}$, clearly indicating the dominance of multiple velocity components.

In an attempt to isolate the turbulent velocity dispersion (i.e. motions that are exclusively associated with turbulence), Federrath et al. (2016) subtracted the observed large-scale velocity gradient from the intensity-weighted average velocity field. This yields a value of $\sigma_{v_{\mathrm{m1},gs,\mathrm{1D}}} = 3.9\,\mathrm{km\,s^{-1}}$ (where the subscript '$gs$' stands for 'gradient-subtracted'). As discussed in Section 3.2, given the complex distribution of centroid velocities that is evident in Fig. 2, it is unclear whether the velocity gradient observed in the intensity-weighted velocity field can be exclusively attributed to the ordered motion of the cloud. Velocity gradients derived from an intensity-weighted average velocity field may be exaggerated by independent clouds or subclouds situated along the line of sight, each of which has its own independent velocity gradient. Moreover, it is also unclear, on a pixel-by-pixel level, to what extent the intensity-weighted average velocity field (and by extension $\sigma_{v_{\mathrm{m1},gs,\mathrm{1D}}}$) is influenced by the presence of multiple velocity components in the HNCO data. That is to say that different regions within the cloud do not have the same number of components (as can be inferred from Fig. A1) and so the first-order moment will be affected differently as a function of position. Therefore, the subtraction of a singular velocity gradient from an intensity-weighted velocity field should be approached with caution.

We convert our velocity dispersions, $\sigma_{v_{\mathrm{los,1D}}}$, measured on the scale of the synthesized beam (0.07 pc; Section 2), into an estimate of the turbulent Mach number, $\mathcal{M}_{\sigma_{v_{\mathrm{los}},\mathrm{3D}}}$ using (Henshaw et al. 2016a)

$$\mathcal{M}_{\sigma_{v_{\mathrm{los}},\mathrm{3D}}} \approx \sqrt{3}\frac{\sigma_{v_{\mathrm{turb,1D}}}}{c_s} = \sqrt{3}\left[ \left( \frac{\sigma_{v_{\mathrm{los,1D}}}}{c_s} \right)^2 - \left( \frac{\overline{\mu}_{\mathrm{p}}}{\mu_{\mathrm{obs}}} \right) \right]^{1/2}, \quad (6)$$

[10] Note that the velocity components at the lower end of the distribution are unresolved (the spectral resolution is $\Delta v_{\mathrm{res}} = 3.4\,\mathrm{km\,s^{-1}}$). We allowed these components in the SCOUSEPY decomposition to improve the overall quality of the fit. This affects approximately ∼2.5 per cent of the data. However, they are removed from the analysis in Section 4.

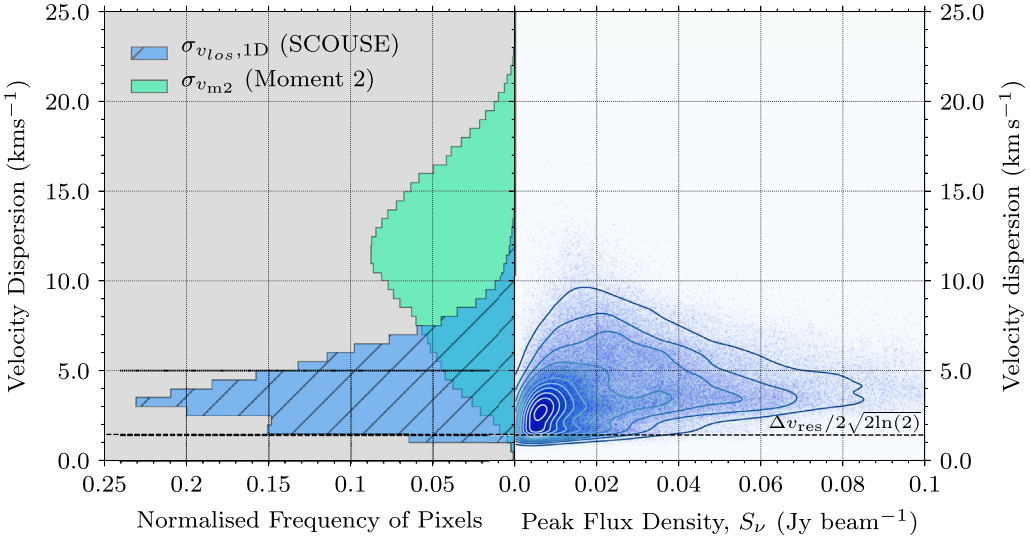


**Figure 4.** Left-hand panel: A histogram of the velocity dispersions extracted using SCOUSEPY ($\sigma_{v_{los},1D}$, blue) compared with the intensity-weighted average velocity dispersion extracted using moment analysis ($\sigma_{v_{m2}}$, green). Right-hand panel: SCOUSEPY velocity dispersion as a function of peak flux density, $S_\nu$. The contours reflect the point density. Note that we have truncated the $x$-axis in order to show the main structure in the $S_\nu - \sigma_{v_{los},1D}$ plane (peak intensities actually go up to $\sim$0.4 Jy beam$^{-1}$). The horizontal line is located at $\Delta v_{res}/2\sqrt{2\ln(2)}$, where $\Delta v_{res} = 3.4\,\mathrm{km\,s^{-1}}$ (the spectral resolution). Note that the tolerance level input during stage 3 of the SCOUSEPY fitting procedure was half of this value to ensure good fits to the data (see Section 3.1 and Henshaw et al. 2016a for further discussion on the input tolerance values for SCOUSEPY). As can be seen, due to the presence of multiple velocity components, moment analysis overestimates the velocity dispersions on average by over a factor of 2 compared to SCOUSEPY.

where $\sigma_{v_{turb},1D}$, in the centre of this equation refers to the 1D turbulent velocity dispersion measured along the line of sight, which we estimate by subtracting the contribution of thermal motions from the observed 1D line-of-sight velocity dispersion in quadrature. The isothermal sound speed is given as $c_s = (k_B T_{kin}/\overline{\mu}_p m_H)^{0.5}$, for a gas with kinetic temperature, $T_{kin}$, and mean molecular mass, $\overline{\mu}_p = 2.33$ amu ($k_B$ and $m_H$ are the Boltzmann constant and the mass of atomic hydrogen, respectively), and $\mu_{obs}$ is the molecular mass of the observed molecule (43 amu in the case of HNCO). Assuming a fixed temperature of 60 K (Ginsburg et al. 2016; Krieger et al. 2017), $c_s = 0.46\,\mathrm{km\,s^{-1}}$. Plugging these values into equation (6), we derive a mean Mach number of $\langle\mathcal{M}_{\sigma_{v_{los},3D}}\rangle = 16.45 \pm 0.01$ (the 25th and 75th percentile are 10.7 and 21.0, respectively), where the uncertainty here reflects the standard error of the mean. This should be taken as an upper bound on the level of turbulent motion, since our assumptions do not take into account the contribution from coherent motions or substructure within the ALMA beam, we assume a uniform temperature, and because of the relatively coarse spectral resolution of the observations $\Delta v_{res} = 3.4\,\mathrm{km\,s^{-1}}$. These factors could combine to result in us overestimating the velocity dispersion and therefore the Mach number throughout the cloud.

This analysis, and the subsequent reduction in measured velocity dispersions and estimated Mach numbers in comparison to other techniques, adds to mounting evidence for the identification of narrow lines in CMZ clouds (Kauffmann et al. 2013, 2017a).[11] This in itself should not come as a surprise, given the increasing spatial resolution of the aforementioned observations. However, despite this, our SCOUSEPY-derived velocity dispersions are broader than those predicted by the observationally derived, steep velocity dispersion–size relationships of the CMZ. Using $\sigma = (\sigma_0/\mathrm{km\,s^{-1}})(r/\mathrm{pc})^\zeta$, where $\sigma_0$ is the absolute scaling of the velocity dispersion and $\zeta$ is the slope, we can predict the magnitude of the velocity dispersions measured on 0.07 pc scales (representative of the ALMA synthesized beam), from the relationships derived by Shetty et al. (2012) and Kauffmann et al. (2017a). Using $\{\sigma_0, \zeta\} = \{2.8\,\mathrm{km\,s^{-1}}, 0.64\}$ (Shetty et al. 2012) and $\{5.5\,\mathrm{km\,s^{-1}}, 0.66\}$ (Kauffmann et al. 2017a), velocity dispersions of the order of $\sim$0.5 and $\sim$1.0 km s$^{-1}$, respectively, are predicted. These are factors of $\sim$9 and $\sim$4 narrower than those measured from our SCOUSEPY decomposition, respectively.

The fact that our mean measured velocity dispersion of $\langle\sigma_{v_{los},1D}\rangle = 4.4\,\mathrm{km\,s^{-1}}$ is fully resolved by ALMA ($2[2\ln(2)]^{1/2}\langle\sigma_{v_{los},1D}\rangle/\Delta v_{res} > 3$, where $\Delta v_{res}$ is the spectral resolution), could indicate that, in contrast to the derived relationships of Shetty et al. (2012) and Kauffmann et al. (2017a), velocity dispersions $\lesssim$1km s$^{-1}$ are not dominant on (projected) $\sim$0.07 pc scales throughout G0.253+0.016. This could imply a shallower velocity dispersion–size relationship. However, this comparison comes with the caveat that although our dispersion measurements are taken on projected scales of the ALMA synthesized beam ($\sim$0.07 pc), we do not know the extent of the cloud along the line of sight. Although this is also true of both the Shetty et al. (2012) and Kauffmann et al. (2017a) studies, the discrepancy between our measured and the predicted velocity dispersions could instead indicate that the depth of the cloud is much greater than the projected spatial extent over which the measurements are taken.

Quantifying both the absolute scaling of non-thermal motions measured at a given spatial scale as well as how the magnitude of non-thermal motions varies as a function of spatial scale throughout the CMZ is of critical importance to understanding star formation in this environment (see Section 1). A steep velocity dispersion–size relationship in the CMZ, if confirmed, may have profound

[11] Although the ALMA cycle 0 data set used here has insufficient spectral resolution to confirm the identification of the more extreme cases (<1km s$^{-1}$) of narrow velocity dispersions presented by Kauffmann et al. (2017a).

implications for how molecular clouds in this environment begin to build their stellar mass.[12] Therefore, higher spatial and spectral resolution observations, those which are capable of resolving the sound speed in the molecular gas (∼0.46 km s$^{-1}$ for 60 K gas), are first required to confirm if the turnover in the SCOUSEPY histogram in the left-hand panel of Fig. 4 is real, and secondly, to fully characterize the gas motions on small spatial scales throughout G0.253+0.016.

## 4 A DETAILED STUDY OF G0.253+0.016'S KINEMATIC SUBSTRUCTURE

### 4.1 ACORNS decomposition of the ALMA HNCO data

To date, analyses of the gas kinematics of G0.253+0.016 have predominantly relied on techniques such as moment analysis (Rathborne et al. 2015; Federrath et al. 2016) and dendrograms (Kauffmann et al. 2013). The former technique is beneficial as it is simple and fast to implement. It returns information on the pixel scale and is an intuitive way of taking a 'first look' at spectroscopic data. However, as is clearly demonstrated in Section 3, detail is easily lost when using moment analysis. Conversely, the latter technique is beneficial in that complex line-of-sight structure is accounted for as the algorithm seeks to build a hierarchy of structure, which can be represented graphically in the form of a dendrogram (see e.g. Rosolowsky et al. 2008). However, kinematic information is provided in the form of intensity-weighted average quantities relating to each structure. Further work is therefore required if one is interested in how those kinematic quantities vary with position within a given structure on the pixel scale. There was previously no publicly available code whose primary function is to extract hierarchical information from spectroscopic data, but which simultaneously retains the pixel scale information needed to study variation in the kinematics throughout each member of the hierarchy.

Our solution to this problem is the development of a new analysis tool, written in Python, named ACORNS (Agglomerative Clustering for ORganising Nested Structures).[13] ACORNS is based on a technique known as hierarchical agglomerative clustering, whose primary function is to generate a hierarchical system of clusters within discrete data. Although ACORNS was designed with the analysis of discrete spectroscopic PPV data in mind (rather than uniformly spaced data cubes), clustering can be performed in $n$-dimensions, and the algorithm can be readily applied using information in addition to PPV measurements. For a full description of the ACORNS algorithm see Appendix B.

In the following sections we use ACORNS to further characterize the velocity structure of the cloud. We perform the ACORNS decomposition only on the most robust spectral velocity components extracted by SCOUSEPY. We define 'robust' as all velocity components whose peak flux density is greater than ∼5 × the typical measured rms noise value[14] and whose velocity dispersion is greater than ∼1.4km s$^{-1}$ (this corresponds to an FWHM of ∼3.4 km s$^{-1}$, which is a single resolution element). The selected data constitute ∼92 per cent of the total data set extracted by SCOUSEPY (420 398 kinematic measurements).

For the clustering, we set the minimum radius of a cluster[15] to be 1.2 arcsec, which is ∼10 per cent larger than the semimajor axis of the ALMA synthesized beam. This is to ensure that all identified clusters are spatially resolved. In addition to spatial information we also include velocity information in the clustering. For two data points to be classified as 'linked' we specify that the Euclidean distance between the points and the absolute difference in both their measured centroid velocity and velocity dispersion can be no greater than 1.2 arcsec and 3.4 km s$^{-1}$, respectively. In summary, these constraints are selected because they reflect our observational limitations.

During the initial phase of the clustering a total of 1152 clusters were identified, representing ∼97 per cent of the subsample selected above.[16] Having fixed these parameters for the initial development of the hierarchy, we then relaxed all linking lengths (position, velocity, and velocity dispersion) by 50 per cent to further develop the clusters. Our final data set contains 1182 clusters, accounting for ∼98 per cent of all data.

As with any hierarchical system of clusters, the result can be displayed graphically as a dendrogram (see e.g. Rosolowsky et al. 2008). In Fig. 5 we display the resultant ACORNS hierarchy for G0.253+0.016. To avoid confusion in star formation nomenclature, we drop the statistical terminology of 'cluster' and instead expand on the nomenclature typically used in describing dendrograms (see e.g. Houlahan & Scalo 1992). We refer to the hierarchical system presented in Fig. 5 as the forest, which itself contains numerous trees. Each tree may then be further subdivided into branches or leaves in a hierarchical fashion (trees with no hierarchical substructure are also classed as leaves). In the case of G0.253+0.016, the forest consists of a total of 195 trees. The forest is dominated by four trees; #3, #22, #85, and #98 (highlighted in red, blue, green, and yellow, respectively). These four trees contain over 50 per cent of all data. In Fig. 6 we display these trees in PPV space (as in Fig. 2). As can be clearly seen in this figure, these trees are associated with the dominant features that are evident in Fig. 2 and discussed in Section 3.2. Given the enormity of the data set, we focus on these dominant trees for the remainder of our analysis. For simplicity, we henceforth refer to the trees as A (red), B (blue), C (green), and D (yellow).

### 4.2 Peak intensity distributions

#### *4.2.1 Tree features: Localized peaks, arcs, and shocks*

In Fig. 7 we display the spatial distribution of peak flux density for each of the main trees to give an impression of their physical structure. While the trees appear to follow the overall distribution and curvature of the cloud, which is commonly observed on large scales in dust continuum maps (see e.g. Johnston et al. 2014; Rathborne et al. 2015), our analysis has also revealed a lot of small-scale structure in the gas distribution.

[12]The shape of the stellar initial mass function, or more specifically, the turnover in the IMF may be closely tied to the sonic length, which is the scale below which thermal or magnetic support dominates over turbulence (see e.g. Offner et al. 2014, and references therein).

[13]ACORNS is publicly available for download here: https://github.com/jdhenshaw/acorns.

[14]This is performed on a pixel-by-pixel basis. The mean rms value is $\langle \sigma_{\rm rms} \rangle = 0.8$ mJy beam$^{-1}$.

[15]Note that here and throughout this paper the term 'cluster' is used in the statistical sense to refer to an agglomeration of data points.

[16]Note that using a linking length of 1.7 km s$^{-1}$ for both the centroid velocity and velocity dispersion (i.e. a single channel), respectively, changes the results only slightly. In this case, the total number of clusters identified is 1231 and these clusters contain ∼95 per cent of the data. This does not however, affect any of the conclusions of this work.

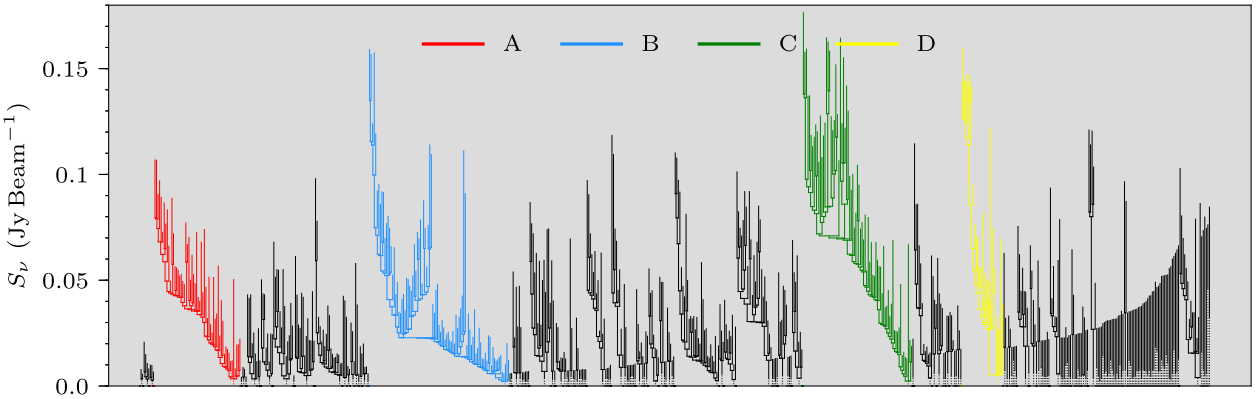


**Figure 5.** ACORNS clustering displayed graphically as a dendrogram. Here we display the full forest of clusters. Each tree in the forest can be further subdivided into branches and leaves in a hierarchical fashion (see Section B1 for a full description of the method and nomenclature). The forest, comprising a total of 195 trees, is dominated by four trees; A, B, C, and D (highlighted in red, blue, green, and yellow, respectively). Together they comprise > 50 per cent of all data. After these first four trees there is a factor of ∼2 drop in the next tree's percentage contribution to the total data set.

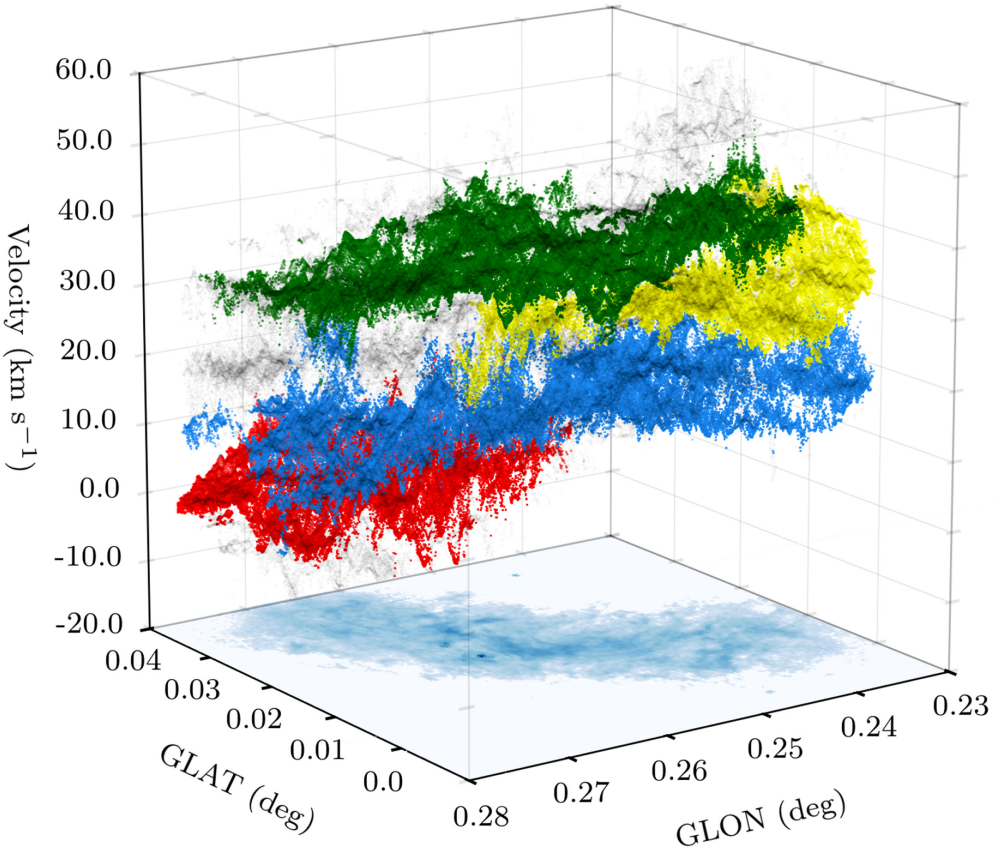


**Figure 6.** In this image we highlight the dominant ACORNS trees in PPV space. Each colour refers to a different tree in the forest: A (red), B (blue), C (green), and D (yellow) [see Section 4.1 and Fig. 5]. These four (out of 195) trees contain > 50 per cent of all data. The full data set is included in this image as small black data points. The image at the base of the plot is equivalent to that presented in Fig. 2.

Trees B and C, appearing in blue and green in Fig. 6 and in the top right and bottom left of Fig. 7, respectively, are the most prominent of the identified trees. Together they dominate the physical appearance of G0.253+0.016, accounting for ∼34 per cent of the data (which is roughly distributed evenly between them). A cursory visual comparison of the two trees in Fig. 7 suggests that the HNCO emission is brighter throughout tree C, on average. This can be inferred from Fig. 5, where C has a greater number of leaves that have greater peak flux density than those associated with B.

Qualitatively, the small-scale peaks of emission (identified as leaves by ACORNS) in tree C show a similar spatial distribution to those observed in the corresponding 3 mm dust continuum image presented by Rathborne et al. (2015) and displayed in Fig. 8. There is, however, a notable exception. The green circle in Fig. 7 denotes the location of the $H_2O$ maser identified by Lis et al. (1994). This

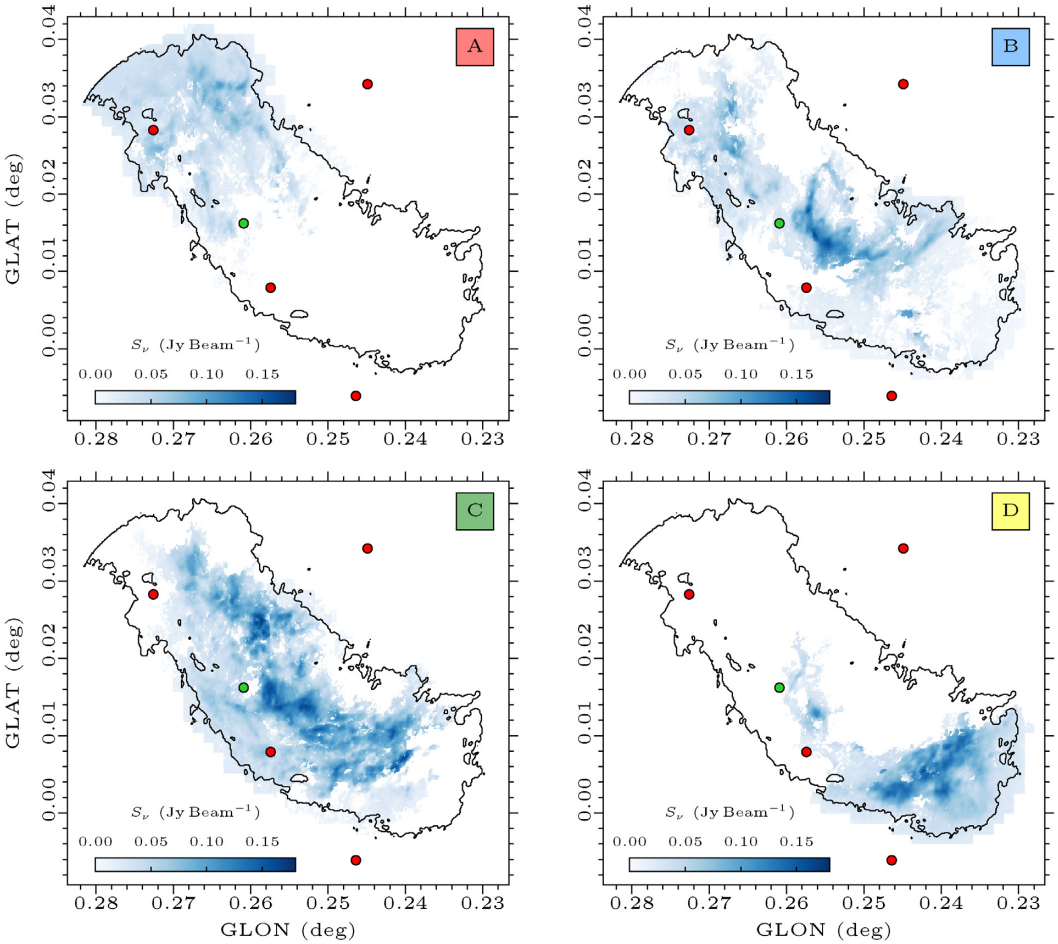


**Figure 7.** Peak flux density distributions of the main ACORNS trees associated with G0.253+0.016. The coloured squares in the top right-hand corner of each image refer to the colouring used in Figs 5 and 6. The black contour highlights the boundary of the cloud as defined during the masking described in Section 3.1. The green circle indicates the location of the $H_2O$ maser identified by Lis et al. (1994). Note that this coincides with a hole in the intensity distribution of tree C (green). This is discussed further in Section 4.2. The red circles indicate the locations of compact radio continuum sources (Rodríguez & Zapata 2013).

coincides with a 'hole' in the emission associated with tree C and we will discuss this further in Section 4.2.2.

Another prominent feature evident in Fig. 7 is the 'C'-shaped arc structure associated with tree B (top-right panel of Fig. 7). The arc was originally discovered with ALMA as a prominent feature traced by sulphur monoxide (specifically the SO $[v = 0,\ 3(2) - 2(1)]$ transition).[17] Higuchi et al. (2014) characterize the arc as being associated with a number of emission peaks (both in the dust continuum and SO), some of which show broad velocity dispersions (of the order of 30–40 km s$^{-1}$) as well as strong velocity gradients. Despite these relatively extreme values, the right-hand panel of Higuchi et al. (2014)'s fig. 2 (which displays the second-order moment map), shows that most of the emission associated with the arc has velocity dispersions up to ∼10 km s$^{-1}$. Mills et al. (2015) later confirmed that the arc is observed in other molecular species and transitions, identifying it clearly in the (peak) emission maps of $NH_3$ transitions from (1,1) up to (7,7). Although the presence of the 'C'-shaped arc was therefore noted in previous studies, ACORNS provides the first evidence that the arc is coherent in both (projected) space and velocity.

Tree D (bottom right-hand panel of Fig. 7) resides at the interface of trees B and C in terms of velocity (see Fig. 6 and Section 4.3.1). This tree is associated with a linear feature referred to as the 'tilted bar' by Mills et al. (2015) contains the bulk of the brightest clumps seen in $NH_3$ (3, 3) and a multitude of 'class I', collisionally excited, and shock tracing $CH_3OH$ masers and maser candidates. The 'tilted bar' is also evident in Johnston et al. (2014)'s fig. 14 which displays the integrated flux line ratio of different $H_2CO$ transitions. Radiative transfer analysis suggests that this region shows elevated gas temperatures (Johnston et al. 2014) consistent with Mills et al. (2015). Moreover, this region is observed to exhibit enhanced emission from shocked and warm (>140 K) gas tracers (e.g. SiO (5-4) and $H_2CO$; Kauffmann et al., in preparation). These features are complemented by the HNCO emission, which is very bright throughout the tree and follows a linear feature running perpendicular to the major axis of G0.253+0.016. This region of the cloud has previously been cited as a potential location for cloud–cloud collisions (Johnston et al. 2014) and the linear feature that is observed may be the result of large-scale shocks (Mills et al. 2015). We will return to this discussion in Section 5.

[17]These data were taken as part of the same ALMA cycle 0 data set as that used in this paper: ADS/JAO.ALMA#2011.0.00217.S. The spatial (∼1.9 arcsec) and spectral (∼3.4km s$^{-1}$) resolutions are therefore approximately equivalent to the HNCO data presented here.

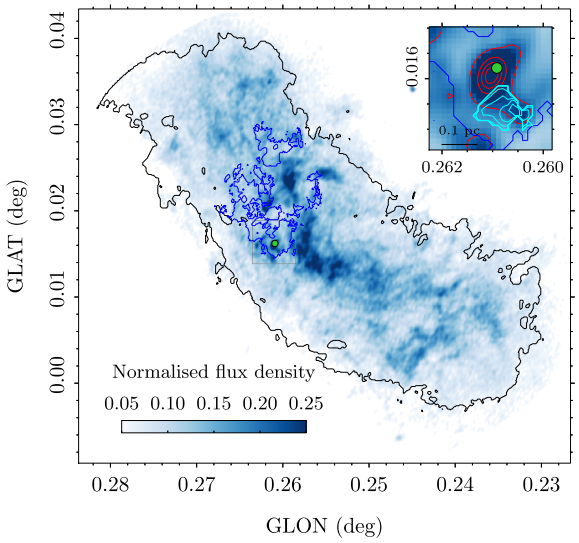


**Figure 8.** ALMA 3 mm dust continuum observations of G0.253+0.016 in colour scale (Rathborne et al. 2014b). The colour scale has been normalized against the peak emission and intentionally truncated in order to emphasize the more diffuse emission. The black contour is equivalent to that presented in Fig. 7. Overlaid in blue is the outline of tree #32 (see Section 4.2.2 for details). In the inset image we zoom in on the dust continuum peak associated with the $H_2O$ maser identified by Lis et al. (1994). In red contours we display the 3 mm dust continuum (from [0.2, 0.4, 0.6, 0.8] × the peak emission ∼0.004 Jy beam$^{-1}$). The blue contour indicates the outline of the tree. In cyan contours we show the leaf that is closest to the bright 3 mm peak seen in dust continuum. The filled green circle in both the main image and the inset indicates the location of the $H_2O$ maser identified by Lis et al. (1994).

Finally, tree A overall has fewer regions of bright emission than the others despite showing a lot of substructure. This is evident in Fig. 5, where the tree is seen to exhibit a complex dendrogram. In larger-scale single-dish observations of G0.253+0.016 emission at the low velocities associated with A extends further north of the cloud in the direction of the dust ridge cloud 'b' (see e.g. Lis et al. 2001), whose mean velocity is measured to be ∼3.4km s$^{-1}$ (Henshaw et al. 2016a). This extension is also evident in dust continuum observations (see e.g. Immer et al. 2012).

#### *4.2.2 Star formation within G0.253+0.016*

In the previous section we noted that there is a lack of emission in tree C at the only (currently) confirmed location of ongoing star formation in G0.253+0.016.[18] To investigate whether or not there is a true absence of emission at this location, we first of all inspected the best-fitting solutions extracted using SCOUSEPY. A cursory inspection indicates that there are several velocity components at this location. We then further explored the ACORNS hierarchy for any trees that spatially overlap with the $H_2O$ maser and are located at the 'appropriate' velocity (Lis et al. 1994 quote velocities of 32.1 and 41.6 km s$^{-1}$ for the maser). Using these criteria we identified two trees (#32 and #108). We then identified all leaves that spatially coincide with the gap in emission associated with tree C.

[18]Note that recently, two additional $H_2O$ masers have recently been discovered. One further to the north of the cloud at 70 km s$^{-1}$ and another ∼10 arcsec to the south of the maser identified by Lis et al. (1994) at a velocity of 28.4 km s$^{-1}$ (Lu et al. 2019).

We identify a centrally concentrated leaf associated with the first of these two trees that fits this criteria. It has a mean centroid velocity of ∼42.0 km s$^{-1}$ and a velocity dispersion of ∼2.8 km s$^{-1}$. Given the spectral resolution of 3.4 km s$^{-1}$, this is in satisfactory agreement with the velocity of the $H_2O$ maser identified by Lis et al. (1994). In Fig. 8 we plot the 3 mm continuum map first presented by Rathborne et al. (2014b). Overlaid on this image we display the contoured outline of the tree (blue). In the inset image we zoom in on the 3 mm dust continuum peak (red contours and background) that is associated with the maser emission identified by Lis et al. (1994). Comparing the ALMA dust continuum with our ACORNS decomposition, we find that the ACORNS leaf (cyan contours) does not trace the main dust continuum peak, but instead traces an extension of this peak observed towards the south.

To further investigate this, we compare our results with new high-resolution (∼0.13 arcsec) ALMA band 6 observations towards the maser region (Walker et al. in preparation). Using a combination of dust continuum observations and $CH_3CN$ emission we find that there is evidence for line emission associated with the 3 mm dust continuum peak at ∼42–43 km s$^{-1}$ consistent with the velocity of tree #32. The reason for the lack of a line emission peak in our 3 mm HNCO data is currently unclear and further investigation at high-angular resolution and with molecular line tracers that probe different critical densities and excitation conditions are necessary. Nevertheless, there is an evidence for a small compact continuum source that coincides with the extension in emission seen in the 3 mm data presented in Fig. 8 (D. L. Walker, private communication), and therefore our ACORNS leaf.

### 4.3 Gas kinematics

#### *4.3.1 Centroid velocities: non-Gaussian velocity PDFs and velocity gradients*

In Fig. 9 we plot velocity probability density functions (PDFs) of the ACORNS trees. In laboratory experiments of incompressible turbulence, the PDF of the velocity field is often very nearly Gaussian (see e.g. Anselmet et al. 1984). This has also been demonstrated in numerical simulations of turbulence (Lis et al. 1996; Klessen 2000; Federrath 2013). Results from observations of the ISM however, have been mixed and largely show some deviation from pure Gaussian behaviour (e.g. Miesch, Scalo & Bally 1999; Ossenkopf & Mac Low 2002; Federrath et al. 2016). To assess this we fit a normal distribution to the centroid velocities measurements associated with each tree and also compute the higher order moments (skewness and kurtosis) of the distributions. The first four central moments of a data set (in our case $v$) with $N$ elements are

$$\text{mean} = \langle v \rangle = \frac{1}{N}\sum_{n=i}^{N} v_i \tag{7}$$

$$\text{dispersion} = \sigma_{v_{pos},\text{1D}} = \sqrt{\langle [v_i - \langle v \rangle]^2 \rangle} \tag{8}$$

$$\text{skewness} = \mathcal{S} = \frac{\langle [v_i - \langle v \rangle]^3 \rangle}{\sigma^3_{v_{pos},\text{1D}}} \tag{9}$$

$$\text{kurtosis} = \mathcal{K} = \frac{\langle [v_i - \langle v \rangle]^4 \rangle}{\sigma^4_{v_{pos},\text{1D}}}. \tag{10}$$

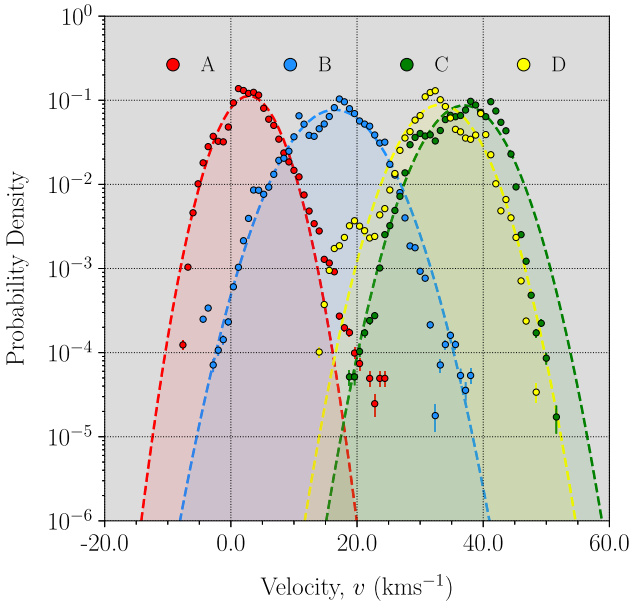


**Figure 9.** Centroid velocity probability density functions (PDFs) of the main ACORNS trees A (red), B (blue), C (green), and D (yellow) [cf. Fig. 5]. The dashed lines are normal distributions fitted to the data. Despite the velocity PDFs appearing broadly consistent with Gaussian distributions (the profiles have a mean kurtosis value of $\langle\mathcal{K}\rangle \sim 3.3$), there are statistically significant deviations from Gaussianity.

Note that the dispersion in equation (8) is a measurement of the dispersion of centroid velocities in the plane of the sky measured across the trees, which we denote $\sigma_{v_{\rm pos},1D}$ (this will be discussed further in Section 4.3.2; cf. $\sigma_{v_{\rm m1},1D}$ in Section 3.3). The skewness and kurtosis are measures of the symmetry and flatness of a distribution, respectively. Negative skewness indicates that the distribution is skewed to the left and a positive skewness the opposite. A Gaussian distribution has a kurtosis of 3. A value larger than 3 implies that the distribution has prominent tails, and therefore rarer, high-amplitude events occur more frequently than would be expected for purely Gaussian behaviour. A value less than 3 implies the opposite.

The trees are mostly well separated in velocity (as can also be seen in Fig. 6) with mean velocities of $\langle v\rangle = \{2.9, 16.5, 33.1, 37.0\}$ km s$^{-1}$ for trees A, B, D, and C, respectively. Note however, that this is not an explicit requirement of ACORNS. For example, trees C and D are more closely related in velocity but are identified as distinct due to their differing velocity dispersions (their median velocity dispersions are separated by $\sim$2 km s$^{-1}$; see Section 4.3.2).

Each of the trees shows a slightly skewed distribution of centroid velocity. Trees B, C, and D are negatively skewed while A is positively skewed. In terms of the kurtosis, A and D have similar values of $\mathcal{K} \sim 3.9$ indicating that the tails of the distribution are more prominent than those expected from a purely Gaussian distribution. Conversely, A has $\mathcal{K} \sim 2.5$. Finally, B has a kurtosis value of $\mathcal{K} \sim 3.0$. Despite all clusters having $\mathcal{S} \lesssim |0.4|$ and $2.5 < \mathcal{K} < 3.9$, the centroid velocities of the clusters are statistically inconsistent with Gaussian distributions based on the computation of the D'Agostino ($p \approx 0.0$, which combines the skewness and kurtosis of the distribution; D'agostino et al. 1990) and the Anderson-Darling statistics ($p \approx 0.0$; Anderson & Darling 1952).

It has been argued that a deviation from Gaussianity can occur when systematic or ordered motions are present within the velocity field. Federrath et al. (2016) recently argued that the large-scale velocity gradient observed across G0.253+0.016 contributes to producing a non-Gaussian velocity PDF. After subtraction of the systematic motions from the velocity field, Federrath et al. (2016) state (following visual inspection of the data) that the velocity PDF is in excellent agreement with a Gaussian profile, and used this as a method to decouple the contribution of turbulent gas motions from the observed velocity dispersion.

It is worth noting that despite appearing consistent with a Gaussian profile, the gradient-subtracted velocity field derived from intensity-weighted mean velocities (see Section 3.2) also produces a non-Gaussian distribution, in a statistical sense. We examine the gradient-subtracted velocity field for the moment 1 map and find: $\langle v\rangle = 0.0$ km s$^{-1}$ (note this is because the gradient has been subtracted); $\sigma_{v_{\rm m1},gs,1D} = 4.0$km s$^{-1}$; $\mathcal{S} = -0.24$; $\mathcal{K} = 3.0$. As with the ACORNS trees, the null hypothesis that the distribution of velocities is drawn from a Gaussian distribution can be rejected following the computation of the D'Agostino and Anderson-Darling statistics ($p$ values $\approx 0.0$). However, with many physical processes at work within the ISM, deviations from Gaussianity are unsurprising (Klessen 2000). Moreover, Federrath et al. (2016) clearly acknowledge that there are residual deviations from their Gaussian fit. These deviations, the authors argue, are most likely due to a combination of noise in the data, the excitation conditions of HNCO, and the fact that small-scale systematic motions may still be present in the data.

In Fig. 10 we plot the velocity fields of the four ACORNS trees. Velocity gradients are clearly evident in the data. Using the methodology outlined in Section 3.2 we compute velocity gradients for each tree. We find $\mathscr{G} = \{1.1, 2.1, 1.9, 4.2\}$ km s$^{-1}$pc$^{-1}$ (corresponding to $\mathscr{G} = \{2.7, 5.0, 4.6, 10.2\}$ km s$^{-1}$arcmin$^{-1}$) in directions $\Theta_{\mathscr{G}} = \{-82.6^\circ, -141.9^\circ, -129.6^\circ, -31.2^\circ\}$ east of north for trees, A, B, C, and D, respectively. The magnitude of the velocity gradients of trees B and C are more consistent with those derived from simulations of molecular clouds following the Kruijssen et al. (2015) orbit (2.4 km s$^{-1}$pc$^{-1}$; Kruijssen et al. 2019; cf. Section 3.2). We display the magnitude and direction of these gradients as arrows in Fig. 10.

### *4.3.2 Velocity dispersions: plane of the sky versus line-of-sight velocity fluctuations*

In this section we focus on the velocity dispersions of the ACORNS trees. The standard deviation of centroid velocities estimated above (equation 8) provides an estimate for $\sigma_{v_{\rm pos},1D}$ for each tree. For trees A, B, C, and D we measure $\sigma_{v_{\rm pos},1D} = \{3.5, 5.2, 4.6, 4.5\}$ km s$^{-1}$, respectively.

If we recompute the dispersions after subtracting a 2-D velocity plane constructed from the velocity field of each tree [cf. the linear model in equation (2) and the gradients displayed in Fig. 10), we find for A, B, C, and D $\sigma_{v_{\rm pos},gs,1D} = \{3.3, 4.1, 3.1, 4.1\}$ km s$^{-1}$, respectively (where the subscript *gs* stands for 'gradient-subtraction'). Accounting for these large-scale systematic motions leads to a reduction of $\sim$17 per cent in the total dispersion of centroid velocities, in contrast to the $\sim$56 per cent reduction inferred by Federrath et al. (2016). This indicates that although large-scale systematic motions, if they are indeed systematic, may contribute to the observed dispersion in the plane of the sky velocity, they do not dominate.

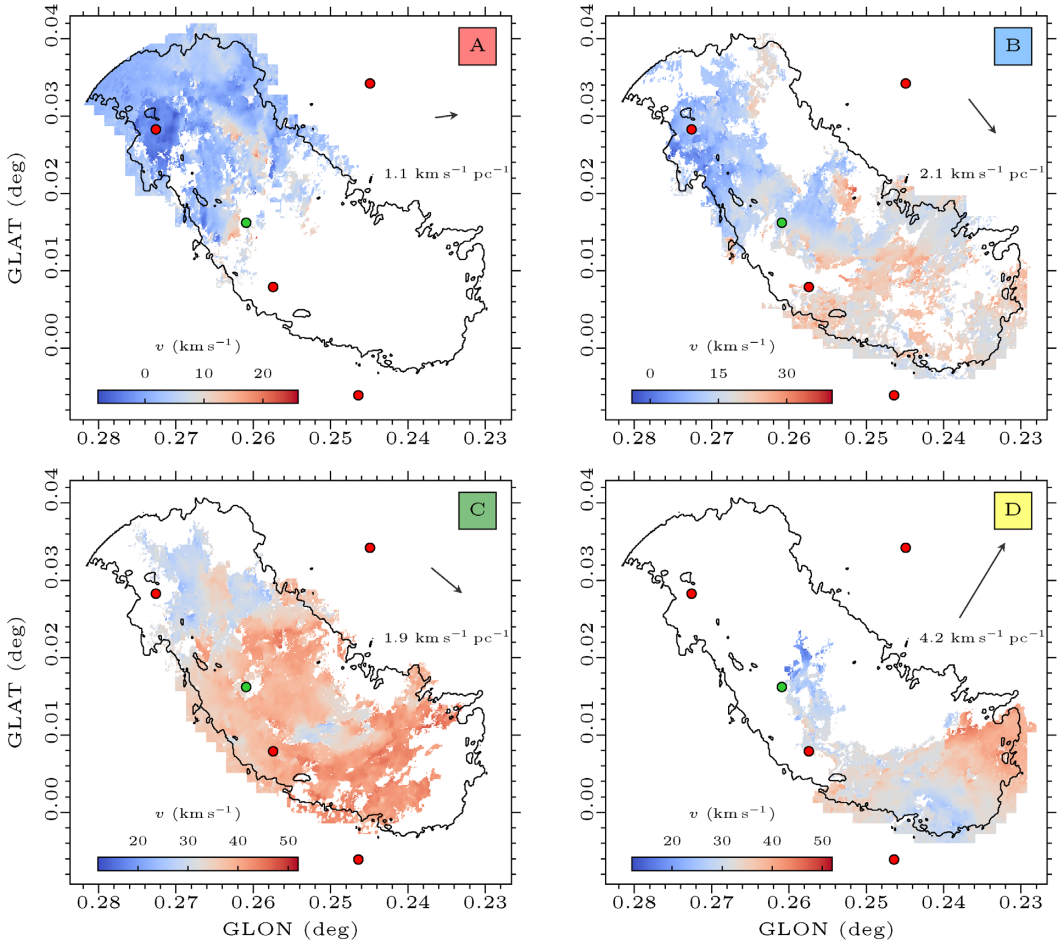


**Figure 10.** Equivalent to Fig. 7 but for the centroid velocities, $v$, measured throughout each ACORNS tree. The size and direction of the arrow in each plot represents the magnitude and direction of the velocity gradient across each tree (pointing in the direction of increasing velocities).

In Fig. 11 we plot histograms of $\sigma_{v_{\rm los,1D}}$ for each of the ACORNS trees. For A, B, C, and D we find $\langle\sigma_{v_{\rm los,1D}}\rangle = \{5.3, 4.9, 4.0, 5.8\}$ km s$^{-1}$, respectively (where the angle brackets indicate that we have taken the mean value over all SCOUSEPY measurements associated with each cluster). As is evident in Fig. 11, the distributions are skewed and so we report median velocity dispersions of $\{5.2, 4.5, 3.8, 5.8\}$ km s$^{-1}$. In the bottom panels, we plot the cumulative histograms of the velocity dispersions. A two-sample Kolmogorov–Smirnov test for each of the six unique parings of the four tree samples reveals that none of the samples are drawn from the same distribution, indicating that there are statistical differences between the clusters in terms of their measured line-of-sight velocity dispersions. The peak of the distribution for D is, for example, shifted rightwards from those of B and C indicating broader velocity dispersions on average. This can be seen in Fig. 12, where we have plotted the spatial distribution of velocity dispersions throughout each tree.

It is notable that taking the ratio of the line-of-sight and plane of the sky velocity dispersions yields $\sigma_{v_{\rm los,1D}}/\sigma_{v_{\rm pos,1D}} = \{1.5, 0.9, 0.9, 1.3\}$ for trees A, B, C, and D, respectively. On average this is $\langle\sigma_{v_{\rm los,1D}}/\sigma_{v_{\rm pos,1D}}\rangle = 1.2 \pm 0.3$, where the uncertainty here reflects the standard deviation. We speculate that this isotropy in the line-of-sight velocity distribution and the line-of-sight fluctuations in the centroid velocity in the plane of the sky may encode information about the cloud geometry. Namely, that the line-of-sight extent of the cloud components are approximately equivalent to that in the plane of the sky. This could perhaps explain some of the discrepancy between our measured velocity dispersions and those predicted from the steep velocity dispersion–size relationship derived for the CMZ (see Section 3.3). However, we hasten to add that this result would need to be tested rigorously with numerical simulations.

## 5 IS 'THE BRICK' *REALLY* A BRICK? A CENTRALLY CONDENSED MOLECULAR CLOUD VERSUS MULTIPLE COLLIDING (SUB-)CLOUDS

G0.253+0.016's moniker, 'the Brick', reflects both its shape on the plane of the sky and the fact that we see it in silhouette against the bright Galactic mid-infrared background at the Galactic Centre (see Fig. 1 and Longmore et al. 2012). However, the analysis presented in Sections 3 and 4 reveals substantial and complex substructure in both position and velocity, consistent with prior analyses that identified cores, filaments, and other coherent features (Bally et al. 2014; Higuchi et al. 2014; Johnston et al. 2014; Rathborne et al. 2014b; Mills et al. 2015; Rathborne et al. 2015; Federrath et al. 2016). In the following sections we discuss the current understanding of the structure of G0.253+0.016 both in

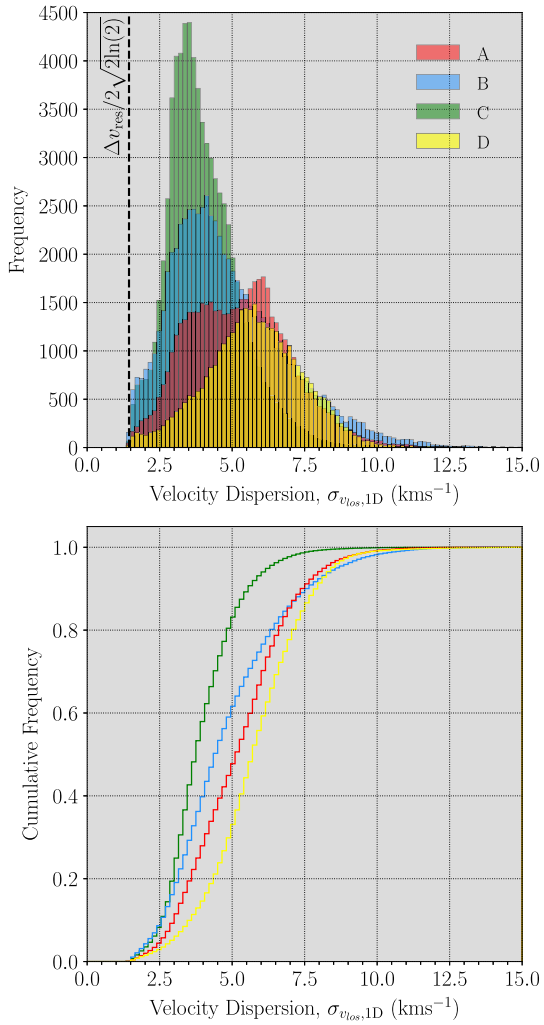


**Figure 11.** Top: Histograms of the velocity dispersion, $\sigma_{v_{\rm los,1D}}$, for the main ACORNS trees (coloured histograms; see Fig. 5). The vertical line is located at $\Delta v_{\rm res}/2\sqrt{2\ln(2)}$, where $\Delta v_{\rm res} = 3.4 {\rm km\,s^{-1}}$ (the spectral resolution). Bottom: Cumulative histograms of the velocity dispersions for the trees.

terms of the kinematic analysis presented in this work and in the global context of the CMZ.

### 5.1 'The Brick': G0.253+0.016 as a centrally condensed molecular cloud

Using single-dish observations from the MALT90 survey (Foster et al. 2011; Jackson et al. 2013), Rathborne et al. (2014a) presented a study of the structure of G0.253+0.016. One of the distinctive features noted by the authors was the presence of multiple velocity components associated with G0.253+0.016 (much like in Fig. 2). Rathborne et al. (2014b) presented two possible explanations for the presence of multiple velocity components in G0.253+0.016: (i) that G0.253+0.016 is a single, coherent, centrally condensed cloud with depletion in its cold interior; (ii) that the two velocity components reflect two clumps colliding. Rathborne et al. (2014a) favour the former of the two scenarios, which we assess in this section. In Section 5.2 we discuss the cloud collision scenario.

#### *5.1.1 Scenario 1a: Optically thin lines: G0.253+0.016 is a centrally condensed molecular cloud with depletion in its cold interior*

The first interpretation was coined the 'Baked Alaska' model by Rathborne et al. (2014a) and was conceived in an attempt to explain the profiles of molecular emission lines observed throughout G0.253+0.016. Conceptually, it is easiest to think of the Baked Alaska model as an adjustment to the classic blue-shifted infall model (see e.g. Evans 1999; Smith et al. 2012 for intuitive diagrammatic explanations). In this idealized picture, an asymmetric self-absorbed line profile occurs in emission lines with high opacity due to the inside-out collapse of a molecular cloud or a core. If the core exhibits a density and temperature gradient such that the excitation temperature increases inwards, emission from the centre can be absorbed by the low-excitation outer envelope, producing a double peaked emission line profile with an emission dip at the centroid velocity of the core. The blue asymmetry (i.e. where the blue peak appears brighter than the red peak) is due to the high excitation point in the red peak being obscured by the lower excitation point (as only the $\tau = 1$ surface is observed). Consequently, a double-peaked line profile with a blue asymmetry in an optically thick line can often be interpreted as a signature of collapse.

In five positions selected by Rathborne et al. (2014a), the line profiles of the optically thick species (e.g. $HCO^+$, HCN, $N_2H^+$) showed redshifted asymmetry (i.e. the opposite of the blue-shifted infall model). Rathborne et al. (2014a) argue that one way to create a redshifted asymmetry would be to invoke the same model, but with a cloud that is externally heated (as is observed in dust emission in G0.253+0.016; Longmore et al. 2012) such that the excitation temperature actually decreases towards the centre (hence the name 'Baked Alaska'). A schematic explanation of this idea is presented in fig. 9 of Rathborne et al. (2014a).[19]

A key to the interpretation of collapse in the aforementioned idealized model is that optically thin tracers peak at the location of the self-absorbed dip in emission in optically thick tracers (see e.g. Contreras et al. 2018, for a recent example). This is crucial because multiple spectral components in optically thin lines may simply indicate the presence of additional cloud components along the line of sight.

Herein lies the problem with G0.253+0.016: the lines that are believed to be optically thin (e.g. $H^{13}CO^+$, $HN^{13}C$) also show a double peak towards the cloud interior. Rathborne et al. (2014a) argue that a plausible explanation for the double peaked optically thin lines is that, if the lines are not optically thick, there must be severe, parsec-scale, chemical gas depletion of molecules in the cloud's high density and low temperature interior. One proposed line of evidence in favour of the aforementioned scenario is that there is an observed anticorrelation between the dust column density and the integrated intensity of various molecular lines towards the centre of the cloud. This gave rise to the interpretation that G0.253+0.016 is a single, coherent, centrally condensed cloud with depletion in its cold interior.

In a later publication, Rathborne et al. (2015) reported a tendency for molecular transitions with higher excitation energies and critical densities to peak towards the centre of the cloud, consistent, they argue, with a cloud with a dense interior. Fig. 8 of Rathborne et al.

[19]Note that another way to create a redshifted asymmetry would be to invoke expansion motions rather than collapse.

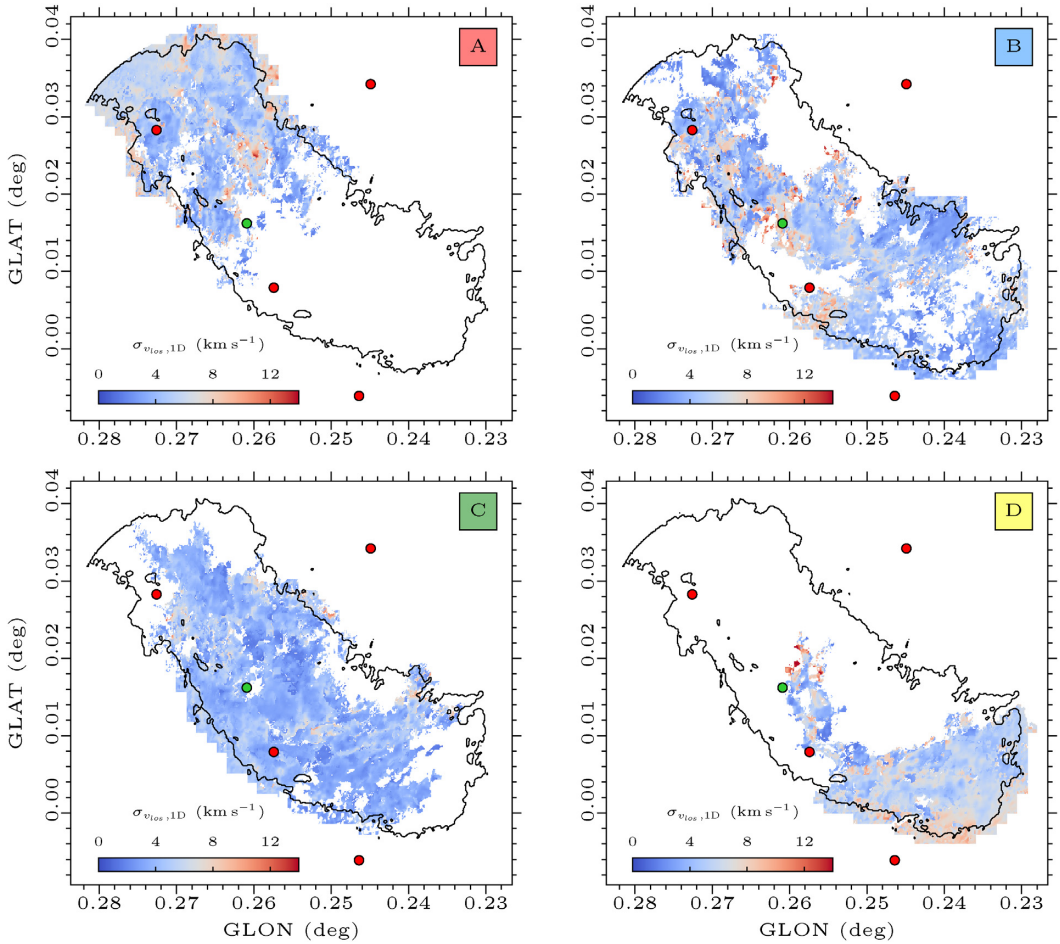


**Figure 12.** Equivalent to Fig. 7 but for the velocity dispersion, $\sigma_{v_{los},1D}$, measured throughout each ACORNS tree.

(2015) shows that PV diagrams of the emission associated with a variety of different molecules, including $C_2H$, SiO, $HN^{13}C$, $H^{13}CN$, $H^{13}CO^+$, $HCC^{13}CN$, SO, $NH_2CHO$, $CH_3CHO$, and $H_2CS$, have similar profiles to that shown in PPV space in Fig. 2 (i.e. two dominant features separated by $\sim$20 km s$^{-1}$ in the north of the cloud that merge in velocity towards to south).

However, while 2 out of the 17 molecules discussed by Rathborne et al. (2015) do display some emission towards the centre of the cloud ($CH_3CHO$ and $NH_2CN$), it is not extended and it does not peak exclusively in the central region. Rather, the emission qualitatively follows that of the other molecular transitions, but with a small peak towards the centre. Moreover, data from independent studies illustrate that the $\sim$20 km s$^{-1}$ gap between the dominant PPV features observed in Fig. 2 is not populated with emission from nitrogen-bearing species such as $N_2H^+$ (Pound & Yusef-Zadeh 2018), which are less susceptible to freeze-out at high densities (Bergin & Tafalla 2007).

The fact that the difference in velocity between the dominant components is largest towards the north of the cloud may also be problematic for this scenario. First, when observed at higher resolution, lines that are presumed to be optically thin show multiply peaked line profiles towards the north and south of the cloud (cf. the singly peaked profiles in the schematic diagram presented by Rathborne et al. 2014a). Secondly, the greatest velocity difference is observed towards the north of the cloud, where we find trees A, B, and C. In the context of widespread depletion, this would necessitate either a density or temperature gradient in G0.253+0.016. Furthermore it would suggest that either the highest density, or alternatively, lowest temperatures, are observed in the north of the cloud (where the absolute difference in the velocity peaks is the greatest; $\sim$35 km s$^{-1}$; Section 4.3.1). Studies of the dust continuum, and therefore the inferred $H_2$ column density towards G0.253+0.016 show no evidence for such a density gradient (Longmore et al. 2012; Johnston et al. 2014; Rathborne et al. 2015). Additionally, although the highest temperatures (>150 K) in G0.253+0.016 are found towards the south of the cloud (i.e. towards tree D), warm gas temperatures (80–100 K) are also found in the north (and generally distributed throughout; Ginsburg et al. 2016; Krieger et al. 2017). There is no clear and monotonic trend in decreasing gas temperature towards the north of the cloud.

It is worth noting that probably the strongest case for complete depletion of molecules within an individual cloud core (although it is yet to be confirmed) comes from Cyganowski et al. (2014). However, this occurs on <1000 AU scales where densities and temperatures are estimated to be $>10^9$ cm$^{-3}$ and $\lesssim$20 K, respectively. Although dust temperatures within G0.253+0.016 are of the order of $\sim$20 K (Longmore et al. 2012), the gas temperatures are actually considerably higher (of the order of $\gtrsim$60 K; Ginsburg et al. 2016; Krieger et al. 2017), consistent with the gas and

dust not being thermally coupled at the derived cloud density of $\sim 10^4$ cm$^{-3}$ (Clark et al. 2013). Therefore, without detailed chemical modelling it is currently difficult to reconcile the concept of parsec-scale depletion throughout the interior of a singular, coherent, and centrally condensed cloud with the absence of either an increasing density gradient or a decreasing temperature gradient towards the northern portion of G0.253+0.016 (as would be required to create the PPV profile observed in Fig. 2).

#### 5.1.2 Scenario 1b: Optically thick lines: G0.253+0.016 is a centrally concentrated cloud whose interior dynamics are masked due to high optical depth

Another conceivable scenario is that the lines that are often considered to be optically thin (e.g. $H^{13}CO^+$, $H^{13}CN$, and $HN^{13}C$), are actually optically thick. If this is the case then the double peaked profile in these lines may simply arise from self-absorption, with the individual peaks representing the outer 'shell' of the cloud at the $\tau = 1$ surface.

We assess the possibility of the HNCO $J = 4(0, 4) - 3(0, 3)$ line being optically thick using radiative transfer modelling. We adopt a kinetic temperature of 60 K (Ginsburg et al. 2016) and a fixed turbulent line width of 4.4 km s$^{-1}$ (i.e. $\langle \sigma_{v_{\rm los,1D}} \rangle$). We treat the molecular abundance and gas number density as free parameters, though the best estimate of the average number density is $10^4$ cm$^{-3}$ (Federrath et al. 2016) and the assumed canonical HNCO abundance is $10^{-9}$ (the typical abundance found towards dense cores, including those in the CMZ, by Churchwell et al. 1986 and Zinchenko, Henkel & Mao 2000).

We perform radiative transfer calculations using both the large velocity gradient (LVG) approximation and a 3D model evaluated on a 1D grid. The LVG approximation assumes that each emitting position in the cloud can only be absorbed by adjacent material, since more distant material is Doppler shifted out of the emission line profile. For the geometric model, we consider a uniform density sphere of fixed radius 2.35 pc (to give a diameter, 4.7 pc, consistent with Federrath et al. 2016) evaluated on a 1D grid.[20] We employ the NLTE statistical equilibrium solver in the Monte Carlo radiation transport code TORUS (Rundle et al. 2010), which is similar to that of Hogerheijde & van der Tak (2000). This approach accounts for the 3D structure of the cloud by assuming spherical symmetry. The level populations are computed in each cell using either LTE or NLTE assumptions. In LTE the level populations are trivially calculated analytically using the Boltzmann distribution. The NLTE level populations are calculated iteratively. They are initialized to LTE conditions, then ray tracing is performed to determine the radiation field and recalculate the level populations. This process is repeated until level populations converge. To estimate the brightness temperature and optical depth, a ray at the line centre is traced through the centre of the sphere along the observers line of sight. All of the material is assumed to be centred on the same rest velocity with a constant 4.4 km s$^{-1}$ turbulent line width.

The resulting grid in the ray tracing approach, both in LTE and NLTE, is given in the upper two panels of Fig. 13. The single white point represents the canonical HNCO abundance and derived mean density of G0.253+0.016. The lower left-hand panel of Fig. 13 shows the NLTE result in the LVG RADEX calculations. In this panel the colour bar also represents the brightness temperature distribution and the black dotted contour in each plot denotes the region where $\tau = 1$.

In the ray tracing models, there is no component of the parameter space that is both optically thick and has a low enough brightness temperature to be consistent with the observed $T_{\rm B}$ distribution extracted using SCOUSEPY throughout G0.253+0.016 (see the bottom right-hand panel). In the LVG models, there is a very small region of the parameter space where a solution is possible (hatched contour; $\tau > 1$ and $T_{\rm B} < 5$ K, where this latter condition represents three standard deviations from the mean SCOUSEPY-derived brightness temperature $\sim$1.75 K). However, the abundance of HNCO would have to be enhanced above the value observed towards dense cores by Churchwell et al. (1986) and Zinchenko et al. (2000) by at least 1–2 orders of magnitude.

The above analysis comes with the caveats that our calculations assume spherical symmetry, as well as a uniform abundance, density, and temperature. For more realistic conditions, there may be localized regions within G0.253+0.016 where the line becomes optically thick. However, we conclude that, in the absence of independent evidence for an extremely elevated HNCO abundance, the line is likely to be optically thin throughout the majority of the cloud. Even if the abundance is highly elevated, there appears to be only specific, unlikely geometric and kinematic structures of the cloud consistent with the line being thick. We therefore conclude that the HNCO line can be used as a reliable tracer of the gas dynamics of G0.253+0.016, where its emission is widespread (as it is throughout the CMZ; Henshaw et al. 2016a).

### 5.2 Cloud–cloud collision hypothesis: G0.253+0.016 is a superposition of two molecular gas clouds undergoing a collision

#### 5.2.1 Scenario 2a: G0.253+0.016 formed following a cloud–cloud collision

It has been argued that collisions between either atomic or molecular gas clouds in the ISM may play a role in both the formation and/or agglomeration of clouds, and in the triggering of star formation events, particularly high-mass star and star cluster formation (see e.g. Dobbs et al. 2014 and references therein). Hence there is considerable interest in categorizing the observational characteristics of such phenomena. However, inferring cloud–cloud collisions from observations is challenging. Numerical simulations can give important insight to some of the characteristics of cloud collisions (Inoue & Fukui 2013; Haworth et al. 2015a,b), however, often these characteristics are not unique.

Higuchi et al. (2014) invoked cloud–cloud collisions as a possible formation mechanism for G0.253+0.016. The authors identified the presence of a shell (radius $\sim$1.3 pc) within G0.253+0.016, in addition to LVGs ($\sim$20 km s$^{-1}$ pc$^{-1}$) and broad velocity dispersions (of the order of 30–40 km s$^{-1}$). Comparing with simulations of cloud–cloud collisions Higuchi et al. (2014) conclude that the shell structure may have been caused by the collision between two clouds of different mass and radii, resulting in the formation of a dense cloud that we now observe as G0.253+0.016.

The shell structure identified is that which we identify as the 'C'-shaped arc belonging to tree B in Section 4.2. Our kinematic analysis reveals that the arc is exclusively associated with tree

[20] Note that the LVG calculation also assumes spherical symmetry.

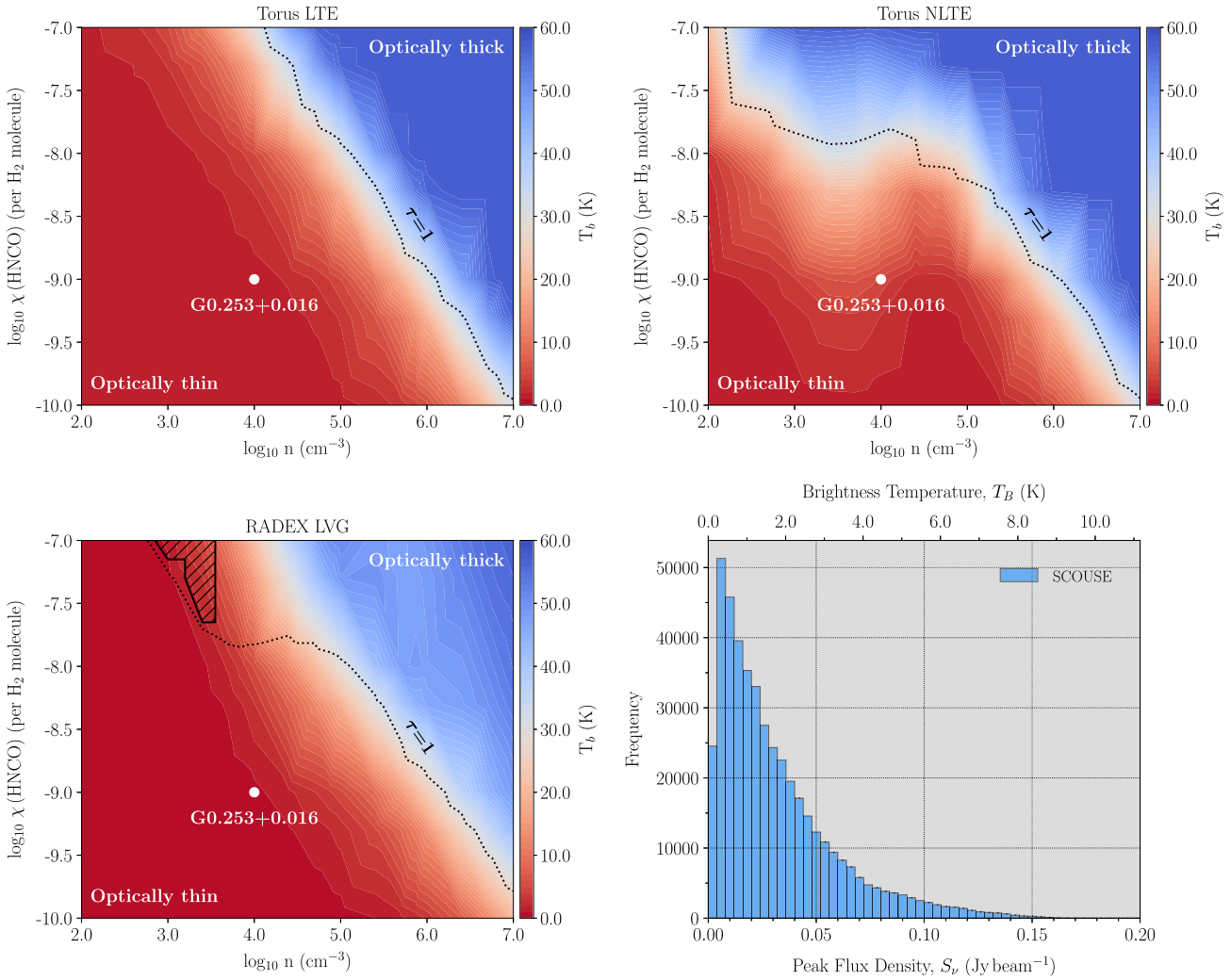


**Figure 13.** Top panels: The brightness temperature as a function of molecular abundance and the number density. The left-hand and right-hand panels follow the ray tracing approach using TORUS (Rundle et al. 2010) in LTE and NLTE, respectively. The dotted black contour denotes an optical depth of unity and the white point represents the likely conditions in G0.253+0.016. Bottom left-hand panel: The lower panel employs an LVG approach computed with RADEX (van der Tak et al. 2007). The hatched region denotes the region in which the medium is optically thick and the brightness temperature is less than 5 K. Bottom right-hand panel: A histogram of the peak flux density, $S_\nu$, of all spectral components extracted using SCOUSEPY. The $x$-axes are given in both Jy beam$^{-1}$ (bottom) and K (top) using a conversion factor of 55.8 K (Jy beam$^{-1}$)$^{-1}$.

B. The fact that this feature only accounts for a small fraction of the total HNCO emission observed throughout G0.253+0.016 (roughly ∼17 per cent of all fitted components) indicates that it is unlikely a relic of the cloud formation process. While we cannot rule out the possibility that G0.253+0.016 has formed via a cloud–cloud collision, based on our combined SCOUSEPY and ACORNS decomposition, we dispute that the presence of the arc is residual evidence of the formation process of the cloud as a whole as hypothesized by Higuchi et al. (2014). More generally, it is unclear whether cloud–cloud collisions occur frequently enough, and on a short enough time-scale, for them to be a dominant physical mechanism in the formation of clouds (Jeffreson et al. 2018; Jeffreson & Kruijssen 2018). Instead, it has recently been suggested that large-scale instabilities may provide a plausible mechanism for the formation of massive and dense molecular clouds in the CMZ, both in observations (Henshaw et al. 2016b) and in simulations (Sormani et al. 2018).

### *5.2.2 Scenario 2b: G0.253+0.016 is currently undergoing a cloud–cloud collision*

The concept of a cloud–cloud collision in G0.253+0.016 is not new. It was first proposed by Lis & Menten (1998; and further expanded by Lis et al. 2001) as a possible explanation for both the presence of multiple line-of-sight velocity components and the observed widespread emission from shocked gas tracers (see also Kauffmann et al. 2013). Lis et al. (2001) argued that the collision occurs between a molecular gas component observed at ∼20 km s$^{-1}$ (cf. tree B) and another at ∼40 km s$^{-1}$ (cf. tree C).

Rathborne et al. (2014a) postulated that for a cloud collision one may expect to observe two velocity components and a central zone of hot and shocked gas at the collision interface. The authors point out that while multiple velocity components are indeed observed in the dense gas tracers in single-dish observations, the same is true for those tracing hot and shocked gas. The hot and shocked

gas tracers (such as SiO) are not isolated to a single region within the cloud. Instead they have a similar distribution and kinematic profile to the optically thin gas tracers. In the absence of a specific collision region, Rathborne et al. (2014a) conclude that the single cloud interpretation is more consistent with their observations (Section 5.1.1).

Using high-spatial resolution interferometric observations however, Johnston et al. (2014) identified the presence of shocked gas tracers and elevated gas temperatures towards the southern portion of the cloud. This emission spatially coincides with our tree D. In investigating the kinematics, the authors noted that there is an additional velocity component, situated at $\sim$70 km s$^{-1}$ which is spatially coincident with the emission from shocked gas. These two velocity components 'connect' in PV space, which led Johnston et al. (2014) to suggest that they may be interacting.

This latter possibility was discussed by Henshaw et al. (2016a) who compared the observed kinematics of the CMZ with three different geometries aiming to describe the 3D structure of the CMZ. Henshaw et al. (2016a) concluded, albeit using much coarser spatial resolution observations (1 arcmin $\sim$ 2.4 pc) than Johnston et al. (2014), that the component observed at $\sim$70 km s$^{-1}$ is unlikely to be associated with G0.253+0.016. The emission from the $\sim$70 km s$^{-1}$ component is morphologically distinct from that of G0.253+0.016 (despite overlapping in projection) and is more extended (with projected extent $>$150 pc), appearing to connect to the molecular clouds closest in projection to Sgr A$*$ (i.e. the 20 and 50 km s$^{-1}$ clouds). In each of the model geometries discussed by Henshaw et al. (2016a), the 70 km s$^{-1}$ component is unrelated to G0.253+0.016. Given the observational evidence that is currently available, we therefore conclude that G0.253+0.016 and the 70 km s$^{-1}$ velocity component are most likely spatially distinct and non-interacting.

Despite the aforementioned discrepancy with the 70 km s$^{-1}$ component, we cannot rule out the possibility of interaction between subclouds within G0.253+0.016. The location of elevated gas temperatures and shocked gas emission identified by Johnston et al. (2014) is spatially coincident with our tree D, which sits at the interface of trees B and C (caution: in PPV space). Indeed, our analysis shows that this location in tree D displays an enhancement of HNCO emission (referred to as the 'tiled bar' in Mills et al. 2015; cf. the bottom right-hand panel of Fig. 7). Moreover, Fig. 11 demonstrates that velocity dispersions ($\sigma_{v_{\rm los},1{\rm D}}$) measured within tree D are on average greater than those measured throughout the other identified components. This could indicate that the interaction of substructure within G0.253+0.016 may play an important role in setting the internal dynamics of the cloud as well as its appearance in shocked gas tracers (see also Lis et al. 2001; Kauffmann et al. 2013).

## 6 THE ACORNS VIEW OF G0.253+0.016: 'THE BRICK' IS NOT A *BRICK*

The kinematic analysis presented in Sections 3 and 4 provides new and unique insight into the structure of G0.253+0.016 and the physical processes that are important (or unimportant) in shaping its appearance. The discussions presented in Sections 5.1.1 and 5.1.2 enable us to conclude that, globally, emission from the HNCO $J = 4(0, 4) - 3(0, 3)$ transition $\sim$3 mm is both likely to be optically thin and not widely depleted. Consequently HNCO is likely a reliable tracer of the internal structure and dynamics of the cloud. Our interpretation is therefore that, rather than a single, coherent, centrally condensed cloud with depletion in its cold interior, G0.253+0.016 is a complex, hierarchically structured molecular cloud exhibiting an intricate network of velocity components situated along the line of sight. 'The Brick' is not a *brick*.

Despite the aforementioned interpretation, one should always approach the connection between PPV space and true physical 3D space with caution (as demonstrated by e.g. Beaumont et al. 2013; Clarke et al. 2018). However, both the arc (top right-hand panel in Fig. 7) and the 'tilted bar' (bottom right) have both been identified in earlier works on G0.253+0.016, in a variety of molecular lines (Higuchi et al. 2014; Mills et al. 2015). ACORNS has uniquely provided the first evidence that these features: (i) were also present in data sets in which they had previously not been identified, but were simply masked by the kinematic complexity of the data; and (ii) are coherent in both (projected) space, velocity, and velocity dispersion. A key result of our analysis therefore is that ACORNS has blindly identified structures that appear to be both physically meaningful and statistically different from one another, evident through their morphologically distinct emission features (Fig. 7) as well as their differing internal dynamics (Figs 3 and 11).

So what is shaping the structure of the cloud? It is likely that G0.253+0.016 is a product of its complex and dynamic environment. A key result of recent hydrodynamical simulations is that the small-scale cloud structure of G0.253+0.016 is consistent with the cloud being sculpted by the Galactic dynamics of the CMZ (Dale, Kruijssen & Longmore 2019; Kruijssen et al. 2019), but see also the simulations of Sormani et al. (2018), where gas clouds are clearly influenced by orbital dynamics. A side-by-side comparison between the dust continuum observations presented in Fig. 8 (Rathborne et al. 2014b) and simulated ALMA observations of clouds orbiting the Galactic centre gives good qualitative agreement in terms of global morphology and the complex spatial structure of G0.253+0.016 (see fig. 6 in Kruijssen et al. 2019). These simulations demonstrate that high column densities, global velocity gradients, flattened cloud morphology, and inclination on the plane of the sky all naturally occur as a result of the influence of the background gravitational potential and shearing motions induced by eccentric orbits.

In addition to the large-scale orbital dynamics that may shape the cloud structure, there may be further external factors that play a significant role in shaping the structure and evolution of G0.253+0.016. In Fig. 14, we show a zoom of the three-colour *Spitzer* GLIMPSE image in Fig. 1, however, here we highlight some of the main features in G0.253+0.016's surrounding environment. As can be seen, the cloud overlaps in projection with the prominent supernova remnant G0.30+0.00 (also, G000.3+00.0, G0.33+0.04, G0.4+0.1; red ellipse; Kassim & Frail 1996; LaRosa et al. 2000). Additionally, Ponti et al. (2015) identify another supernova remnant candidate, G0.224+0.032,[21] located directly to the (Galactic) west of G0.253+0.016 (dashed green ellipse). The high extinction of the cloud means that the soft X-ray emission is partially obscured by the cloud. Nevertheless, Ponti et al. (2015) argue that the properties of G0.224+0.032 are consistent with those of a supernova remnant, but that the true size and energy are difficult to estimate due to the obscuration.

Also indicated in Fig. 14 are the positions of massive stars located towards G0.253+0.016. The filled cyan points indicate the locations of Paschen $\alpha$ emitting sources obtained with the *Hubble Space Telescope*/Near-Infrared Camera and Multi-Object Spectrometer and Multi-Object Spectrometer (*HST*/NICMOS) identified by Dong et al. (2011) and the filled red circles indicate the locations of Wolf–Rayet stars, O supergiants, and B supergiants, obtained by

[21] Note this is labelled as G0.224–0.032 in Ponti et al. (2015).

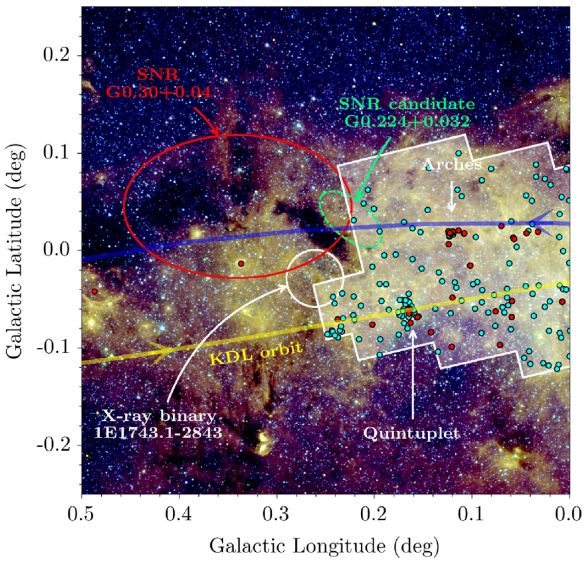


**Figure 14.** A close up view of the immediate environment of G0.253+0.016. The background image is equivalent to Fig. 1, but here we display some of the additional external factors that may play a significant role in the evolution of G0.253+0.016. The red ellipse highlights the prominent supernova remnant, G0.30+0.04 (Kassim & Frail 1996; LaRosa et al. 2000) and the dashed green ellipse is a supernova remnant *candidate*, G0.224+0.032 (Ponti et al. 2015). The white circle highlights the location of X-ray binary 1E1743.1 − 2843 (Porquet et al. 2003). The size of the circle corresponds approximately to the spatial extent of the emission observed with *XMM−Newton* (see Ponti et al. 2015). The filled cyan points indicate the locations of Paschen $\alpha$ emitting sources obtained with the *HST* (Dong et al. 2011). The white shaded region shows the footprint of the *HST* observations. The filled red circles indicate the locations of Wolf-Rayet stars, O supergiants, and B supergiants, obtained by Mauerhan et al. (2010). Finally, the blue (near side) and yellow (far side) lines indicate the orbital model of the CMZ derived by Kruijssen et al. (2015), with the arrows depicting the direction of gas motion.

Mauerhan et al. (2010). Dong et al. (2011) argue that the majority of these sources are most likely evolved massive stars ($M_* > 7\,\mathrm{M}_\odot$) with strong stellar winds. The source locations are categorized into four different groups: (i) & (ii) those associated with the young massive clusters the Arches and Quintuplet; (iii) those located with the nuclear star cluster; (iv) and field sources outside the main clusters. Although the footprint of the observations does not include G0.253+0.016, there is a considerable number of field massive stars spread throughout the observed region. Feedback from such massive stars has the potential to influence the molecular gas in this environment. Indeed, it has been argued that the O4-6 supergiant, which is situated immediately to the (Galactic) east of G0.253+0.016 (that which lies within the boundary of the red ellipse in Fig. 14), may be responsible for the ionization of the exterior of the cloud in this direction (Mills et al. 2015).

Although projection effects may play a role in determining whether or not these features indeed influence the structure of G0.253+0.016; the fact remains that G0.253+0.016 displays complex internal dynamics, both in terms of velocity gradients and supersonic velocity dispersions, as well as elevated gas temperatures, and a prevalence of emission from tracers of shocked gas. The complex interplay of these large-scale (e.g. Galactic dynamics) and comparatively small-scale (e.g. feedback) effects may all contribute in sculpting the physical structure of G0.253+0.016, and therefore its star formation potential.

## 7 CONCLUSIONS

We have performed a comprehensive study of the dynamics and physical structure of G0.253+0.016. To facilitate this study we have developed two pieces of software, both of which we make available to the community. The first, SCOUSEPY, is a redevelopment of the spectral line fitting algorithm first presented by Henshaw et al. (2016a). The second, ACORNS, is a hierarchical clustering algorithm designed specifically for use with discrete data such as that output by SCOUSEPY. Combined, these algorithms have helped us to develop a new view of G0.253+0.016. Our main conclusions are summarized below.

(i) We have performed a full kinematic decomposition of the HNCO ALMA data, quantifying and measuring the spectral lines despite their well-known complexity. Globally, the kinematic structure of G0.253+0.016 appears to show two dominant features in PPV space, one situated at ∼35–50 km s$^{-1}$ and another that ranges from ∼0–30 km s$^{-1}$. Both features have global velocity gradients in the north–south direction (in equatorial coordinates or north east-south west in Galactic) following the major axis of the cloud. However, the magnitude of the velocity gradient across the latter feature is about a factor of ∼2 greater than that across the former. This presents a more complex picture than that of a singular cloud exhibiting the hallmarks of rotation as has been suggested in previous works (e.g. Rathborne et al. 2014a; Federrath et al. 2016).

(ii) A striking feature of our SCOUSEPY decomposition is the 'wiggly' nature of the kinematic substructure. Oscillatory velocity gradients appear ubiquitously throughout the cloud. However, unlike those identified on larger scales (Henshaw et al. 2016b), which display a characteristic wavelength and amplitude, these oscillations appear more stochastic. We will quantify these oscillations further in a future publication (Henshaw et al., in preparation).

(iii) Velocity dispersions measured along the line of sight (extracted directly from spectral line fitting) are a factor of ∼2 below those derived from moment analysis due to the presence of multiple velocity components identified within the spectra. On average we measure $\langle\sigma_{v_{\rm los},1\rm D}\rangle = 4.4\,\mathrm{km\,s^{-1}}$, with a standard deviation of $2.1\,\mathrm{km\,s^{-1}}$. Assuming a fixed temperature of ∼60 K, this translates into a Mach number estimate of $\mathcal{M}_{\sigma_{v_{\rm los},3\rm D}} \sim 16.5$. Although these velocity dispersions are broader than those predicted from the steep linewidth–size relationships of Shetty et al. (2012) and Kauffmann et al. (2017a), these results add to mounting evidence for the existence of narrow ($\lesssim$a few km s$^{-1}$) lines on small spatial scales in CMZ clouds.

(iv) ∼98 per cent of the SCOUSEPY decomposition data are clustered using ACORNS. We find that the dynamics are dominated by four main features containing > 50 per cent of the data.

(v) There are important differences between the four main hierarchical structures (referred to as 'trees'). The dominant tree (C), situated at a mean velocity of $\langle v\rangle \sim 37.0\,\mathrm{km\,s^{-1}}$, is most similar to the intensity distribution observed in dust continuum observations giving the cloud its physical appearance as we observe it on the plane of the sky. Tree B ($\langle v\rangle \sim 16.5\,\mathrm{km\,s^{-1}}$) exhibits a prominent arc shaped feature that has been noted in previous studies (Higuchi et al. 2014; Mills et al. 2015). Out of the two smaller trees, D ($\langle v\rangle \sim 33.1\,\mathrm{km\,s^{-1}}$) displays a prominent linear feature associated with elevated gas temperatures and velocity dispersions. Finally, tree A ($\langle v\rangle \sim 2.9\,\mathrm{km\,s^{-1}}$) extends towards the north of the cloud in the direction of dust ridge cloud 'b' which has a similar velocity $\sim 3.4\,\mathrm{km\,s^{-1}}$ (Henshaw et al. 2016a). While many of these features have been identified previously in the literature, a key and unique element of our analysis is that ACORNS

provides the first evidence that these features are coherent in both (projected) space and velocity. Moreover, ACORNS has extracted these features blindly from the observational data. This indicates that these features were already present in data such as the HNCO emission initially presented by Rathborne et al. (2015), but were masked by the kinematic complexity of the cloud.

(vi) We compare the trees' mean line-of-sight velocity dispersions with the fluctuations in the centroid velocity across the plane of the sky, finding $\langle\sigma_{v_{\rm los,1D}}\rangle = \{5.3, 4.9, 4.0, 5.8\}{\rm km\,s^{-1}}$ and $\sigma_{v_{\rm pos,1D}} = \{3.5, 5.2, 4.6, 4.5\}{\rm km\,s^{-1}}$, respectively. The ratio of these two measurements yields $\langle\sigma_{v_{\rm los,1D}}/\sigma_{v_{\rm pos,1D}}\rangle = 1.2 \pm 0.3$. We speculate that this isotropy in the velocity fluctuations may contain important information regarding the cloud geometry. Namely, that the line-of-sight extent of the cloud components are approximately equivalent to that in the plane of the sky.

(vii) We argue that emission from the $J = 4(0, 4) - 3(0, 3)$ transition of HNCO is (globally) optically thin, and therefore is a good tracer of the internal dynamics of the cloud overall. We disfavour the interpretation that G0.253+0.016 is a centrally condensed molecular cloud with depletion in its cold interior, as was proposed by Rathborne et al. (2014a), since the PPV profile would necessitate either a strong increasing gradient in density from south to the north of the cloud (or alternatively decreasing temperature), which is not observed.

(viii) We do not rule out the possibility that the merger of substructures within G0.253+0.016 may play an important role in producing shocked gas emission, elevating the gas temperature, and raising the velocity dispersion of the gas. However, we dispute the conclusion of Higuchi et al. (2014) that the arc emission feature is evidence that G0.253+0.016 has formed via cloud–cloud collisions. Our kinematic analysis demonstrates that emission from the arc feature is just a small fraction of the total cloud emission. Therefore, it is unlikely that this is a relic signature of the formation mechanism of the cloud as a whole.

(ix) Finally, we discuss our findings in the context of the large-scale kinematics of the CMZ. G0.253+0.016 is a complex, hierarchically structured molecular cloud exhibiting an intricate network of velocity components situated along the line of sight; 'the Brick' is not a *brick*. We argue that the morphology is most likely a product of the tangled interplay of both Galactic dynamics and feedback present in the CMZ. Recent simulations of molecular clouds orbiting galactic centres indicate that complex cloud structure is a natural outcome of the influence of the background gravitational potential and shearing motions induced by eccentric orbits (Sormani et al. 2018; Dale et al. 2019; Kruijssen et al. 2019). Detailed kinematic analysis of such simulations is highly promising for further constraining the physical mechanisms shaping molecular cloud structure within the CMZ.

In the near future, studies such as the CMZoom survey (the Sub-Millimeter Array's legacy survey of the CMZ; Battersby et al. 2017, Battersby et al., in preparation) as well as future ALMA surveys will facilitate a uniform description of molecular cloud dynamics throughout the CMZ. This will help to provide a statistical understanding of the earliest phases of star formation in this complex and dynamic environment.

## FACILITIES AND DATA

This research has made use of NASA's Astrophysics Data System. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2011.0.00217.S. ALMA is a partnership of ESO (representing its member states), NSF (USA), and NINS (Japan), together with NRC (Canada) and NSC and ASIAA (Taiwan), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO, and NAOJ.

## SOFTWARE

This research made use of

(i) ACORNS (https://github.com/jdhenshaw/acorns)
(ii) ASTROPY (Astropy Collaboration 2013)
(iii) MATPLOTLIB (Hunter 2007)
(iv) NUMPY (Van Der Walt, Colbert & Varoquaux 2011)
(v) PYSPECKIT (http://pyspeckit.bitbucket.org)
(vi) SCIKIT-LEARN (Pedregosa et al. 2011)
(vii) SCIPY (Jones et al. 2001)
(viii) SCOUSEPY (https://github.com/jdhenshaw/scousepy)
(ix) SPECTRALCUBE (part of the radio-astro-tools package hosted at http://radio-astro-tools.github.io/)

## ACKNOWLEDGEMENTS

We thank the referee for their careful reading of the manuscript and comments. We would like to thank Jill Rathborne for making the data product used in this paper available for use. Additionally, JDH would like to thank Paola Caselli, Seamus Clarke, Christoph Federrath, Morgan Fouesneau, Iskren Georgiev, John Ilee, Dimitry Semenov, and Juan Soler, for fruitful discussions during the preparation of this paper.

ATB acknowledges funding from the European Union's Horizon 2020 research and innovation programme (grant agreement No. 726384). CB gratefully acknowledges support by the National Science Foundation under Grant No. 1816715. HB acknowledges support from the European Research Council under the Horizon 2020 Framework Program via the ERC Consolidator Grant CSF-648505. TJH is funded by an Imperial College junior research fellowship. JMDK gratefully acknowledges funding from the German Research Foundation (DFG) in the form of an Emmy Noether Research Group (grant number KR4801/1-1) and from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme via the ERC Starting Grant MUSTANG (grant agreement number 714907). MR gratefully acknowledges funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 639459 (PROMISE).

## REFERENCES

Akaike H., 1974, IEEE Trans. Autom. Control, 19, 716
Anderson T. W., Darling D. A., 1952, Ann. Math. Stat., 23, 193
André P., Di Francesco J., Ward-Thompson D., Inutsuka S.-I., Pudritz R. E., Pineda J. E., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. University of Arizona Press, Tucson, p. 27
Anselmet F., Gagne Y., Hopfinger E. J., Antonia R. A., 1984, J. Fluid Mech., 140, 63
Ao Y. et al., 2013, A&A, 550, A135
Astropy Collaboration, 2013, A&A, 558, A33
Bally J., Stark A. A., Wilson R. W., Henkel C., 1987, ApJS, 65, 13
Bally J., Stark A. A., Wilson R. W., Henkel C., 1988, ApJ, 324, 223
Bally J. et al., 2014, ApJ, 795, 28
Barnes A. T., Longmore S. N., Battersby C., Bally J., Kruijssen J. M. D., Henshaw J. D., Walker D. L., 2017, MNRAS, 469, 2263

Battersby C. et al., 2017, in Crocker R. M., Longmore S. N., Bicknell G. V., eds, IAU Symp. Vol. 322, The Multi-Messenger Astrophysics of the Galactic Centre. Cambridge Univ. Press, Cairns, Australia, p. 90
Beaumont C. N., Offner S. S. R., Shetty R., Glover S. C. O., Goodman A. A., 2013, ApJ, 777, 173
Bergin E. A., Tafalla M., 2007, ARA&A, 45, 339
Berry D. S., 2015, Astron. Comput., 10, 22
Churchwell E., Wood D., Myers P. C., Myers R. V., 1986, ApJ, 305, 405
Churchwell E. et al., 2009, PASP, 121, 213
Clark P. C., Glover S. C. O., Ragan S. E., Shetty R., Klessen R. S., 2013, ApJ, 768, L34
Clarke S. D., Whitworth A. P., Spowage R. L., Duarte-Cabral A., Suri S. T., Jaffa S. E., Walch S., Clark P. C., 2018, MNRAS, 479, 1722
Colombo D., Rosolowsky E., Ginsburg A., Duarte-Cabral A., Hughes A., 2015, MNRAS, 454, 2067
Contreras Y. et al., 2018, ApJ, 861, 14
Cyganowski C. J. et al., 2014, ApJ, 796, L2
D'agostino R. B., Belanger A., D'agostino R. B., Jr., 1990, Am. Stat., 44, 316
Dahmen G. et al., 1997, A&AS, 126, 197
Dahmen G., Huttemeister S., Wilson T. L., Mauersberger R., 1998, A&A, 331, 959
Dale J. E., Kruijssen J. M. D., Longmore S. N., 2019, MNRASsubmitted
Dobbs C. L. et al., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. University of Arizona Press, Tucson, p. 3
Dong H. et al., 2011, MNRAS, 417, 114
Emsellem E., Renaud F., Bournaud F., Elmegreen B., Combes F., Gabor J. M., 2015, MNRAS, 446, 2468
Evans N. J., II, 1999, ARA&A, 37, 311
Federrath C., 2013, MNRAS, 436, 1245
Federrath C., Klessen R. S., 2012, ApJ, 761, 156
Federrath C. et al., 2016, ApJ, 832, 143
Figer D. F., Kim S. S., Morris M., Serabyn E., Rich R. M., McLean I. S., 1999, ApJ, 525, 750
Foster J. B. et al., 2011, ApJS, 197, 25
Ginsburg A. et al., 2016, A&A, 586, A50
Ginsburg A. et al., 2018, ApJ, 853, 171
Ginsburg A., Bressert E., Bally J., Battersby C., 2012, ApJ, 758, L29
Goodman A. A., Benson P. J., Fuller G. A., Myers P. C., 1993, ApJ, 406, 528
Goss W. M., Schwarz U. J., van Gorkom J. H., Ekers R. D., 1985, MNRAS, 215, 69P
Guesten R., Downes D., 1983, A&A, 117, 343
Guesten R., Henkel C., 1983, A&A, 125, 136
Hacar A., Tafalla M., Kauffmann J., Kovács A., 2013, A&A, 554, A55
Haworth T. J. et al., 2015a, MNRAS, 450, 10
Haworth T. J., Shima K., Tasker E. J., Fukui Y., Torii K., Dale J. E., Takahira K., Habe A., 2015b, MNRAS, 454, 1634
Hennebelle P., Chabrier G., 2011, ApJ, 743, L29
Henshaw J. D. et al., 2016a, MNRAS, 457, 2675
Henshaw J. D., Longmore S. N., Kruijssen J. M. D., 2016b, MNRAS, 463, L122
Higuchi A. E., Chibueze J. O., Habe A., Takahira K., Takano S., 2014, AJ, 147, 141
Hinz J. L., Rieke G. H., Yusef-Zadeh F., Hewitt J., Balog Z., Block M., 2009, ApJS, 181, 227
Hogerheijde M. R., van der Tak F. F. S., 2000, A&A, 362, 697
Houlahan P., Scalo J., 1992, ApJ, 393, 172
Huettemeister S., Wilson T. L., Bania T. M., Martin-Pintado J., 1993, A&A, 280, 255
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Immer K., Menten K. M., Schuller F., Lis D. C., 2012, A&A, 548, A120
Inoue T., Fukui Y., 2013, ApJ, 774, L31
Jackson J. M. et al., 2013, Publ. Astron. Soc. Aust., 30, e057
Jeffreson S. M. R., Kruijssen J. M. D., 2018, MNRAS, 476, 3688
Jeffreson S. M. R., Kruijssen J. M. D., Krumholz M. R., Longmore S. N., 2018, MNRAS, 478, 3380
Johnston K. G., Beuther H., Linz H., Schmiedeke A., Ragan S. E., Henning T., 2014, A&A, 568, A56
Jones E., Oliphant T., Peterson P., Others, 2001, SciPy: Open source scientific tools for Python. Available at: http://www.scipy.org/
Jones P. A. et al., 2012, MNRAS, 419, 2961
Kassim N. E., Frail D. A., 1996, MNRAS, 283, L51
Kauffmann J., Pillai T., Zhang Q., 2013, ApJ, 765, L35
Kauffmann J., Pillai T., Zhang Q., Menten K. M., Goldsmith P. F., Lu X., Guzmán A. E., 2017a, A&A, 603, A89
Kauffmann J., Pillai T., Zhang Q., Menten K. M., Goldsmith P. F., Lu X., Guzmán A. E., Schmiedeke A., 2017b, A&A, 603, A90
Kendrew S., Ginsburg A., Johnston K., Beuther H., Bally J., Cyganowski C. J., Battersby C., 2013, ApJ, 775, L50
Klessen R. S., 2000, ApJ, 535, 869
Koch E. W., Rosolowsky E. W., 2015, MNRAS, 452, 3435
Koepferl C. M., Robitaille T. P., Morales E. F. E., Johnston K. G., 2015, ApJ, 799, 53
Krieger N. et al., 2017, ApJ, 850, 77
Kruijssen J. M. D., Longmore S. N., 2013, MNRAS, 435, 2598
Kruijssen J. M. D., Longmore S. N., 2014, MNRAS, 439, 3239
Kruijssen J. M. D., Longmore S. N., Elmegreen B. G., Murray N., Bally J., Testi L., Kennicutt R. C., 2014, MNRAS, 440, 3370
Kruijssen J. M. D., Dale J. E., Longmore S. N., 2015, MNRAS, 447, 1059
Kruijssen J. M. D. et al., 2019, MNRAS, 484, 5734
Krumholz M. R., Kruijssen J. M. D., 2015, MNRAS, 453, 739
Krumholz M. R., McKee C. F., 2005, ApJ, 630, 250
Krumholz M. R., Kruijssen J. M. D., Crocker R. M., 2017, MNRAS, 466, 1213
Lada C. J., Lombardi M., Alves J. F., 2010, ApJ, 724, 687
Lada C. J., Forbrich J., Lombardi M., Alves J. F., 2012, ApJ, 745, 190
LaRosa T. N., Kassim N. E., Lazio T. J. W., Hyman S. D., 2000, AJ, 119, 207
Lis D. C., Menten K. M., 1998, ApJ, 507, 794
Lis D. C., Menten K. M., Serabyn E., Zylka R., 1994, ApJ, 423, L39
Lis D. C., Pety J., Phillips T. G., Falgarone E., 1996, ApJ, 463, 623
Lis D. C., Serabyn E., Zylka R., Li Y., 2001, ApJ, 550, 761
Longmore S. N. et al., 2012, ApJ, 746, 117
Longmore S. N. et al., 2013a, MNRAS, 429, 987
Longmore S. N. et al., 2013b, MNRAS, 433, L15
Longmore S. N. et al., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. University of Arizona Press, Tucson, p. 291
Longmore S. N. et al., 2017, MNRAS, 470, 1462
Lu X. et al., 2019, ApJ, 872, 171
Madau P., Dickinson M., 2014, ARA&A, 52, 415
Manning C. D., Raghavan P., Schütze H., 2008, Introduction to Information Retrieval. Cambridge Univ. Press, New York, NY, USA
Marsh K. A., Ragan S. E., Whitworth A. P., Clark P. C., 2016, MNRAS, 461, L16
Mauerhan J. C., Muno M. P., Morris M. R., Stolovy S. R., Cotera A., 2010, ApJ, 710, 706
Mehringer D. M., Yusef-Zadeh F., Palmer P., Goss W. M., 1992, ApJ, 401, 168
Mehringer D. M., Palmer P., Goss W. M., 1993, ApJ, 402, L69
Meier D. S., Turner J. L., 2005, ApJ, 618, 259
Miesch M. S., Scalo J., Bally J., 1999, ApJ, 524, 895
Mills E. A. C., Morris M. R., 2013, ApJ, 772, 105
Mills E. A. C., Butterfield N., Ludovici D. A., Lang C. C., Ott J., Morris M. R., Schmitz S., 2015, ApJ, 805, 72
Mills E. A. C., Ginsburg A., Immer K., Barnes J. M., Wiesenfeld L., Faure A., Morris M. R., Requena-Torres M. A., 2018, ApJ, 868, 7
Miville-Deschênes M.-A., Murray N., Lee E. J., 2017, ApJ, 834, 57
Molinari S. et al., 2011, ApJ, 735, L33
Offner S. S. R., Clark P. C., Hennebelle P., Bastian N., Bate M. R., Hopkins P. F., Moraux E., Whitworth A. P., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. University of Arizona Press, Tucson, p. 53

Oka T., Hasegawa T., Sato F., Tsuboi M., Miyazaki A., Sugimoto M., 2001, ApJ, 562, 348
Oka T., Geballe T. R., Goto M., Usuda T., McCall B. J., 2005, ApJ, 632, 882
Ossenkopf V., Mac Low M.-M., 2002, A&A, 390, 307
Ott J., Weiß A., Staveley-Smith L., Henkel C., Meier D. S., 2014, ApJ, 785, 55
Padoan P., Nordlund, Å., 2011, ApJ, 730, 40
Pedregosa F. et al., 2011, J. Mach. Learn. Res., 12, 2825
Ponti G. et al., 2015, MNRAS, 453, 172
Porquet D., Rodriguez J., Corbel S., Goldoni P., Warwick R. S., Goldwurm A., Decourchelle A., 2003, A&A, 406, 299
Pound M. W., Yusef-Zadeh F., 2018, MNRAS, 473, 2899
Rathborne J. M. et al., 2014a, ApJ, 786, 140
Rathborne J. M. et al., 2014b, ApJ, 795, L25
Rathborne J. M. et al., 2015, ApJ, 802, 125
Reid M. J. et al., 2014, ApJ, 783, 130
Ridley M. G. L., Sormani M. C., Treß R. G., Magorrian J., Klessen R. S., 2017, MNRAS, 469, 2251
Rodríguez L. F., Zapata L. A., 2013, ApJ, 767, L13
Rosolowsky E., Leroy A., 2006, PASP, 118, 590
Rosolowsky E. W., Pineda J. E., Kauffmann J., Goodman A. A., 2008, ApJ, 679, 1338
Rundle D., Harries T. J., Acreman D. M., Bate M. R., 2010, MNRAS, 407, 986
Sawada T., Hasegawa T., Handa T., Cohen R. J., 2004, MNRAS, 349, 1167
Shetty R., Beaumont C. N., Burton M. G., Kelly B. C., Klessen R. S., 2012, MNRAS, 425, 720
Sidorin V., 2017, Quickclump: Identify clumps within a 3D FITS datacube, Astrophysics Source Code Library , record ascl:1704.006
Simpson J. P., 2018, ApJ, 857, 59
Smith R. J., Shetty R., Stutz A. M., Klessen R. S., 2012, ApJ, 750, 64
Sofue Y., 1995, PASJ, 47, 527
Sormani M. C., Barnes A. T., 2019, MNRAS, 484, 1213
Sormani M. C., Treß R. G., Ridley M., Glover S. C. O., Klessen R. S., Binney J., Magorrian J., Smith R., 2018, MNRAS, 475, 2383
Sousbie T., 2011, MNRAS, 414, 350
Sousbie T., Pichon C., Kawahara H., 2011, MNRAS, 414, 384
Stutzki J., Guesten R., 1990, ApJ, 356, 513
Swinbank A. M., Smail I., Sobral D., Theuns T., Best P. N., Geach J. E., 2012, ApJ, 760, 130
Tackenberg J. et al., 2012, A&A, 540, A113
Urquhart J. S. et al., 2014, MNRAS, 443, 1555
van der Tak F. F. S., Black J. H., Schöier F. L., Jansen D. J., van Dishoeck E. F., 2007, A&A, 468, 627
Van Der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
Walker D. L. et al., 2018, MNRAS, 474, 2373
Williams J. P., Blitz L., Stark A. A., 1995, ApJ, 451, 252
Wünsch R., Jáchym P., Sidorin V., Ehlerová S., Palouš J., Dale J., Dawson J. R., Fukui Y., 2012, A&A, 539, A116
Yusef-Zadeh F., Muno M., Wardle M., Lis D. C., 2007, ApJ, 656, 847
Yusef-Zadeh F. et al., 2009, ApJ, 702, 178
Zinchenko I., Henkel C., Mao R. Q., 2000, A&A, 361, 1079

## APPENDIX A: SCOUSEPY DECOMPOSITION

Here we include additional information regarding the SCOUSEPY fitting procedure. The left-hand panel of Fig. A1 displays the result of the new implementation for setting variable (SAA) sizes based on spectral complexity. The procedure is outlined in Section 3.1. Briefly however, we plot a map of $\Delta v_{\rm m} \equiv |v_1 - v_{\rm peak}| \sim 0$, where $v_1$ is the first-order moment and $v_{\rm peak}$ is the velocity of the channel containing the peak emission. We also plot a histogram of the individual pixel values. In the case of the ALMA HNCO observations of G0.253+0.016 we divide the data up into three logarithmically spaced $\Delta v_{\rm m}$ bins, which we use to define the size of our SAAs and are overlaid on the $\Delta v_{\rm m}$ map. This enables the user to pay close attention to regions that have line profiles with a greater degree of complexity. The right-hand panel of Fig. A1 highlights the locations that have best-fitting solutions, with each pixel being colour-coded according the number of velocity components

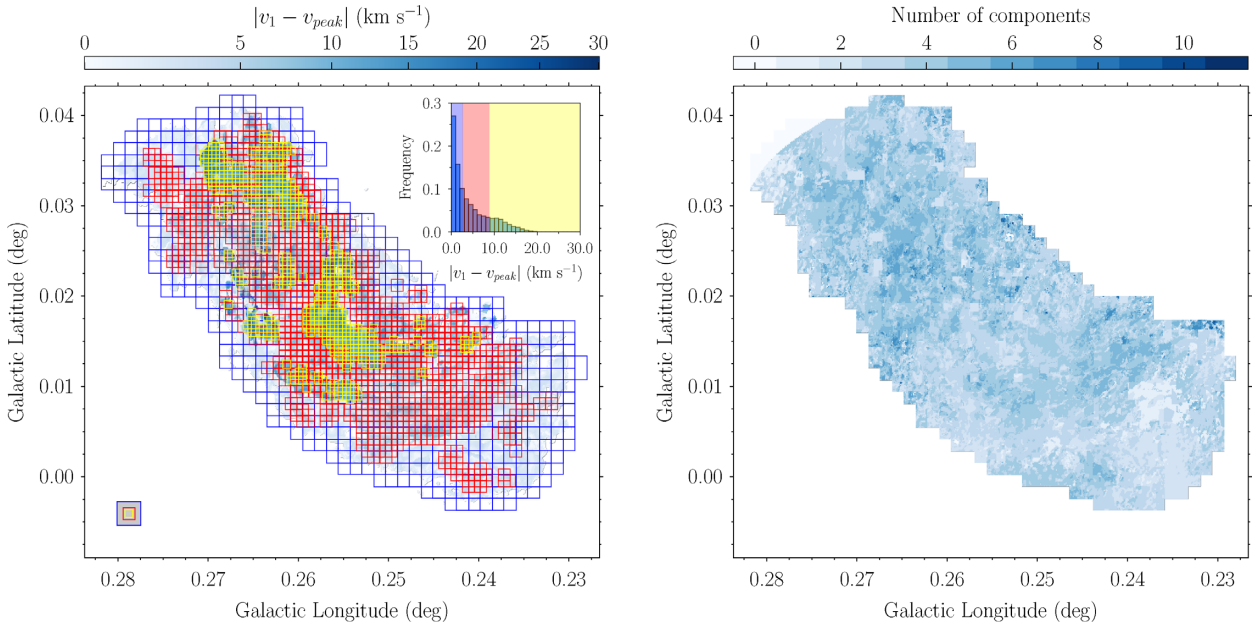


**Figure A1.** Left-hand panel: SCOUSEPY spectral averaging area (SAA) refinement based on spectral 'complexity'. The background image shows a map of the absolute difference between the intensity-weighted average velocity (moment 1; $v_1$) and the velocity at peak emission ($v_{\rm peak}$), which serves as a metric for spectral complexity ($\Delta v_{\rm m} = |v_1 - v_{\rm peak}| \sim 0.0$ for a singly peaked 'simple' spectrum). The inset image is a histogram of the pixel values. SCOUSEPY generates $n$ logarithmically spaced bins based on this complexity measure. The size of the SAA is then refined according to how complex the spectra are at a given location in the map (coloured boxes). Right-hand panel: The number of velocity components extracted by SCOUSEPY from each location in the map.

**Table A1.** SCOUSEPY: Global fitting statistics (see Section 3.1 for details).

| | Output statistic | Full resolution |
|---|---|---|
| | $N_{\rm tot}$ | 315 219 |
| | $N_{\rm tot,\,SAA}$ | 135 020 |
| | $N_{\rm SAA}$ | 2355 |
| | $N_{\rm SAArefine}$ | [527, 1141, 687] |
| | $N_{\rm fit}$ | 133 065 |
| | $N_{\rm comp}$ | 457 264 |
| | $N_{\rm comp}/N_{\rm fit}$ | 3.4 |
| | $N_{\rm orig}/N_{\rm fit}$ (per cent) | 96.4 |
| | $N_{\rm refit}/N_{\rm fit}$ (per cent) | 2.5 |
| | $N_{\rm alt}/N_{\rm fit}$ (per cent) | 1.1 |
| $N_{\rm tot}$ | Total number of pixels in the mapped area | |
| $N_{\rm tot,\,SAA}$ | Total number of pixels included in the coverage | |
| $N_{\rm SAA}$ | Total number of SAAs | |
| $N_{\rm SAArefine}$ | Number of SAAs at each level of refinement (see the text) | |
| $N_{\rm fit}$ | Total number of pixels fitted (automated) | |
| $N_{\rm comp}$ | Total number of components fitted | |
| $N_{\rm comp}/N_{\rm fit}$ | Mean number of Gaussian components per position | |
| $N_{\rm orig}/N_{\rm fit}$ | Percentage of original fits in the final data product | |
| $N_{\rm refit}/N_{\rm fit}$ | Percentage of refitted spectra in the final data product | |
| $N_{\rm alt}/N_{\rm fit}$ | Percentage of spectra with alternative solutions selected | |

identified at that location. Table A1 contains the statistics of our SCOUSEPY decomposition.

## APPENDIX B: EXTRACTING MOLECULAR GAS KINEMATICS FOLLOWING SPECTRAL DECOMPOSITION

As discussed in Section 1, the primary aim of this study is to obtain a detailed description of the molecular gas kinematics of G0.253+0.016. To date, analyses of the gas kinematics of G0.253+0.016 have predominantly relied on techniques such as moment analysis (Rathborne et al. 2015; Federrath et al. 2016), and dendrograms (Kauffmann et al. 2013). The former technique is beneficial as it is simple and fast to implement, and it returns information on the pixel scale. However, taking an intensity-weighted average velocity along the LO results in information being lost, particularly in regions with complex LOS density and velocity structure. Conversely, the latter technique is beneficial in that complex line-of-sight structure is accounted for as the algorithm seeks to build a hierarchy of structure, which can be represented graphically in the form of a dendrogram (see e.g. Rosolowsky et al. 2008). However, kinematic information is provided in the form of intensity-weighted average quantities relating to each structure. Further work is therefore required if one is interested in how those kinematic quantities vary with position within a given structure on the pixel scale.

More generally, there is an array of automatic algorithms whose primary function is to parse and extract information regarding the structure of molecular clouds and their internal dynamics. These include, but are not limited to, those designed to segment and extract isolated peaks of emission for example 'cores', 'clumps', or 'fibres' (e.g. CLUMPFIND, Williams, Blitz & Stark 1995; GAUSSCLUMPS, Stutzki & Guesten 1990; FELLWALKER, Berry 2015; FIVE, Hacar et al. 2013), those which target the hierarchical structure of molecular clouds (e.g. ASTRODENDRO, www.dendrograms.org; DENDROFIND/QUICKCLUMP, Wünsch et al. 2012; Sidorin 2017; see also Miville-Deschênes, Murray & Lee 2017), those which specifically aim to extract molecular clouds (e.g. CPROPS, Rosolowsky & Leroy 2006; SCIMES, Colombo et al. 2015), and those which have been used to target structure with a particular geometry, for instance, filaments (e.g. DISPERSE, Sousbie 2011; Sousbie, Pichon & Kawahara 2011; FILFINDER, Koch & Rosolowsky 2015). Despite this, there is currently no publicly available code whose primary function is to extract hierarchical structure within molecular clouds, thereby providing the important connection between cores, clumps, and clouds, but which simultaneously retains the pixel scale information needed to study variation in the kinematics throughout each member of the hierarchy.

Our solution to this problem is the development of a new analysis tool, written in Python, named ACORNS (Agglomerative Clustering for ORganising Nested Structures). The primary function of ACORNS is to generate a hierarchical system of clusters within discrete data. Although ACORNS was designed with the analysis of spectroscopic (position–position–velocity; PPV) data in mind, it can readily be implemented to other data sets, providing many applications.[22] The following section is dedicated to describing the methodology used by ACORNS.

### B1 ACORNS: Agglomerative Clustering for ORganising Nested Structures

#### *B1.1 Introduction and description of the input parameters*

ACORNS follows the philosophy of hierarchical agglomerative clustering (HAC).[23] HAC methods fall into two main categories: 'bottom-up' or 'top-down'. ACORNS follows the bottom-up approach in that each singleton data point begins its life as a 'cluster'. Traditionally, clusters then merge until only a single cluster remains that contains all of the data. The output of this technique is often visualized graphically as a dendrogram, which has become popular in astronomy as a convenient way of representing and interpreting the hierarchical nature of molecular clouds (e.g. Houlahan & Scalo 1992; Rosolowsky et al. 2008).

Briefly, clustering in ACORNS commences with the most significant data point. In the analysis presented in this work, this refers to the data point with the greatest peak intensity. However, given the applicability of ACORNS to different systems, this may instead refer to, for example, a density, column density, or mass. ACORNS then descends in significance, merging clusters based on physically motivated user-provided criteria, until a hierarchy is established.

Input to ACORNS is an array of $n \times m$ dimensions, where $n$ is the number of parameters, at minimum 4, but in principle has no upper limit and depends on how many parameters the user wishes to use during the clustering procedure. $m$ refers to the number of data points in the sample. As an example, in its simplest form (clustering in two spatial dimensions), this array should consist of $x$ position, $y$ position, intensity (or equivalent), and an uncertainty on the intensity (or equivalent). If linking in PPV, an additional column for the velocity is a mandatory requirement.

The linking of clusters is handled via the supply of an array containing $n - 3$ elements (or $n - 4$ if linking in PPP) which describes the clustering criteria (`cluster_criteria`). Here, the

[22] ACORNS is publicly available for download here: https://github.com/jdhenshaw/acorns.

[23] More information on this technique and its philosophy can be found in Manning, Raghavan & Schütze (2008).

user must supply the maximum spatial Euclidean distance between data points, as well as the maximum absolute difference in any other variable used for linking. If the separation between two data points satisfy these criteria, the data points are considered to be linked. We note however, that no two data points extracted from the same location can be linked to the same cluster.

In addition the user must supply the following parameters:

(i) The pixel size in equivalent units to the positional information in the input array (`pixel_size`).
(ii) The radius of the smallest structures the user would like ACORNS to identify (`min_radius`).
(iii) The minimum height above a merge level for a cluster to be considered as a separate structure (`min_height`).
(iv) The stopping criteria, given as a multiple of the rms noise level (`stop`).

In the following sections we provide a qualitative description of the ACORNS algorithm.[24] We begin with a description of the overall methodology before expanding on some of the individual steps.

### *B1.2 A description of the method*

The main steps taken by ACORNS in developing the hierarchy are as follows (these are also illustrated in the flow diagram presented in Fig. B1):

(1) ACORNS begins by creating a catalogue of the currently unassigned data. All data whose intensity, $I$, satisfies the following criterion are added to this catalogue:

$$I > \texttt{stop} \times \sigma_{\rm rms}, \tag{B1}$$

where $\sigma_{\rm rms}$ refers to the noise level at that position. The unassigned data is then rearranged in descending order of $I$.

(2) These data are used to generate a k-d tree,[25] which can be queried to return the nearest neighbours to a given point.

(3) Starting with the first data point in the unassigned catalogue, and looping over all data points in the unassigned catalogue, ACORNS implements the following steps:

(i) ACORNS first generates a 'bud cluster'. Extending the nomenclature of Houlahan & Scalo (1992), a bud cluster refers to a structure that has not yet met the criteria to become a 'leaf' in its own right (where leaves are the clusters situated at the top of the hierarchical system).
(ii) ACORNS queries the k-d tree to find all data points that are within some maximum Euclidean distance (provided in `cluster_criteria`) from the bud cluster (see Section B1.1). If additional linking criteria are supplied by the user, ACORNS then computes the maximum absolute difference in the desired property between the bud cluster and these data points. This is then also checked against the linking criteria supplied within `cluster_criteria`.
(iii) All data satisfying the clustering criteria are then cross-referenced against the current cluster catalogue to see if they belong to an already established cluster within the hierarchy. If so, a link is established and the hierarchy grows (we will expand on this methodology in Section B1.3).

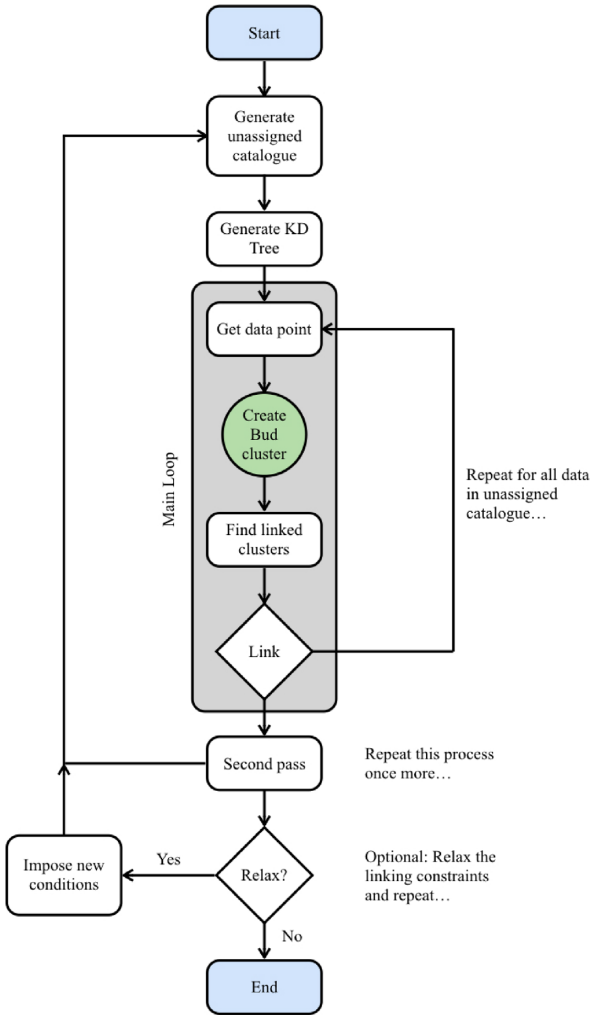


**Figure B1.** A flow chart depicting the steps followed by ACORNS during the clustering procedure. The main loop is indicated in dark grey. The creation of new clusters appears in green. The procedure employed during the 'Link' phase is described in Section B1.3 and Fig. B2.

(4) Once ACORNS has cycled through all data points in the unassigned catalogue, it begins a second loop. The cluster catalogue is first cleaned of any bud clusters and these data are used to generate a new unassigned catalogue. This step picks up any data points that were unable to be linked during the first pass of the algorithm.

(5) If specified by the user (`relax`), the clustering criteria are relaxed and ACORNS performs additional loops based on this new criteria. This helps further develop the hierarchy and this method is described in more detail in Section B1.4.

(6) ACORNS then discards all remaining bud clusters since they did not meet the criteria to become fully fledged clusters.

ACORNS returns a system of clusters as its output. In a given hierarchy, the antecessor is the largest common ancestor of all clusters within that hierarchy (note that for a given data set, there may be multiple antecessors and each of them may or may not have descendant substructure). Expanding the nomenclature typically used in describing dendrograms (see e.g. Houlahan & Scalo 1992),

[24] Throughout this description we will use intensity as an example; however, in principle this can be exchanged for an equivalent parameter.
[25] A data structure used to organize a number of points in a space with k-dimensions.

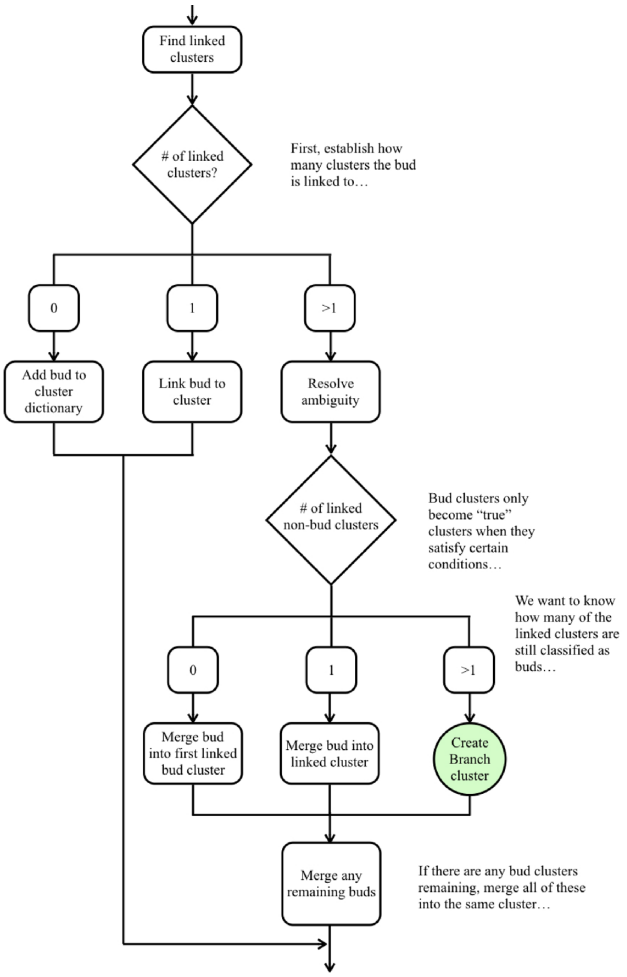


**Figure B2.** A flow chart describing the growth and merging of clusters. This strategy follows the methods of ASTRODENDRO (www.dendrograms.org) and QUICKCLUMP (Sidorin 2017), see Section B1.3 for more details.

an antecessor refers to a tree in a forest of clusters. Each tree may or may not exhibit substructure, referred to as branches and leaves.

### *B1.3 The growth of the hierarchy*

The procedure employed by ACORNS during the growth of the hierarchy is described in the flow chart in Fig. B2. This growth strategy is developed following the methods of ASTRODENDRO (www.dendrograms.org) and QUICKCLUMP (Sidorin 2017). However, key differences in the algorithms (namely working with discrete data, rather than uniformly spaced data cubes) necessitate important differences in the details of each step. After establishing a link between the bud cluster (see Section B1.2) and already-established clusters in the hierarchy [see step (iii)(c) in Section B1.2], the next step depends on the number of linked clusters:

(1) If no linked clusters are identified, the bud cluster is added to the cluster catalogue as a new cluster.

(2) If only a single cluster is identified as linked, the bud cluster is merged into this already-established cluster.

(3) If multiple linked clusters are identified, further decision making is required (see 'Resolve ambiguity' in Fig. B2). ACORNS first determines how many of the linked clusters are 'true' clusters (i.e. not bud clusters). Once this has been determined, what happens next depends on how many fully fledged clusters our bud cluster is linked to:

(i) If none, then this tells us that all of the clusters linked to our bud must also be bud clusters. We merge our bud cluster into the first of the other buds.

(ii) If there is only a single linked cluster, we merge the bud cluster into this already established cluster.

(iii) If there are multiple linked clusters, we generate a branch between these fully fledged clusters – a new level in the hierarchy.

All remaining bud clusters (if any) are then merged with the same cluster as our original bud cluster, be it a bud (i), a fully fledged cluster (ii), or clusters (iii).

### *B1.4 Relaxing the linking constraints*

This (optional) second phase of the algorithm can be of importance when working with discrete and irregularly spaced data such as the velocity information output following the decomposition of spectroscopic data (and not, for example, regularly spaced velocity channels within a data cube). Conceptually, the idea behind this phase is to relax the linking constraints used during stage 1 (Section B1.2) in order to further develop the hierarchy. This can be implemented in multiple ways. The user has the option to relax the constraints either in a single step or incrementally (either interactively or non-interactively).

ACORNS first generates a new catalogue of data points that were not assigned to clusters during the first and second passes of the algorithm using the initial clustering criteria.[26] As with the implementation discussed in Section B1.2, ACORNS starts with the most significant (in this example, that with the greatest peak intensity) data point in the unassigned catalogue. Steps (i)–(iii) are implemented as in Fig. B1, but this time using the new relaxed criteria. The main differences during the relax phase relate to the steps labelled 'Find linked clusters' and 'Link' in Fig. B1, and are outlined as follows:

(i) During this phase, ACORNS attempts to link bud clusters (Fig. B1) to an already-established forest (see Section B1.2). It is important to ensure that any links that are created are still strong despite having relaxed the linking constraints.

Often the user wants to link data based on more than just positional (and intensity) information. Therefore, if additional properties are considered when searching for linked clusters (e.g. the centroid velocity or velocity dispersion), ACORNS checks these properties against those of the linked clusters. If the properties of the bud cluster lie $>3\sigma$ away from the mean of the linked cluster properties (where $\sigma$ refers to the standard deviation of that property), then these linked clusters are prevented from creating links. This ensures that even despite relaxing the linking constraints, only strong links are forged.

(ii) During the relax phase, a bud cluster may be linked to multiple trees within the forest and, in some cases, it may be linked to multiple clusters belonging to the same tree. ACORNS first establishes whether or not it is possible to insert the bud cluster into the correct position in the hierarchy. This is governed by the peak intensity of

[26]Incidentally, the 'second pass' of the algorithm [see step (iv) in Section B1.2] follows the exact strategy outlined in this section; however, the linking constraints are not relaxed.

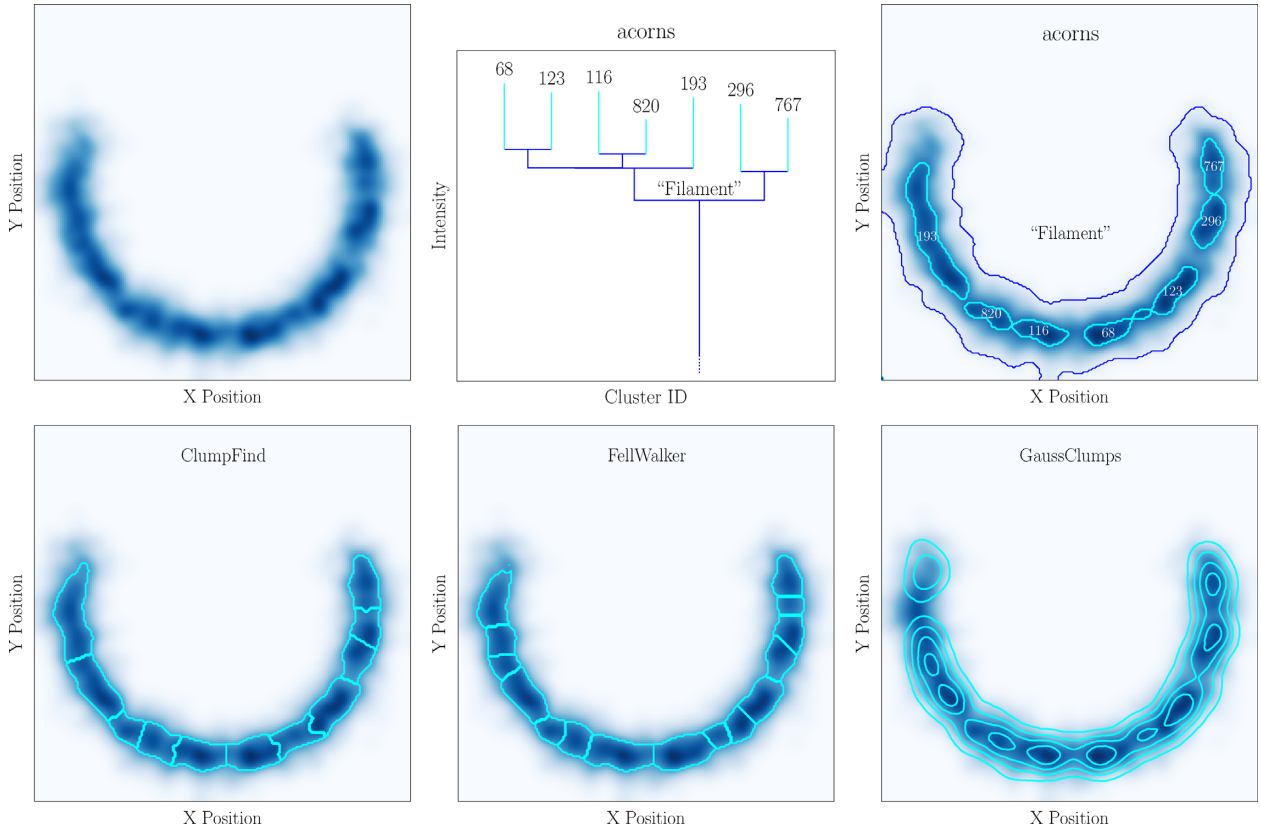


**Figure B3.** Top panels: Demonstration of ACORNS on 2-D data. Top left-hand panel: A fake clumpy 'filament'. The colour-scale is a proxy for intensity. Top central panel: A graphical representation of the hierarchy found by ACORNS. Top right-hand panel: ACORNS clusters displayed as contours. In the centre and right-hand panels, the leaves, i.e. the clusters situated at the top of the hierarchy, are displayed in cyan. The tree, corresponding to the 'filament' is displayed as a dark blue contour. Bottom panels: Structure finding methods from the literature for comparison. Bottom left-hand panel: CLUMPFIND (Williams et al. 1995). Bottom centre panel: FELLWALKER (Berry 2015). Bottom right-hand panel: GAUSSCLUMPS (Stutzki & Guesten 1990). A key difference between ACORNS (also ASTRODENDRO) and the algorithms presented in the bottom panels is that the latter methods search for discretized islands of emission, whereas the former methods extract hierarchical structural information.

the bud cluster, and the minimum and maximum intensity levels of each linked cluster. If the bud cluster cannot be inserted into the linked cluster, ACORNS searches downwards in the hierarchical tree (if possible) to establish a link. If the bud cluster cannot be slotted in at the correct level in any established tree, these linked clusters are ignored. ACORNS returns a single linked cluster *per tree* to which the bud cluster will be linked.

(iii) Step (iii)(c) (Fig. B2) is implemented as above, with a key difference during the branching procedure. If a branch is to be created, the bud cluster is firstly merged with the closest matching cluster out of all the available linked clusters. A new branch (between multiple trees) is then created at the base of the parent hierarchies.

### B2 ACORNS: Clustering in 2-D

In this section we demonstrate the application of ACORNS to 2-D data. The top left-hand panel of Fig. B3 depicts a clumpy 'filament' from which we wish to extract structural information. The filament was generated using the clustering examples in Python's scikit-learn package.[27] We first generate a 2-D set of data points distributed randomly within the confines of a semicircle with finite width. We then convert the point density into an image by convolving the point density with a Gaussian kernel.

The top central panel of Fig. B3 shows a graphical representation of the hierarchical system identified by ACORNS known as a dendrogram. ACORNS picks out a total of seven 'leaves', which are situated at the top of the hierarchy and highlighted in cyan, all of which belong to a single 'tree' (i.e. the 'filament'). The top right-hand panel highlights this information on the filament image. The leaves are indicated by cyan contours and the filament appears in dark blue. In this particular instance we chose to search for clusters using only the distance between data points as linking criteria. Consequently this solution is identical to that found with ASTRODENDRO using equivalent input parameters. However, in principle (i.e. if available), additional constraints could be added to the ACORNS linking procedure, which would result in the solutions from the two algorithms diverging. As an example, if one also had a measurement of temperature at each position, that could also be included in the clustering procedure.

The bottom panels compare this result with other structure finding algorithms commonly used in the literature, namely CLUMPFIND (left; Williams et al. 1995), FELLWALKER (centre; Berry 2015), and *gaussclumps* (right; Stutzki & Guesten 1990). Each of these algorithms seeks to identify discretized islands of emission, breaking

[27] http://scikit-learn.org/stable/modules/clustering.html.

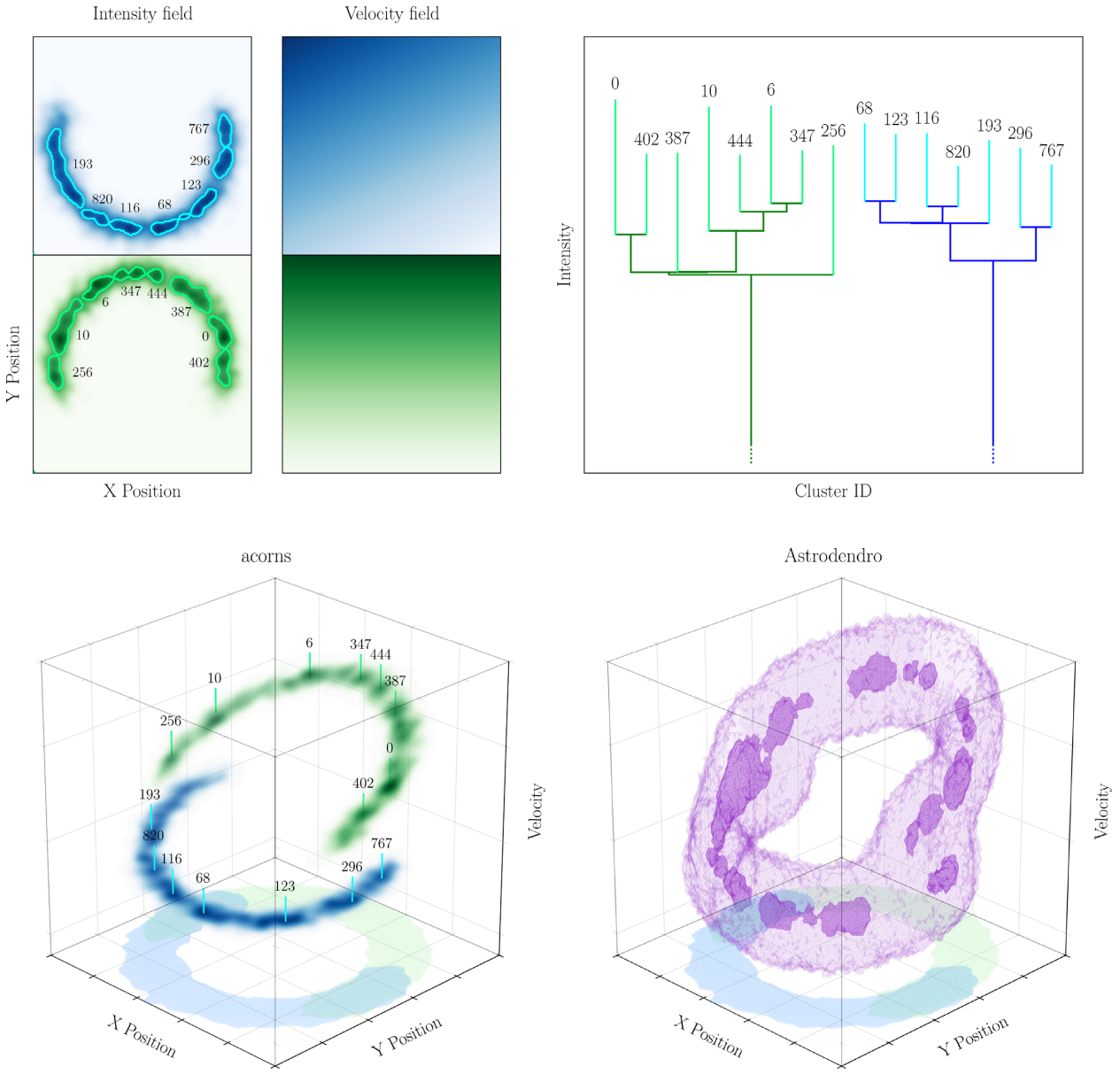


**Figure B4.** A figure demonstrating the use of ACORNS in position–position–velocity space. Top-left-hand panels: two clumpy 'filaments'. The colour-scale in the left-hand plots is a proxy for intensity. The right-hand plots show the corresponding velocity fields of the filaments (velocity increases from light to dark). The contours overlaid on the intensity-field highlight the leaves identified by ACORNS. Top-right-hand panel: The corresponding dendrogram. Bottom left-hand panel: The PPV structure of the clusters. ACORNS identified leaves are marked. Bottom right-hand panel: Demonstrating the use of ASTRODENDRO on the same data. The semitransparent purple rendering highlights the tree identified by ASTRODENDRO. ASTRODENDRO finds a single isosurface due to the merging of the filaments in PPV space. At no point in the ASTRODENDRO hierarchy are the same two filamentary structures identified in contrast to the ACORNS solution. The dark purple structures highlight the leaves. Encouragingly, there is close correspondence between many of the leaves identified by both algorithms. However, a crucial difference is that leaves #193 and #256, identified with ACORNS, merge in PPV space and therefore are also merged in the ASTRODENDRO solution.

the map up into 'clumps', whereas ACORNS (also ASTRODENDRO) searches for hierarchical information within data.

### B3 ACORNS: Clustering in 3-D

A key difference between ACORNS and the algorithms mentioned in Section B1.1 is that ACORNS is designed to work on decomposed spectroscopic data rather than data cubes. Analysis with ACORNS can therefore be performed in unison with such algorithms, complementing their results by providing a detailed description of the gas kinematics.

This is illustrated in Fig. B4. Here we have generated two 'filaments' (one of which is identical to that shown in Fig. B3). The intensity field of both filaments is illustrated in the top-left-hand panels of Fig. B4. In this example, we also impose a velocity field. The filaments have differing velocity gradients (also shown in the top-left-hand panels) and a uniform velocity dispersion (not shown). The intensity distribution and velocity field of the filaments are designed in such a way that the filaments overlap in PPV space.

The top panels of Fig. B4 display the result of applying ACORNS to this configuration. ACORNS identifies two clusters in the decomposed data. Importantly, the hierarchy associated with the blue filament is *identical* to that found in Section B2. A corresponding hierarchy is identified for the green filament. The top-right-hand image displays a representative dendrogram of this hierarchical system.

The bottom left-hand image in Fig. B4 displays structures recovered by ACORNS in PPV space. At the base of the image we demonstrate how the two filaments overlap in projection and appear as a ring. In PPV space, we plot the velocity centroids; the data that ACORNS uses for clustering. As can be seen, despite the velocity dispersion of the filaments being large enough such that they overlap in PPV space, the two clusters are distinguishable when focusing on their centroids.

To illustrate the difference in approach between ACORNS and ASTRODENDRO, the bottom right-hand image of Fig. B4 displays the result of running ASTRODENDRO on the same data cube. The light-coloured semitransparent feature is a volume rendering of the main structure identified by ASTRODENDRO (i.e. the trunk of the hierarchy), and the darker shaded structures refer to the leaves. Herein lies the key difference between the algorithms. Because of the blending in PPV space between the two filaments, ASTRODENDRO, which classifies structure as independent isosurfaces, returns a singular doughnut-shaped structure. At no point in the hierarchy are the two input filaments returned by ASTRODENDRO.

Encouragingly, there is close correspondence between many of the leaves identified using both algorithms. There are some very slight differences owing to the differences in the input parameters, but these are small. The key difference occurs where blending in PPV space is observed, for example, leaves #193 and #256. The reason these are picked out by ACORNS is because these two features are identifiable (albeit blended) as multiple velocity components in the spectra and are extracted as such during the spectral decomposition.

Consequently, the two methods complement each other nicely. While tools such as ASTRODENDRO and SCIMES (which uses ASTRODENDRO for structure identification) are able to pick out molecular clouds as isosurfaces in spectroscopic data, ACORNS can be used to search for regions of statistical similarity *within* such clouds.

[1]*Max Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany*
[2]*National Radio Astronomy Observatory, 1003 Lopezville Rd., Socorro, NM 87801, USA*
[3]*Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK*
[4]*Astrophysics Research Institute, Liverpool John Moores University, Liverpool L3 5RF, UK*
[5]*Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstraße 12-14, D-69120 Heidelberg, Germany*
[6]*Physics Department, Brandeis University, 415 South Street, Waltham, MA 02453, USA*
[7]*Max Planck Institute for Extraterrestrial Physics, Gießenbachstraße 1, D-85748 Garching bei München, Germany*
[8]*Joint ALMA Observatory, Alonso de Cordova 3107, Vitacura, Santiago, Chile*
[9]*National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan*
[10]*Argelander Institute for Astronomy, University of Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany*
[11]*Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*
[12]*Center for Astrophysics and Space Astronomy, Astrophysical and Planetary Sciences Department, University of Colorado, Boulder, CO 80309, USA*
[13]*National Radio Astronomy Observatory, PO Box 2, Green Bank, WV 24944, USA*
[14]*Department of Physics, University of Connecticut, Storrs, CT 06269, USA*
[15]*Centre for Astrophysics Research, University of Hertfordshire, College Lane, Hatfield, AL10 9AB, UK*
[16]*School of Mathematical and Physical Sciences, University of Newcastle, University Drive, Callaghan NSW 2308, Australia*
[17]*Hay Stack Observatory, Massachusetts Institute of Technology, 99 Millstone Road, Westford, MA 01886, USA*
[18]*Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53212 Bonn, Germany*
[19]*Institute for Astrophysical Research, Boston University, 725 Commonwealth Avenue, Boston, MA 02215, USA*
[20]*School of Physics and Astronomy, Cardiff University, Queen's Buildings, The Parade, Cardiff CF24 3AA, UK*
[21]*Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*

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