# LEGO – II. A 3 mm molecular line study covering 100 pc of one of the most actively star-forming portions within the Milky Way disc

A. T. Barnes,[1]★ J. Kauffmann,[2] F. Bigiel,[1] N. Brinkmann,[3] D. Colombo,[3] A. E Guzmán ®,[4] W. J. Kim,[5] L. Szűcs,[6] V. Wakelam ®,[7] S. Aalto,[8] T. Albertsson,[3] N. J. Evans,[9] S. C. O. Glover,[10] P. F. Goldsmith,[11] C. Kramer,[12] K. Menten,[3] Y. Nishimura,[13,14] S. Viti,[15] Y. Watanabe ®,[16] A. Weiss,[3] M. Wienen,[3,17] H. Wiesemeyer[3] and F. Wyrowski[3]

*Affiliations are listed at the end of the paper*

## ABSTRACT

The current generation of (sub)mm-telescopes has allowed molecular line emission to become a major tool for studying the physical, kinematic, and chemical properties of extragalactic systems, yet exploiting these observations requires a detailed understanding of where emission lines originate within the Milky Way. In this paper, we present 60 arcsec ($\sim$3 pc) resolution observations of many 3 mm band molecular lines across a large map of the W49 massive star-forming region ($\sim$100 pc $\times$ 100 pc at 11 kpc), which were taken as part of the 'LEGO' IRAM-30m large project. We find that the spatial extent or brightness of the molecular line transitions are not well correlated with their critical densities, highlighting abundance and optical depth must be considered when estimating line emission characteristics. We explore how the total emission and emission efficiency (i.e. line brightness per $H_2$ column density) of the line emission vary as a function of molecular hydrogen column density and dust temperature. We find that there is not a single region of this parameter space responsible for the brightest and most efficiently emitting gas for all species. For example, we find that the HCN transition shows high emission efficiency at high column density ($10^{22}$ cm$^{-2}$) and moderate temperatures (35 K), whilst e.g. $N_2H^+$ emits most efficiently towards lower temperatures ($10^{22}$ cm$^{-2}$; $<20$ K). We determine $X_{CO(1-0)} \sim 0.3 \times 10^{20}$ cm$^{-2}$ (K km s$^{-1}$)$^{-1}$, and $\alpha_{HCN(1-0)} \sim 30$ M$_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$, which both differ significantly from the commonly adopted values. In all, these results suggest caution should be taken when interpreting molecular line emission.

**Key words:** stars: formation – ISM: clouds – ISM: molecules – galaxies: evolution – galaxies: ISM – galaxies: star formation.

## 1 INTRODUCTION

A full understanding of the process of star formation from dense molecular clouds is still one of the major outstanding problems in astronomy. Many fundamental open questions are still hotly debated, such as understanding the relative balance of turbulence, magnetic field, and gravitational energy (e.g. Krumholz & McKee 2005; Padoan & Nordlund 2011; Hennebelle & Chabrier 2013), how these properties can be observationally determined (e.g. Federrath et al. 2010; Orkisz et al. 2017), how these properties may vary with Galactic environment (e.g. Federrath et al. 2016; Barnes et al. 2017), and if they are important for the fraction of 'dense' gas that is eventually converted into stars (e.g. Lada, Lombardi & Alves 2010; Lada et al. 2012).

The luminosities, $L_Q$, from the now extensive collection of known molecular line transitions, $Q$,[1] originating from molecular clouds are useful tools for studying the above-mentioned properties. In

particular, transitions with high critical densities are regularly used in the literature to selectively study the relatively denser gas within molecular clouds. For example, it is typically assumed that the HCN ($J = 1-0$) molecular line transition is a particularly good tracer of 'dense gas' (number density of $H_2$ above $10^4$ cm$^{-3}$), owing to its high critical density for emission, and a chemical formation pathway that produces abundant HCN molecules within cold, dense gas (e.g. Boger & Sternberg 2005). Under specific physical assumptions, the luminosity of HCN ($J = 1-0$), $L_{HCN(1-0)}$ can be used to estimate the dense molecular gas mass, $M_{dg}$, within star-forming regions (Gao & Solomon 2004a,b; Wu et al. 2010). A measure of the dense gas mass, along with the star formation rate (SFR), $\dot{M}_*$, has been used to constrain the models for star formation through, for example, the star formation efficiency (SFE) of dense gas ($\dot{M}_*/M_{dg}$). However, with the advent of more sensitive radio telescopes that are capable of efficiently mapping large regions of the sky, it has been pointed out that the $L_{HCN(1-0)} \propto M_{dg}$ relation may not hold across environments (e.g. Kauffmann et al. 2017; Pety et al. 2017; Shimajiri et al. 2017; Evans et al. 2020; Nguyen-Luong et al. 2020). The results from these authors suggest that the underlying assumption that molecular line emission originates solely from gas around and above the critical density is fundamentally flawed,

★ E-mail: ashleybarnes.astro@gmail.com

[1] As of November 2019, around 210 molecules have been detected in the interstellar medium (Endres et al. 2016; https://cdms.astro.uni-koeln.de/classic/molecules).
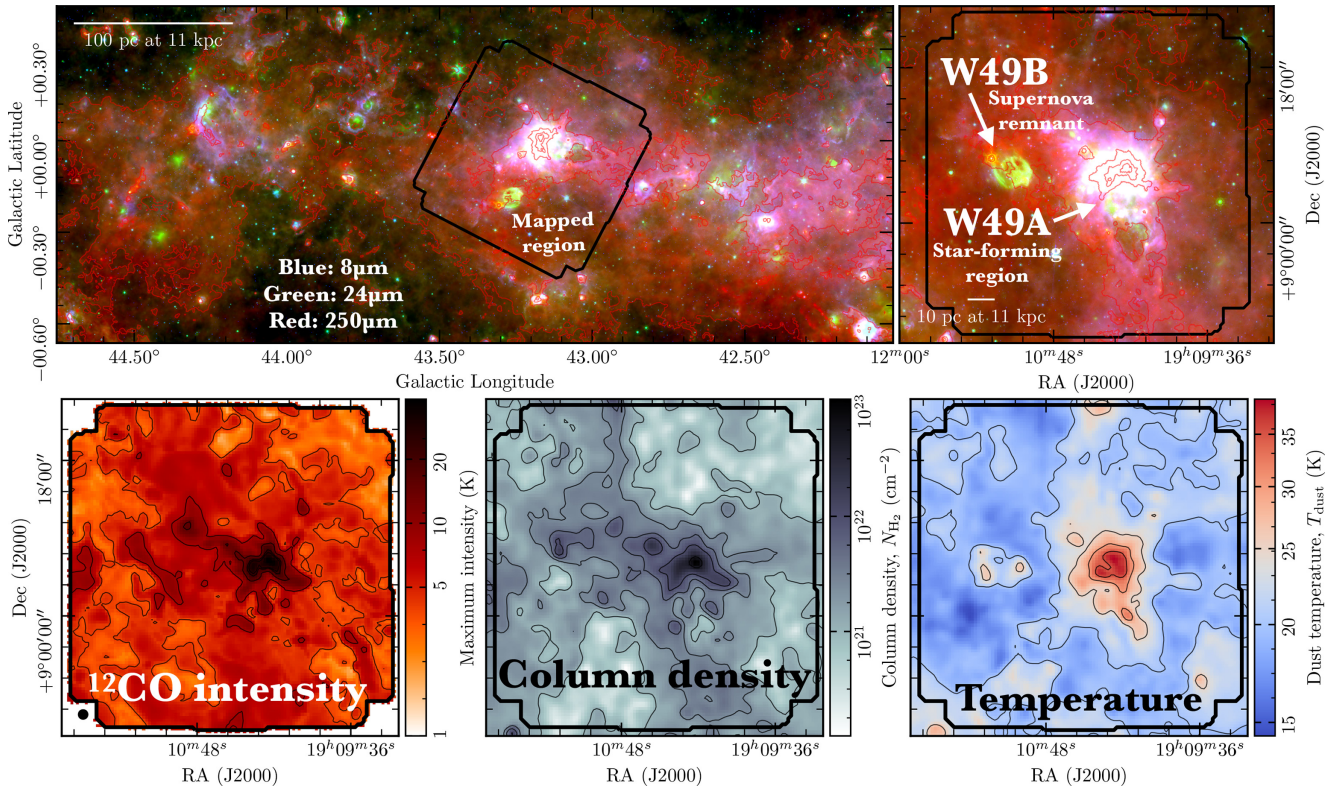
**Figure 1.** The region around W49 within the plane of the Milky Way. The upper panel shows a three-colour image across the Galactic plane, and a zoom-in on the mapped region of W49 indicated by the heavy black line (see Section 2). Shown are the *Spitzer* 8 μm in blue, *Spitzer* 24 μm in green, and *Herschel* 250 μm in red (Carey et al. 2009; Churchwell et al. 2009; Molinari et al. 2010). Overlaid as red contours is the 250 μm map in levels of 100, 200, 400, 1000, 2000, 10000 MJy sr$^{-1}$, which were chosen to best highlight the dense dust/gas distribution (Molinari et al. 2010). The labels on the zoom-in panel show the positions of the W49A star-forming region and the W49B supernova remnant. Shown in the main and zoom-in panels are scale bars of 100 and 10 pc at a distance of 11 kpc, respectively. The lower row, left-hand panel shows a map of the peak brightness temperature from the $^{12}$CO molecular line, overlaid with black contours of 4, 8, 15, 22, 25 K. This $^{12}$CO map highlights the extended nature of the molecular emission, which completely fills the mapped region. The angular beam size of 60 arcsec, or ~3 pc at the distance of W49A (~11 kpc; Gwinn, Moran & Reid 1992; Zhang et al. 2013), is shown as a black circle in the lower left of the lower left panel. The lower row, centre and right-hand panels show maps of the molecular hydrogen column density $N_{H_2}$ (centre panel) and dust temperature $T_{dust}$ (right-hand panel) derived from *Herschel* observations. These maps have been smoothed to an angular resolution of 60 arcsec to match the resolution of the IRAM-30m observations. Black contours of $N_{H_2}$ have been overlaid on the $N_{H_2}$ map in levels of 1, 2.5, 5, 10, 50, 100 × 10$^{21}$ cm$^{-2}$, and black contours of $T_{dust}$ have been overlaid on the $T_{dust}$ map in levels of 10, 15, 20, 25, and 30 K.

despite this being made by many molecular line studies. Indeed, Evans (1989) highlighted three decades ago that such an assumption breaks down out of the Rayleigh–Jeans limit in the sub-mm regime, where sub-thermal emission is the typical mode for emission before quickly becoming optically thick just above the critical density. If anything then, the opposite is true, and emission from molecular lines within the sub-mm regime traces gas with densities up to the critical density. Moreover, there could be a contribution to $L_Q$ from various additional excitation mechanisms (e.g. electron excitation; Goldsmith & Kauffmann 2017).

Investigating the emission properties of these dense gas tracers – and many other commonly observed molecular line transitions – across a broad range of environments is, therefore, of current critical importance (e.g. Nishimura et al. 2017; Watanabe et al. 2017). This is of particular relevance now more than ever, as these dense molecular gas tracers are becoming routinely observable in other galaxies (e.g. Bigiel et al. 2016), where, even at the highest achievable spatial resolution for the closest disc galaxies, each line of sight contains a very broad range of environments (e.g. gas densities and temperatures; e.g. Jiménez-Donaire et al. 2019).

One of the primary aims of the Molecular Line Emission as a Tool for Galaxy Observations (LEGO) project is to further investigate the above issue by obtaining large maps of many commonly observed 3 mm band molecular lines, towards a selection of Milky Way star-forming regions that span a range of environmental properties (e.g. cloud density, star formation feedback, galactic environment, and metallicity). Building on the initial results from the Orion nebula presented by Kauffmann et al. (2017), we present in this work observations of the W49 massive star-forming region, which represents the most distant and, in terms of star formation rate, extreme source within the LEGO sample.

A three-colour image of the W49 region and its position within the galactic plane of the Milky Way is presented in the upper row of Fig. 1. The bright blues (*Spitzer* 8 μm emission) and greens (*Spitzer* 24 μm emission) within this image highlight the regions of ongoing star formation, whilst the reds (*Herschel* 250 μm emission) show the cold dust (Carey et al. 2009; Churchwell et al. 2009; Molinari et al. 2010). The mapped region (indicated by the black box) is of particular interest for a study of how the gas emission properties vary with environment, thanks to the large range of physical regimes

**Table 1.** Source properties of the W49 region.

| Parameter | Value | Notes |
|---|---|---|
| Distance | $11.11^{+0.79}_{-0.69}$ kpc | 60 arcsec $\approx$ 3 pc (Gwinn et al. 1992; Zhang et al. 2013) |
| Velocity range of all emission | [−20.0 to 90.0] km s$^{-1}$ | LSR, radio convention |
| Systemic velocity of W49A | [−10.0 to 20.0] km s$^{-1}$ | LSR, radio convention (e.g. Galván-Madrid et al. 2013) |
| Projection centre | $19^h10^m32^s.487$, $9°05'36''.532$ | RA, Dec. (J2000), or $l$, $b$ = 43.19, −0.06 (see Fig. 1) |
| Field of view | $0°.54 \times 0°.54$ | $\sim$100 × 100 pc at 11 kpc (see Fig. 1) |
| Column density, $N_{H_2}$ | [0.1, 2.1, 210] × $10^{21}$ cm$^{-2}$ | $N_{H_2}^{min}$, $N_{H_2}^{mean}$, $N_{H_2}^{max}$ (Section 3.1 and Appendix C) |
| Dust temperature, $T_{dust}$ | [14.4, 20.5, 38.9] K | $T_{dust}^{min}$, $T_{dust}^{mean}$, $T_{dust}^{max}$ (Section 3.1 and Appendix C) |

present within such a relatively small projected area on the sky. Table 1 summarizes several properties of interest for the mapped region.

The W49A region highlighted in Fig. 1 is a very well-studied actively star-forming region that contains high-mass stars and associated H II regions (total stellar mass of $10^{4-5}$ M$_\odot$; e.g. Homeier & Alves 2005). Thanks to these H II regions, it is one of the brightest known radio sources outside of the Galactic Centre (e.g. Westerhout 1958). Therefore, W49A has been the subject of many radio continuum and recombination line investigations that have tried to characterize the embedded stellar objects and ionized gas properties (Anantharamaiah 1985; De Pree, Mehringer & Goss 1997; De Pree et al. 2000; Rugel et al. 2019). Moreover, and of particular interest to the study presented here, the central part of this region has been the subject of several mm-wavelength spectroscopic observations, which find a wealth of molecular line emission (Roberts et al. 2011; Nagy et al. 2012; Galván-Madrid et al. 2013; Nagy et al. 2015). Parallax distance measurements using H$_2$O maser emission towards W49A place it at a distance of $11.11^{+0.79}_{-0.69}$ kpc, locating it in a distant section of the Perseus arm near the solar circle in the first Galactic quadrant (Gwinn et al. 1992; Zhang et al. 2013). The W49B supernova remnant is also highlighted in Fig. 1, and stands out as a bright feature in 24 μm emission (i.e. green in the upper right panel).

The paper is organized as follows. Section 2 describes the IRAM-30m observations, the procedure for creating the integrated intensity maps, and the *Herschel* derived molecular hydrogen column density and dust temperature maps. Section 3 presents the results of the comparison of integrated intensities of the various molecular lines to the column density. Section 4 presents an analysis of where the majority of the emission is emitted within our mapped region, created by applying regional and dust temperature masks. Section 5 places these in the context of our understanding of molecular line emission. In Section 6, we summarize the main results of this work.

## 2 OBSERVATIONS

### 2.1 Observations with the IRAM 30 m telescope

#### 2.1.1 Data reduction

The observations presented here were taken as part of the Molecular Line Emission as a Tool for Galaxy Observations (LEGO) survey (IRAM project code: 183-17). The full observational information of this survey, which covers $\sim$25 regions throughout the Galactic disc, will be summarized in a future publication (Kauffmann et al., in preparation). This survey was conducted using the Eight Mixer Receiver (EMIR; Carter et al. 2012) on the 30 m telescope of the

Instituto de Radioastronomía Milimétrica (IRAM) on Pico Veleta, Spain. Data in the horizontal and vertical polarizations were acquired with the Fast Fourier Transform Spectrometer (FTS) operating in a wide–band mode that provides a spectral resolution of 200 kHz (i.e. 0.6 km s$^{-1}$ · [$\nu$/100 GHz]$^{-1}$ in velocity). The EMIR receiver was operated in two set-ups, centred at local oscillator frequencies of 97.830 and 103.840 GHz. As a result, the observations cover in combination continuous frequency ranges of 86.10–99.89 and 101.78–115.57 GHz. Relatively bright emission lines from a variety of astrophysically relevant molecules can be found at these frequencies. The transitions investigated by LEGO are listed in Table 2, which is created as described in Section 2.1.2. The telescope's intrinsic beam size[2] is 24''.6 · ($\nu$/100 GHz)$^{-1}$. The angular resolution of our data is lower, though, due to the processing described below.

The observations cover a total area of $\sim$30 arcmin × 30 arcmin size around a central reference coordinate of $19^h10^m32^s.487$, $9°05'36''.532$. The On-The-Fly (OTF) mapping technique was used to create the mosaic image, using a dump time of 0.25 s and a scanning speed of 90 arcsec s$^{-1}$. The rows of the OTF map were spaced by 10 arcsec, corresponding to less than half the telescope's intrinsic resolution at the frequencies $\lesssim$115 GHz covered by our project. The OTF maps were taken by scanning along either the RA or Dec. axis of the equatorial coordinate system. Every point in the map was covered by at least two orthogonal scans, in order to reduce artefacts due to scan patterns.

The LEGO data reduction pipeline uses functionality from the Python programming language and from IRAM's GILDAS software suite (i.e. the CLASS package).[3] This approach permits use of Python's superior bookkeeping functionality to control the data flow, while it also takes advantage of the rich and well-tested data processing functionality within CLASS for low-level tasks. The pipeline, as well as the derived data products, will be released as part of future publications. In the very first processing step, observing logs are generated on the basis of header information of observed scans and data from IRAM's Telescope Access for Public Archive System (TAPAS).[4] These logs constitute the basis for all further data processing steps.

The pipeline processes the data separately for every emission line recorded in Table 2. Python scripts are used to identify, for a given emission line, the observations covering a particular source. CLASS is then employed to extract a velocity range of $\gtrsim$350 km s$^{-1}$ total width around a given spectral line (a larger range is extracted around

**Table 2.** Information on the selected observed molecular lines, ordered by increasing rest frequency. Section 2.1.2 describes how the lines recorded in this table were selected, and how the line characteristics recorded here were obtained. Columns 1–10 show the name of each molecule, the transition information (used to refer to each transition throughout), the frequency of the transition, the upper energy level of the transition, the Einstein spontaneous decay coefficient, the collisional deexcitation rate coefficients at a kinetic temperature of 20 K, and the critical and effective densities for emission.[a] Transitions that are not available within the LAMDA data base have the corresponding information blanked. Additional information on the molecular line data base used within this work can be found in Table A1. The full, machine-readable version of this table can be obtained from the supplementary online material.

| Species | Transition | Frequencies (GHz) | $E_u/k_B$ (K) | $A_{ul}$ (s$^{-1}$) | $C_{ul}^{2lvl}$ (cm$^3$ s$^{-1}$) | $C_{ul}$ (cm$^3$ s$^{-1}$) | $n_{crit}^{2lvl}$ (cm$^{-3}$) | $n_{crit}$ (cm$^{-3}$) | $n_{eff}$ (cm$^{-3}$) |
|---|---|---|---|---|---|---|---|---|---|
| H$^{13}$CN | $1-0$ | 86.3401764 | ... | ... | ... | ... | ... | ... | $1.6 \times 10^5$ |
| H$^{13}$CO$^+$ | $1-0$ | 86.7542880 | 4.16 | $3.9 \times 10^{-5}$ | $2.3 \times 10^{-10}$ | $9.3 \times 10^{-10}$ | $1.7 \times 10^5$ | $4.1 \times 10^4$ | $2.2 \times 10^4$ |
| SiO | $2-1$ | 86.8469950 | 6.25 | $2.9 \times 10^{-5}$ | $1.1 \times 10^{-10}$ | $2.8 \times 10^{-10}$ | $2.7 \times 10^5$ | $1.0 \times 10^5$ | ... |
| HN$^{13}$C | $1-0$ | 87.0908590 | ... | ... | ... | ... | ... | ... | ... |
| CCH | $N = 1-0, J = 3/2-1/2$ | 87.3169250 | 4.193 | $1.5 \times 10^{-6}$ | $1.4 \times 10^{-11}$ | $5.8 \times 10^{-11}$ | $1.1 \times 10^5$ | $2.7 \times 10^4$ | ... |
| CCH | $N = 1-0, J = 1/2-1/2$ | 87.4020040 | 4.197 | $1.3 \times 10^{-6}$ | $8.5 \times 10^{-12}$ | $5.9 \times 10^{-11}$ | $1.5 \times 10^5$ | $2.2 \times 10^4$ | ... |
| HNCO | $J_{K_a,K_c} = 4_{0,4} - 3_{0,3}$ | 87.9252380 | 10.55 | $9.0 \times 10^{-6}$ | $6.8 \times 10^{-11}$ | $4.3 \times 10^{-10}$ | $1.3 \times 10^5$ | $2.1 \times 10^4$ | ... |
| HCN | $1-0$ | 88.6318473 | 4.25 | $2.4 \times 10^{-5}$ | $1.9 \times 10^{-11}$ | $8.1 \times 10^{-11}$ | $1.3 \times 10^6$ | $3.0 \times 10^5$ | $4.5 \times 10^3$ |
| HCO$^+$ | $1-0$ | 89.1885260 | 4.28 | $4.3 \times 10^{-5}$ | $2.3 \times 10^{-10}$ | $9.2 \times 10^{-10}$ | $1.8 \times 10^5$ | $4.6 \times 10^4$ | $5.3 \times 10^2$ |
| HNC | $1-0$ | 90.6635640 | 4.35 | $2.7 \times 10^{-5}$ | $9.0 \times 10^{-11}$ | $2.6 \times 10^{-10}$ | $3.0 \times 10^5$ | $1.1 \times 10^5$ | $2.3 \times 10^3$ |
| H | $41\alpha$ (42 − 41) | 92.0344340 | ... | ... | ... | ... | ... | ... | ... |
| N$_2$H$^+$ | $1-0$ | 93.1737770 | 4.47 | $3.6 \times 10^{-5}$ | $2.3 \times 10^{-10}$ | $8.9 \times 10^{-10}$ | $1.6 \times 10^5$ | $4.1 \times 10^4$ | $5.5 \times 10^3$ |
| C$^{34}$S | $2-1$ | 96.4129500 | ... | ... | ... | ... | ... | ... | ... |
| CH$_3$OH-E | $J_K = 2_{-1}-1_{-1}$ | 96.7393630 | 12.5 | $2.6 \times 10^{-6}$ | $9.0 \times 10^{-11}$ | $2.9 \times 10^{-10}$ | $2.8 \times 10^4$ | $8.7 \times 10^3$ | ... |
| CH$_3$OH-A | $J_K = 2_0-1_0$ | 96.7413770 | 7.0 | $3.4 \times 10^{-6}$ | $1.0 \times 10^{-10}$ | $3.5 \times 10^{-10}$ | $3.4 \times 10^4$ | $9.8 \times 10^3$ | ... |
| CS | $2-1$ | 97.9809530 | 7.1 | $1.7 \times 10^{-5}$ | $4.6 \times 10^{-11}$ | $1.6 \times 10^{-10}$ | $3.6 \times 10^5$ | $1.0 \times 10^5$ | $1.2 \times 10^4$ |
| SO | $J_K = 3_2-2_1$ | 99.2999050 | 9.2 | $1.1 \times 10^{-5}$ | $3.9 \times 10^{-11}$ | $3.4 \times 10^{-10}$ | $2.9 \times 10^5$ | $3.3 \times 10^4$ | ... |
| HC$_3$N | $12-11$ | 109.1736380 | 34.058 | $1.0 \times 10^{-4}$ | $5.2 \times 10^{-11}$ | $6.4 \times 10^{-10}$ | $2.0 \times 10^6$ | $1.6 \times 10^5$ | $1.1 \times 10^5$ |
| C$^{18}$O | $1-0$ | 109.7821760 | 5.27 | $6.3 \times 10^{-8}$ | $3.2 \times 10^{-11}$ | $1.3 \times 10^{-10}$ | $1.9 \times 10^3$ | $4.8 \times 10^2$ | ... |
| HNCO | $J_{K_a,K_c} = 5_{0,5} - 4_{0,4}$ | 109.9057530 | 15.82 | $1.8 \times 10^{-5}$ | $7.4 \times 10^{-11}$ | $4.4 \times 10^{-10}$ | $2.4 \times 10^5$ | $4.1 \times 10^4$ | ... |
| $^{13}$CO | $1-0$ | 110.2013540 | 5.29 | $6.3 \times 10^{-8}$ | $3.2 \times 10^{-11}$ | $1.3 \times 10^{-10}$ | $1.9 \times 10^3$ | $4.8 \times 10^2$ | ... |
| C$^{17}$O | $1-0$ | 112.3589880 | 5.39 | $6.7 \times 10^{-8}$ | $3.2 \times 10^{-11}$ | $1.3 \times 10^{-10}$ | $2.1 \times 10^3$ | $5.2 \times 10^2$ | ... |
| CN | $N = 1-0, J = 1/2-1/2$ | 113.1913170 | 5.43 | $1.2 \times 10^{-5}$ | $7.2 \times 10^{-12}$ | $6.2 \times 10^{-11}$ | $1.6 \times 10^6$ | $1.9 \times 10^5$ | ... |
| CN | $N = 1-0, J = 3/2-1/2$ | 113.4909820 | 5.45 | $1.2 \times 10^{-5}$ | $7.2 \times 10^{-12}$ | $5.0 \times 10^{-11}$ | $1.7 \times 10^6$ | $2.4 \times 10^5$ | $1.7 \times 10^4$ |
| CO | $1-0$ | 115.2712020 | 5.53 | $7.2 \times 10^{-8}$ | $3.2 \times 10^{-11}$ | $1.3 \times 10^{-10}$ | $2.2 \times 10^3$ | $5.7 \times 10^2$ | ... |

*Note.* [a]The two-level approximation of the critical density has been calculated by accounting for only the single downward collisional rate coefficient from the initial to final energy level ($C_{ul}^{2lvl}$ and $n_{crit}^{2lvl}$). The full critical density calculation accounts for all the possible downward transition collisional rate coefficients to and from the final energy level ($C_{ul}$ and $n_{crit}$). The effective excitation densities ($n_{eff}$) are taken from Shirley (2015), and have been defined by radiative transfer modelling as the density that results in a molecular line with an integrated intensity of 1 K km s$^{-1}$ (also see Evans 1999).

transitions with substructure in order to contain all lines), to grid all data to a common set of velocity channels of 0.6 km s$^{-1}$ width, and to subtract a spectral baseline of third order from the data. Velocities in the range $-50$ to $+80$ km s$^{-1}$ are ignored when fitting the baseline. This range is further expanded in case of emission lines with hyperfine structure, provided this line substructure is more compact than 20 km s$^{-1}$. The original velocity range of $-50$ to $+80$ km s$^{-1}$ is not modified for baseline fitting in cases where line substructure extends over a range $>20$ km s$^{-1}$: we hypothesize that line substructure spread out over a large velocity range is less likely to bias baseline fits, and the absence of substantial artefacts in manually inspected processed data validates this approach. Future versions of the pipeline may include more sophisticated algorithms for automatic flagging of significant line emission during baseline fitting, and future publications will include a detailed assessment of the baseline quality. The spectra delivered by the telescope's data processing system are calibrated in forward-beam brightness temperatures (the $T_A^*$-scale). We convert these to the main beam brightness scale using the relation $T_{mb} = F_{eff}/B_{eff} \cdot T_A^*$. We adopt frequency-independent forward and main beam efficiencies of $F_{eff} = 0.95$ and $B_{eff} = 0.8$, respectively, consistent with the available calibration data.[2] The baselined data were eventually gridded into a data cube with pixels of 10 arcsec × 10 arcsec size. This effectively also averages data from the two

independent polarizations in every part of the map. This step uses a frequency-dependent Gaussian gridding kernel that increases the intrinsic resolution of the data (i.e. impact of frequency-dependent telescope beam and gridding) to 30 arcsec.

The angular resolution of our data is, however, further reduced due to the high scan speed and slow sampling pattern employed by our project. We pursue this fast mapping strategy because it allows us to cover a large area on the sky at the expense of somewhat reduced angular resolution. To be specific, along the scan direction, we obtain one data sample every 90 arcsec s$^{-1}$ · 0.25 s = 22.5 arcsec. The impact of this sampling pattern on the angular resolution can, to first order, be gauged by modelling it by a Gaussian smoothing kernel of 22.5 arcsec width at half peak value that is extended along the scan direction. The resulting angular resolution in this model – along the sampling direction of a single scan – would be ([30 arcsec]$^2$ + [22.5 arcsec]$^2$)$^{1/2}$ = 37.5 arcsec. The angular size of the beam would be smaller when averaging perpendicular scans. A future data-oriented LEGO paper will provide a more detailed discussion of this issue.

To allow direct comparison to the molecular hydrogen column density and dust temperature maps (see Section 3.1), we further smooth data cubes with a Gaussian kernel of 52 arcsec width at half peak value to achieve an angular resolution of 60 arcsec, and

re-sample the data set on to the same 20 arcsec pixel size spatial grid. Further smoothing the data set also has the benefit of increasing the signal-to-noise of the lines across the mapped region, whilst somewhat mitigating the beam smearing effect caused by the fast mapping speed.

### 2.1.2 Selection of emission lines

The LEGO pipeline generates data cubes spanning a few hundred channels for a selected set of emission lines. This approach is necessary because processing the full data set would be computationally too demanding. Table 2 lists the transitions for which data cubes are produced. Additional details on quantum numbers and frequencies are presented in Table A1. All of the lines are emitted by molecules, with the exception of the H41$\alpha$ hydrogen recombination line.

The catalogue is systematically constructed using the following steps. At a minimum, LEGO should cover all emission lines that can be studied in nearby galaxies. We thus start our line catalogue by including all lines that Watanabe et al. (2014) detect in M51 at a signal-to-noise ratio >2. To this we add all transitions studied by Pety et al. (2017) in Orion B, to be consistent with their work. We then include rare isotopologues of species where the abundant line is potentially optically thick. (This is not relevant in practice, since rare forms of HCN, HNC, and HCO$^+$ are already included due to previous selection steps.) We further include bright lines from selected species that intuitively seem interesting, such as long and simple carbon chains. In practice, these lines are taken from Watanabe et al. (2015). Finally, the lowest-lying hydrogen $\alpha$ recombination line above a frequency of 86 GHz is included in the table.

We are dealing with a large number of transitions, and the nature of our experiments does not require a particularly high level of accuracy in transition properties. We, therefore, obtain transition parameters from catalogues that cover a large number of lines, but we do not make any effort to refine the accuracy of parameters on the basis of the latest literature. Rest frequencies are taken from the Lovas/NIST data base (Lovas 2004), as accessed via the Splatalogue[5] data base for astronomical spectroscopy. This catalogue is also used to characterize the substructure of a transition. Provided splitting is present in a given transition, the emission lines were split into groups of <100 MHz width (i.e. <300 km s$^{-1}$ · [$\nu$/100 GHz]$^{-1}$ in velocity). The group's rest frequency, $\nu_{rest}$, is taken to be the frequency of the line with the highest 'intensity' in the Lovas (2004) catalogue. The full quantum numbers of this transition are recorded in Table A1. The lowest and highest frequencies of lines in a given group (i.e. $\nu_{min}$ and $\nu_{max}$) are also collected in Table A1, for example, to guide baseline subtraction and to build data cubes that are wide enough to contain all transitions within a given line group. Table 2 lists the quantum numbers common within a given line group, as well as $\nu_{rest}$.

The Leiden Atomic and Molecular Database (Schöier et al. 2005; LAMDA, accessed in November 2019) is used to obtain the upper energy levels, Einstein coefficients and downward collisional rates coefficients for the lines covered by LEGO. When calculating the collisional rates coefficients, we ignore line splitting, where present, and use LAMBDA data files collapsing transition substructure in those cases. The critical density for each of the transitions has been calculated using two approximations. The first is the simple two-level approximation, which only accounts for the

downward collisional rate of the initial upper ($u$) to final lower ($l$) energy level such that $n_{crit}^{2lvl} = A_{ul}/C_{ul}$. The second approximation accounts for all the possible upward and downward transition collisional rate coefficients from the initial energy level, which when neglecting background contribution and assuming optically thin conditions can be defined as $n_{crit} = A_{ul}/\sum_{u \neq k} C_{uk}$ (subscript $uk$ represents a transition from the 'upper' energy level, $u$, to any allowed energy level, $k$, even where $E_k > E_u$). Collisional rates of upward transitions have been calculated using the formalism outlined by Shirley (2015, equation 4). In both the calculations, we use the collision rates determined at a kinetic temperature 20 K, assuming that $H_2$ is the dominant collisional partner, and assume that the emission is optically thin. These results are presented in Table 2.

Also given in Table 2 are the effective excitation densities ($n_{eff}$) taken from Shirley (2015). These have been determined using the RADEX radiative transfer modelling code (van der Tak et al. 2007), and are defined as the density that results in a molecular line with an integrated intensity of 1 K km s$^{-1}$ (also see Evans 1999). These estimates then account for radiative trapping effects that lower the critical density within the optically thick regime. In Table 2, we show $n_{eff}$ for a kinetic temperature 20 K and the reference column densities for each molecule from Shirley (2015, table 1).

### 2.1.3 Spectra, noise properties, and integrated intensity maps

A brief inspection of the cubes shows that they contain a varying degree of complexity in their position–position–velocity space structure. Fig. 2 shows the spectrum from each molecular transition averaged across the mapped region. Here, we can see that several of the lines contain multiple velocity components, which vary in spatial structure and peak intensity (also see channel maps presented in Fig. D1).

Highlighted at the top of Fig. 2 is the range of velocities ($\sim -10$ to 20 km s$^{-1}$) observed towards W49A, and represented as a vertical dashed black line is the mean systemic velocity of W49A ($\sim 11$ km s$^{-1}$; e.g. Nagy et al. 2012, 2015). Within this velocity range, we find asymmetric line profiles, and, in some cases, multiple narrowly separated (<10 km s$^{-1}$) velocity components. These line profiles are likely due to the unresolved internal structure and kinematics (e.g. outflows) within W49A, which is currently being heavily influenced by stellar feedback (e.g. see Galván-Madrid et al. 2013). Also labelled in Fig. 2 are examples of several more broadly spaced (>10 km s$^{-1}$) velocity components from sources unrelated to W49A, which, given their velocities, are thought to arise from emission in intervening spiral arms (e.g. W49B at $\sim 60$ km s$^{-1}$).

There are a couple of additional potential causes for the multiple velocity components seen in averaged spectra. The first is hyperfine splitting of molecular transitions (also see Table A1). Labelled in Fig. 2 are several examples for lines that have narrow (e.g. $N_2H^+$) and broadly (e.g. CCH) spaced hyperfine transitions. The second is self-absorption within regions of high optical depth, which can result in the narrow (<10 km s$^{-1}$) spaced velocity components. This would be particularly relevant towards the centre of W49A, and for the more abundant molecules (e.g. CO; Nagy et al. 2012; Galván-Madrid et al. 2013).

The final feature to note within Fig. 2, is the blending of the $CH_3OH$ spectra. The $CH_3OH$-E and $CH_3OH$-A transitions shown here are very close in frequency (corresponding to a velocity
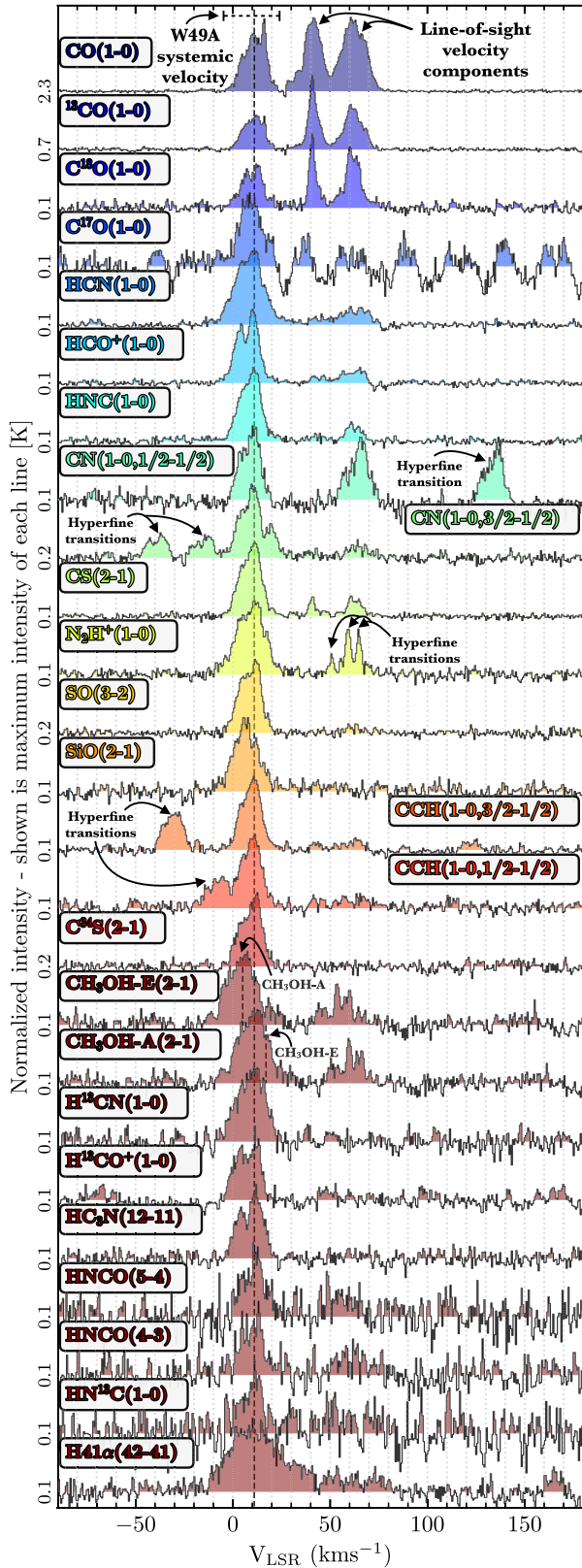
difference of $\sim 5$ km s$^{-1}$), and cannot be fully resolved towards W49A (also see Fig. 3).[6] Further analysis of the CH$_3$OH lines in this work is, therefore, limited to the significantly brighter A-type line, and all references to CH$_3$OH (2-1) henceforth correspond to the CH$_3$OH-A (2-1) line.

Despite identifying several velocity components within the spatially averaged spectra, when examining the cubes we find that, in general, there is a single velocity component that dominates the emission along a given line of sight. Such that the total emission at any position within a molecular line cube can typically be attributed to a single velocity range (see Fig. D1). It then appears the brightest sources within the mapped region are generally distinct in both velocity space and spatially. Throughout this work, we then make the simple assumption that at each position there is a single source that is responsible for the emission seen in molecular line emission. Additionally, the analysis presented in this work relies on a comparison to the total hydrogen column density and dust temperature maps, which also integrate emission from all molecular material along the line of sight (see Section 2.2). Hence, not limiting our analysis to a given velocity range allows us to make a consistent comparison across all the available data sets.

The aim of this work is to produce intensity maps integrated along all velocities that (a) trace the compact and bright emission, (b) recover the diffuse and extended emission, (c) have a constant noise profile where emission is not present. To do so, we follow a two-step masking procedure for the data cubes to reduce the noise within positions with significant line emission. Briefly, for this procedure, we firstly create a mask including only the voxels (velocity pixels) with significant CO emission, which we then apply to a given line cube. As the CO line shows significant emission within the majority of velocity channels where emission is found from the other lines, this masking procedure provides a convenient way of producing a generous integration velocity range at each line-of-sight such that both the noise is reduced in pixels containing emission and pixels without emission show a flat noise profile. In the case of lines with hyperfine structure that extends outside of the CO mask (e.g. CCH), we choose to expand the mask to also include voxels with significant emission from the line cube (i.e. producing a final mask where each position has either significant emission from the line hyperfine structure or CO emission).

We follow the method outlined by Dame (2011) to produce the initial CO data cube mask. For this method, we firstly smooth the cubes by a factor of two spatially (i.e. to 120 arcsec), and a factor of 10 spectrally (i.e. to 6 km s$^{-1}$). We then determine the rms value, $\sigma_{rms}$, along each line of sight using two spectral windows, covering velocities of $-150$ to $-20$ km s$^{-1}$ and 90–150 km s$^{-1}$. These velocity ranges were chosen for their lack of significant emission (see Fig. 2). An initial smoothed CO cube mask is then produced to include all channels above a $5\sigma_{rms}$ threshold. If hyperfine structure is present within a given cube, but outside of this CO mask, then this smoothing procedure is applied to the cube. Significant emission from this smoothed cube is then included in the final mask for that cube only. These final smoothed masks are then applied to the full resolution cubes, which are used to create the integrated intensity maps ($W_Q = \int T_{Q,mb} \, dv$).

**Figure 2.** Normalized spectra from each molecule averaged over the mapped region, for pixels with a integrated intensity $>3\,\sigma_{W_Q}$ (see Fig. 3). The peak brightness temperature of each spectrum is given on the left axis. The coloured shaded region of each spectrum shows $>0$ K intensities. Highlighted are several features of interest within the spectra, such as the systemic velocity of W49A, hyperfine transitions, and line blending.

---

[6]It is worth considering that there are several (typically weaker) CH$_3$OH transitions within the frequency range that we have not considered within this work. These include the 2(0)–1(0) and 2(1)–1(1) transitions of CH$_3$OH-E at rest frequencies of 96.744 545 GHz and 96.755 501 GHz, respectively (e.g. Menten et al. 1988; Leurini et al. 2004). This is noted within Table A1 within the Appendix of this work.
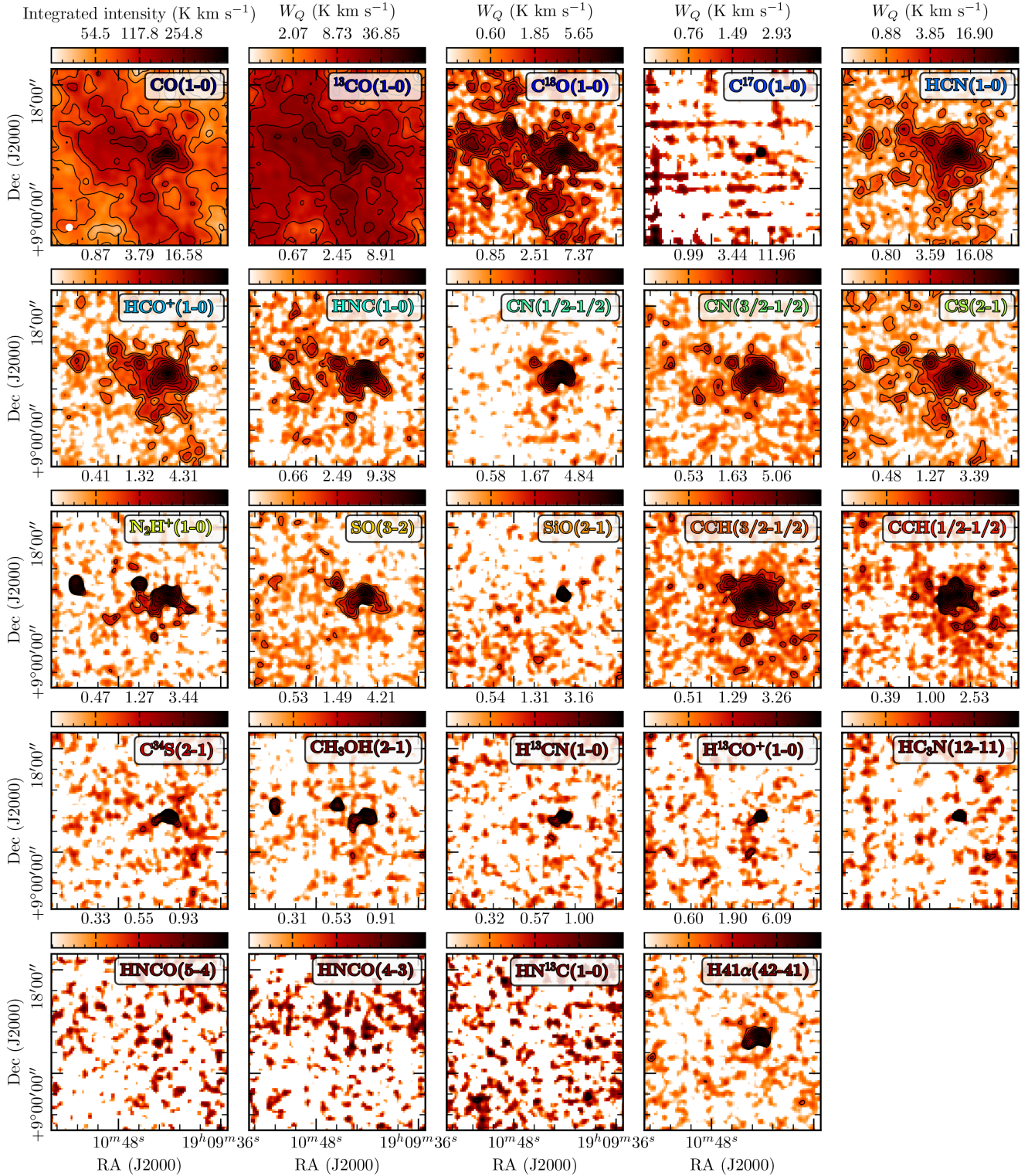
Integrated intensity (K km s$^{-1}$)
54.5  117.8  254.8

$W_Q$ (K km s$^{-1}$)
2.07  8.73  36.85

$W_Q$ (K km s$^{-1}$)
0.60  1.85  5.65

$W_Q$ (K km s$^{-1}$)
0.76  1.49  2.93

$W_Q$ (K km s$^{-1}$)
0.88  3.85  16.90

CO(1–0)   $^{13}$CO(1–0)   C$^{18}$O(1–0)   C$^{17}$O(1–0)   HCN(1–0)

0.87  3.79  16.58     0.67  2.45  8.91     0.85  2.51  7.37     0.99  3.44  11.96     0.80  3.59  16.08

HCO$^+$(1–0)   HNC(1–0)   CN(1/2–1/2)   CN(3/2–1/2)   CS(2–1)

0.41  1.32  4.31     0.66  2.49  9.38     0.58  1.67  4.84     0.53  1.63  5.06     0.48  1.27  3.39

N$_2$H$^+$(1–0)   SO(3–2)   SiO(2–1)   CCH(3/2–1/2)   CCH(1/2–1/2)

0.47  1.27  3.44     0.53  1.49  4.21     0.54  1.31  3.16     0.51  1.29  3.26     0.39  1.00  2.53

C$^{34}$S(2–1)   CH$_3$OH(2–1)   H$^{13}$CN(1–0)   H$^{13}$CO$^+$(1–0)   HC$_3$N(12–11)

0.33  0.55  0.93     0.31  0.53  0.91     0.32  0.57  1.00     0.60  1.90  6.09

HNCO(5–4)   HNCO(4–3)   HN$^{13}$C(1–0)   H41$\alpha$(42–41)

Dec (J2000)   18'00"   +9°00'00"

RA (J2000)   10$^m$48$^s$   19$^h$09$^m$36$^s$

**Figure 3.** Maps of the integrated intensity, $W_Q$, for the selected molecular transitions, $Q$, across the W49 region. Overlaid on each panel are signal-to-noise contours in levels of 5 $\sigma_{W_Q}$, increasing up to the maximum $\sigma_{W_Q}$ value within the mapped region (see Table 3). The angular beam size of 60 arcsec, or ~3 pc at the distance of W49A (~11 kpc; Gwinn et al. 1992; Zhang et al. 2013), is shown as a white circle in the lower left of the first panel.

**Table 3.** Observational properties across the mapped region (i.e. that covered with both vertical and horizontal on-the-fly scans). The columns show the molecule name, the average cube rms (in a $1\,\mathrm{km\,s^{-1}}$ channel), mean uncertainty of the integrated intensity, and mean value of the integrated intensity. The final column shows the area percentage, $A_{\mathrm{cov}}$, within the mapped region that has an integrated intensity above five times the uncertainty; i.e. $W_Q > 5\sigma_{W_Q}$. The table has been ordered by decreasing $A_{\mathrm{cov}}$ value (cf. table 8 of Pety et al. 2017). The information given is for maps that have been smoothed to an angular resolution of 60 arcsec, and have a spectral resolution of 0.6 km s$^{-1}$. See Table B1 for additional statistical properties of the molecular line integrated intensity maps. The full, machine-readable version of this table can be obtained from the supplementary online material.

| Line | $\sigma_{\mathrm{rms}}$ (1 km s$^{-1}$) (K) | $\sigma_W$ | $W_Q^{\mathrm{mean}}$ (K km s$^{-1}$) | $A_{\mathrm{cov}}$ (%) |
|---|---|---|---|---|
| CO (1–0) | 0.15 | 0.62 | 84.56 | 100.0 |
| $^{13}$CO (1–0) | 0.06 | 0.26 | 13.75 | 99.9 |
| C$^{18}$O (1–0) | 0.06 | 0.24 | 1.05 | 34.1 |
| HCN (1–0) | 0.05 | 0.22 | 1.38 | 31.9 |
| N$_2$H$^+$ (1–0) | 0.04 | 0.15 | 0.17 | 7.0 |
| ... | ... | ... | ... | ... |

The uncertainty on the integrated intensity at each pixel was calculated as,

$$\sigma_{W_Q} = \sigma_{\mathrm{rms}} \Delta v_{\mathrm{res}} (N_{\mathrm{ch}})^{1/2},  \tag{1}$$

where $\Delta_{v_{\mathrm{res}}}$ is the velocity resolution, and $N_{\mathrm{ch}}$ is the number of unmasked channels used to create the moment maps. Mean values for the $\sigma_{\mathrm{rms}}$ and $\sigma_{W_Q}$ for each molecular line are shown in Table 3. The final moment maps that are used throughout this work are displayed in Fig. 3.

### 2.2 Dust column density and temperature maps

The molecular hydrogen column density and dust temperature maps, together with their uncertainties, were obtained by fitting modified blackbody models on a pixel-by-pixel basis to the *Herschel* Space Observatory large program Hi-Gal data (Molinari et al. 2011). For that, we follow a similar procedure as described in Guzmán et al. (2015). We also subtract dust emission arising from diffuse gas in the Galactic plane attributable to clouds located in the foreground or background of W49 using the same method as in Guzmán et al. (2015). Here, however, the background is smoothed to a much larger angular scale of 660 arcsec (34 pc at a distance of 11 kpc), to approximately match the size of the mapped region. The subtracted maps balance preserving small-scale structures with an angular resolution equal to the longest *Herschel* wavelength map (32 arcsec), whilst removing large-scale variations due to the background (>660 arcsec). A smoother background would leave too much diffuse emission not associated with W49 but to the Galactic plane, and a background with more structure on smaller scales would filter out emission associated with W49 itself. The column density and temperature maps produced using this method, however, suffer from minor small-scale artefacts, such as saturated values (e.g. towards the centre of W49A). To correct for these issues, we spatially smooth the maps using the ASTROPY.CONVOLUTION package, which accurately interpolates over bad or missing values within an image. We find that a Gaussian kernal of 52 arcsec provided the optimal trade-off between final map resolution and a complete interpolation over missing values. The final molecular hydrogen column density and dust temperature have an angular resolution of 60 arcsec, and have

been resampled on to a 20 arcsec pixel grid. These maps are presented within the lower row of Fig. 1.

In Appendix C, we present an analysis of the column density and dust temperature uncertainties produced by random noise in the *Herschel* measurements. We find that mean fractional uncertainties across the mapped region studied in this work are $\Delta N_{\mathrm{H_2}}/N_{\mathrm{H_2}} = \pm 30$ per cent and $\Delta T_{\mathrm{dust}}/T_{\mathrm{dust}} = \pm 10$ per cent (see Fig. C2).

It is useful to convert the $H_2$ column density map into units of visual extinction, $A_{\mathrm{v}}$, for comparison with other works (e.g. Kauffmann et al. 2017). To do so, we adopt the commonly used extinction relation $A_{\mathrm{v}} = 1.1 \times 10^{-21} N_{\mathrm{H_2}}$, where $A_{\mathrm{v}}$ is in units of mag, and $N_{\mathrm{H_2}}$ is in units of cm$^{-2}$ (Bohlin, Savage & Drake 1978; Fitzpatrick 1999; Lacy et al. 2017).

## 3 RESULTS

In this section, we present the results from the molecular hydrogen column density, dust temperature, and molecular line integrated intensity maps. For ease of comparison, all the data presented in this work have been smoothed to the same angular resolution of 60 arcsec, which corresponds to a spatial resolution of ∼3 pc at the distance of the W49A star-forming region (11 kpc; Zhang et al. 2013). It is worth noting that, given the low spatial resolution of these observations, we are most likely not fully resolving any structures within our source; e.g the central star-forming core of the W49A is contained within a single pixel (e.g. Rugel et al. 2019). Each pixel within our maps is expected, therefore, to contain a range of physical properties (e.g. densities and temperatures), and should be considered as parsec scale averages across a representative star-forming region of the Galactic Plane. The maps presented throughout this work contain all emission integrated along the line of sight. Similar to the effect of the low spatial resolution, this may cause averaging of physical properties within any given pixel.

### 3.1 Distribution of the molecular hydrogen and dust temperature

In Section 2, we discussed the $H_2$ column density uncertainty produced by the random noise in the *Herschel* maps. There is, however, a second systematic uncertainty produced by the unrelated fore- and back-ground clouds along the line of sight. This has already been somewhat accounted for by subtracting the diffuse emission from the individual *Herschel* wavelength maps, yet unrelated Galactic plane emission may be still present within the final column density map.

To account for the remaining foreground and background contamination, we implement a location-dependent minimum $N_{\mathrm{H_2}}$ ($A_{\mathrm{v}}$) threshold for Galactic plane contamination. This threshold represents the Galactic background level on top of which the column density enhancements can be reliably attributed to the emission within the molecular line cubes. This threshold is defined, $N_{\mathrm{H_2}}^{\mathrm{thresh}}$ ($A_{\mathrm{v}}^{\mathrm{thresh}}$), such that 90 per cent of the pixels within an area of the Galactic plane have column densities less than this threshold (i.e. $N_{\mathrm{H_2}}$(90 per cent) < $N_{\mathrm{H_2}}^{\mathrm{thresh}}$). This analysis has been limited to the $-0.4 < l < 0.3°$, $43.5 < b < 47.5°$ region for the Galactic plane (see Fig. 1). This is a large section of the Galactic plane directly adjacent to the LEGO mapped region, and hence should provide a representative background Galactic plane column density distribution for the same latitude range of the mapped region.

The upper panel of Fig. 4 shows the cumulative (pixel) distribution of molecular hydrogen column density (extinction on upper axis) across the mapped region and the Galactic plane. Highlighted as a horizontal dashed line is the 90 percentile for the Galactic plane,
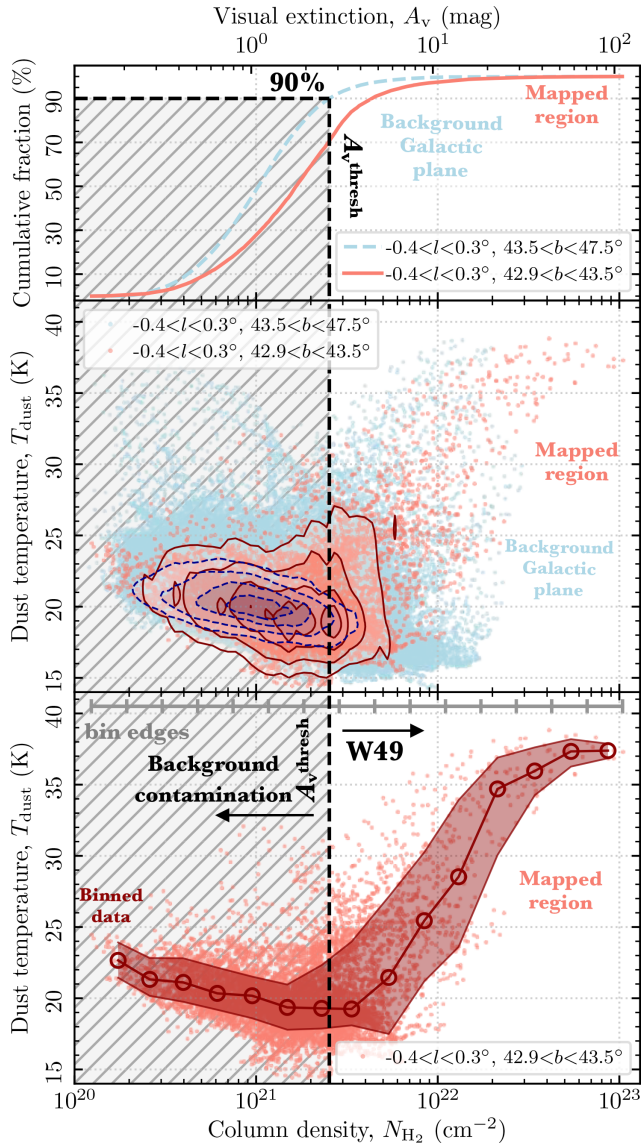
**Figure 4.** Line-of-sight contamination in the molecular hydrogen column density map. The upper panel shows the cumulative (pixel) distribution of molecular hydrogen column density (extinction on upper axis) across the mapped region (red) and Galactic plane (blue) region (see Fig. 1). Highlighted as a horizontal dashed line and shaded region is the 90 per cent percentile for the Galactic plane, which corresponds to a column density $N_{H_2}^{thresh} = 2.5 \times 10^{21}$ cm$^{-2}$ ($A_v^{thresh} = 2.7$ mag). The central panel shows the dust temperature distribution as a function of column density across both the mapped region and Galactic plane. The overlaid contours show the pixel density of all the points in levels of 10, 25, 50, and 75 per cent. The lower panel shows the dust temperature distribution as a function of column density for only the mapped region. The joined open circles show the median values within equally logarithmically spaced column density (extinction) bins, as shown at the top of the panel. The red shaded region shows the standard deviation within the bins.

which corresponds to a column density $N_{H_2}^{thresh} = 2.5 \times 10^{21}$ cm$^{-2}$ ($A_v^{thresh} = 2.7$ mag). We find minimum–mean–maximum values of the column density within the mapped region of $0.1$–$2.5$–$104.2 \times 10^{21}$ cm$^{-2}$, and dust temperatures of $14.4$–$20.6$–$38.8$ K. Across the Galactic plane region, we find minimum–mean–maximum values of $0.1$–$1.4$–$29.0 \times 10^{21}$ cm$^{-2}$, and $14.3$–$20.5$–$38.7$ K. Whilst above the

column density threshold within the mapped region we find $2.5$–$5.5$–$104.2 \times 10^{21}$ cm$^{-2}$, and $14.7$–$21.8$–$38.8$ K.

The central panel of Fig. 4 shows the dust temperature distribution as a function of column density for the mapped region and the Galactic plane. The overlaid contours highlight where the majority of data points lie within both distributions. Comparing these contours we see that both regions appear to be qualitatively similar, with the mapped region shifted higher by a factor of a few in column density. Moreover, below the column density threshold, there appears to be a shallow negative correlation between the column density and dust temperature for both regions. Such a negative correlation is typically expected for quiescent molecular gas, where the highest column densities have the lowest temperatures due to the lack of a heating source and efficient cooling mechanisms (i.e. via dust radiation). Given the aforementioned line-of-sight contamination at low column densities, it is, however, difficult to assess the significance of this relation.

Finally, the lower panel of Fig. 4 shows the dust temperature distribution as a function of column density for only the mapped region. This clearly shows that above the column density threshold there is an increase in the dust temperature with increasing column density, which originates from positions associated with the W49A massive star-forming region (cf. Fig. 1). This is then indicative of an internal heating source associated with star formation (an increased interstellar radiation field; e.g. Pety et al. 2017).

We now compare to an independent measure of the molecular hydrogen column density. In doing so, we aim to test the fore- and back-ground contamination threshold of $N_{H_2}^{thresh} = 2.5 \times 10^{21}$ cm$^{-2}$ ($A_v^{thresh} = 2.7$ mag) in the *Herschel* molecular hydrogen column density map. To do so, we follow the method presented in Heiderman et al. (2010, their equations 11–13) for converting the $^{13}$CO(1−0) integrated intensity into a $^{13}$CO column density. This method solves for the excitation temperature and optical depth of $^{13}$CO(1−0) using both $^{12}$CO(1−0) and $^{13}$CO(1−0) peak intensities, by assuming LTE conditions, that both lines have equal excitation temperatures, and that the $^{12}$CO(1−0) and $^{13}$CO(1−0) are optically thick and thin, respectively. We then convert to a molecular hydrogen column density assuming a $^{13}$CO abundance of $5 \times 10^{-6}$.

Fig. 5 shows the molecular hydrogen column density determined from the $^{13}$CO(1−0) observations, $N_{H_2(^{13}CO)}$, as a function of the molecular hydrogen column density determined from the *Herschel* observations, $N_{H_2}$. The overlaid solid red line shows the running median within logarithmically spaced $N_{H_2}$ bins, and the diagonal black dashed line shows $N_{H_2(^{13}CO)} = N_{H_2}$. Comparing these two lines, we find a remarkably good agreement between $N_{H_2(^{13}CO)}$ and $N_{H_2}$ across around two orders of magnitude in molecular hydrogen column density. This indicates that the method for subtracting foreground and background emission within in the *Herschel* maps has correctly removed the Galactic cirrus emission across the map, leaving only the emission associated with the molecular hydrogen. Moreover, this agreement between $N_{H_2(^{13}CO)}$ and $N_{H_2}$ highlights that the same sources within each line of sight are recovered within both the integrated intensity and the molecular hydrogen column density maps. This then confirms our choice to integrate over all velocities within the molecular line cubes to produce the integrated intensity maps (Section 2.1.3).

We find that the $N_{H_2(^{13}CO)}$ and $N_{H_2}$ in Fig. 5 begin to deviate by more than the uncertainty on the *Herschel* column density at around $N_{H_2} < 10^{21}$ cm$^{-2}$ (see Section C). This value is significantly below $N_{H_2}^{thresh} = 2.5 \times 10^{21}$ cm$^{-2}$, which highlights this as a reasonable threshold for the increased uncertainty $>30$ per cent within the *Herschel* molecular hydrogen map.
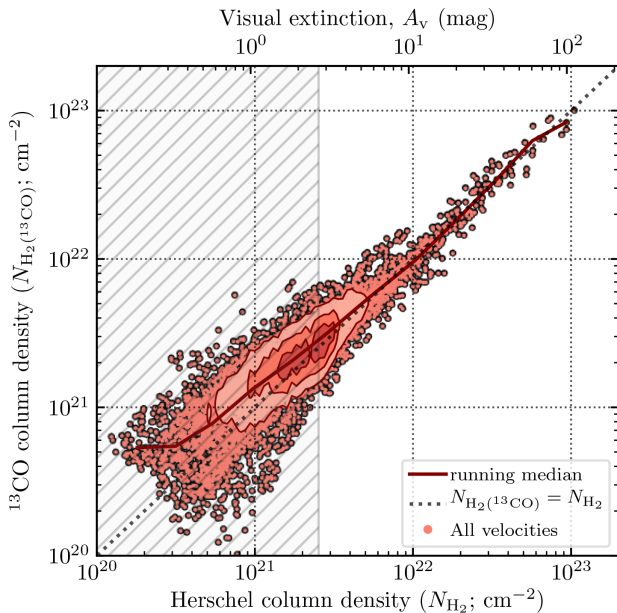
**Figure 5.** A comparison of the molecular hydrogen column density determined from the $^{13}$CO (1-0), $N_{H_2(^{13}CO)}$, and *Herschel*, $N_{H_2}$, observations. The overlaid contours show the pixel density of all the points at levels of 10, 25, 50, and 75 per cent. The solid red line shows the running median within equally logarithmically spaced *Herschel* column density (extinction) bins. We highlight here values of the column density below $N_{H_2}^{thresh} = 2.5 \times 10^{21}$ cm$^{-2}$ ($A_v^{thresh} = 2.7$ mag) that suffer from foreground and background contamination (see Section 3.1). The diagonal dashed line shows $N_{H_2(^{13}CO)} = N_{H_2}$. This figure highlights that (1) the *Herschel* column density reliably recovers only the molecular gas (i.e. not atomic or Galactic cirrus) down to a column density below $N_{H_2}^{thresh}$ (Section 2.2), and (2) integrating all emission within the molecular line cubes is well correlated with the molecular hydrogen column density (Section 2.1.3).

In summary, in this section, we have been able to differentiate the column density peaks within the mapped region from the systematically lower column density material associated with the background Galactic plane. We defined a column density threshold above which we expect that the measured column density and temperature values can be predominantly attributed to a single source along any given line of sight, and, therefore, a more direct comparison can be made to the molecular line integrated intensity maps. This limit is highlighted on all figures within this work that make use of the column density within the variable plotted on the *y*-axis. Any values below the threshold should be assumed to have an associated uncertainty of >30 per cent and, therefore, be taken with caution.

### 3.2 Distribution of the integrated intensities

The integrated intensity maps for each molecular transition are shown in Fig. 3 (as labelled). The logarithmic colour-scale for each panel has been adjusted to best display the features within each map. Signal-to-noise contours are overlaid at levels of $5\,\sigma_{W_Q}$, increasing up to the maximum $\sigma_{W_Q}$ value within the mapped region (see Table 3).

We find that all the lines have peak integrated intensities towards the centre of W49A, at a position of $19^h10^m13^s$, $9°06'16''$. This corresponds to the position of a well-studied ring-like cluster of ultra-compact H II regions, which is typically referred to as the 'Welch ring' (e.g. Welch et al. 1987; Rugel et al. 2019). This is in agreement with many extensive molecular line studies towards this central

region, who find that both higher-*J* transitions and less abundant isotopologues of the molecule also peak towards this position (e.g. Nagy et al. 2012, 2015; Galván-Madrid et al. 2013).

To assess the spatial extent of each line, we define its coverage as the percentage of pixels across the mapped region that have integrated intensities at least five times higher than their associated uncertainty values; $A_{cov} = A(W_Q > 5\,\sigma_{W_Q})$, where $A$ is the percentage of the total mapped area. We find that the CO lines are the most spatially extended across the mapped region, with both $^{12}$CO and $^{13}$CO covering 100 per cent and C$^{18}$O covering around $A_{cov} \sim 35$ per cent of the main mapped region (i.e. that covered with both vertical and horizontal on-the-fly scans). The next most extended lines are HCN, HCO$^+$ and CS, which each have coverages of between $A_{cov} \sim 20$–30 per cent. These are followed by the HNC, CN, SO, CCH, and N$_2$H$^+$ lines that cover around $A_{cov} \sim 10$ per cent of the mapped region. Finally, the remaining 14 lines studied in this work cover only a few per cent of the mapped region. These coverage values, along with the minimum, mean, and maximum integrated intensity values measured within each mapped region are summarized in Table 3, which is ordered by decreasing coverage values.

Fig. 6 presents the mean integrated intensity and coverage as a function of critical density. Shown as triangles and circles are the critical densities determined using the two-level approximation ($n_{crit}^{2lvl}$) and when accounting for all the possible downward collisional rates ($n_{crit}$), respectively. For the lines given by Shirley (2015), we also plot the effective excitation density ($n_{eff}$) (see Table 3). This figure shows that beyond the $^{12}$CO, $^{13}$CO, and C$^{18}$O lines there does not appear to be a clear correlation between the strength or coverage of an observed molecular transition and its critical or effective excitation density. This highlights that the simple approximations we have used to estimate the broad range of characteristic densities are not sufficient to explain the distribution of all the observed molecular lines. Such a result is not surprising, and could be easily explained by the lower abundance and, therefore, optical depth, of rarer molecules. This can be simply seen by studying the CO isotopologues in Fig. 6, which show a stark decrease in mean brightness and coverage with increasing rareness despite having very similar critical densities (i.e. $^{12}$CO, $^{13}$CO, C$^{18}$O, and C$^{17}$O). This behaviour is discussed further in Sections 5.1 and 5.2.

### 3.3 Molecular hydrogen column density versus integrated intensities

In Fig. 7, we present scatter plots of the integrated intensities as a function of molecular hydrogen column density (or extinction as shown on the upper *x*-axis; Section 2.2). The black and grey circular points show the data with an integrated intensity value above and below $5\sigma_{W_Q}$, respectively. We plot solid coloured lines that show the median values of the integrated intensity within equally spaced logarithmic column density (extinction) bins (shown at the top of each panel), where the shaded region shows the standard deviation range with each bin. To allow reliable detections of the integrated intensity down to low values of the column density, we stack the individual integrated intensity values within the logarithmic column density (extinction) bins. When doing so, the uncertainty on the median value, $\sigma_{bin}$(med.), is defined as the propagated uncertainty from all the points contained within the bin,

$$\sigma_{bin}(\text{med.}) = 1.25\sigma_{bin}(\text{mean}) = 1.25\frac{f_{os}}{n}\sqrt{\sum_i^n \left(\sigma_{W_Q}(i)\right)^2}, \quad (2)$$
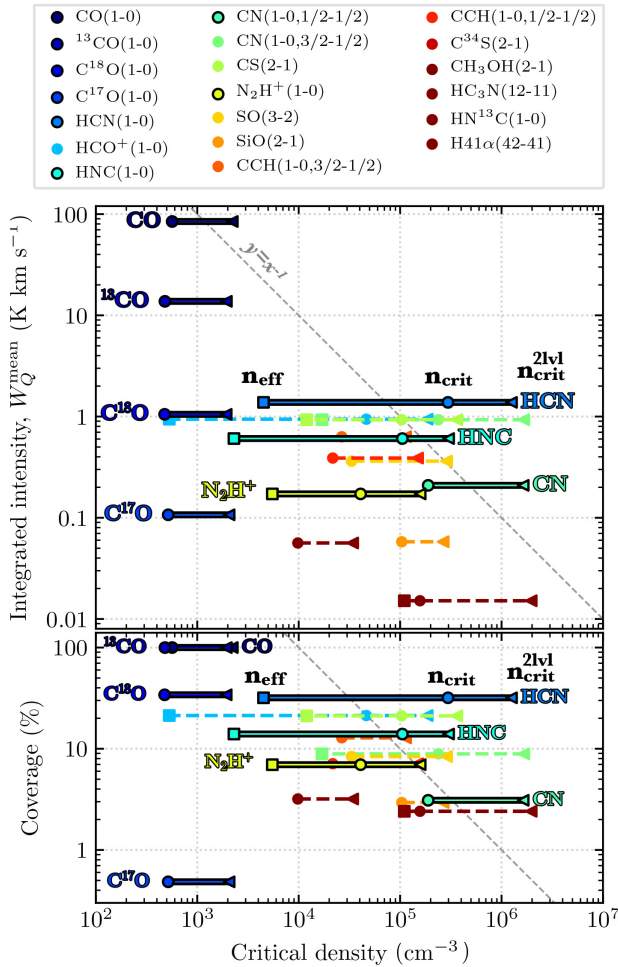
**Figure 6.** The mean integrated intensity (upper panel) and coverage (lower panel) of each molecular line observed as a function of the critical density. The triangles, circles, and squares show the critical density when using the two-level approximation ($n_{\rm crit}^{\rm 2lvl}$) and accounting for all the possible collisional rates ($n_{\rm crit}$), and the effective excitation density ($n_{\rm eff}$), respectively (see Table 3). The coverage is defined as the percentage of pixels within the mapped region that have an integrated intensity higher than at least five times its associated uncertainty ($W_Q > 5\sigma_{W_Q}$). The colours of each marker correspond to the molecular lines shown in the legend, and highlighted with labels in the figure are several molecular lines of interest. A linear $y \propto x^{-1}$ relationship has been plotted for comparison. This plot highlights that using the critical density as a predictor for line brightness or coverage across a mapped region is not trivial. Rather, the abundance of a given molecule should be taken into account when predicting its emission characteristics.

where $n$ is the number of points within the bin, and $\sigma_{W_Q}(i)$ is the uncertainty associated with the $i$th value within the bin (see equation 1). The oversampling factor is defined as the ratio of the angular beam area to the pixel area, $f_{\rm os} = 1.13\Omega_{\rm beam}/\Omega_{\rm pixel}$, and corrects for the number of independent measurements within each bin. Using the stacking method we find that even within bins that contain a significant fraction of points with $<5\sigma_{W_Q}$ (grey points), we can still extract a median integrated intensity value with a high level of significance. The result of this can be seen as the open coloured circles, which show the column density bins that have a median value above three times the associated uncertainty.

The first obvious trend from the binned lines in Fig. 7 is that on average there is a positive correlation between the integrated

intensities and the $H_2$ column density. Indeed, we find that for each line the highest value of the integrated intensity is found at the position of the highest column density. This shows that, in general, the integrated intensities are relatively well correlated with the amount of material along the line of sight. However, there are significant deviations from a simple linear relationship.

To investigate these deviations from a simple linear relation, we normalize the integrated intensity by the $H_2$ column density. Kauffmann et al. (2017) define this fraction $h_Q = W(Q)/N_{H_2}$, which can be thought of as a proxy for the line emissivity, or the efficiency of an emitting transition per $H_2$ molecule. Fig. 8 shows the emission efficiency ratio as a function of column density (extinction; cf. fig. 2 of Kauffmann et al. 2017). As in Fig. 7, we stack within logarithmically spaced column density bins to increase the significant detections at lower column densities. This is shown as the solid coloured lines and shaded regions, and the significant bins are shown with an open circle symbol. In Fig. 8, the $y$-axis of each panel has been normalized such that the maximum value of the significant logarithmically spaced bins is equal to unity.

We find that the $^{12}CO$, $^{13}CO$, and $C^{18}O$ show a near-constant decrease in $h_Q$ with increasing column density, which is to be expected based on the overall sub-linear relations shown in Fig. 7. Focusing on the significant median bins, we find that CS, $CH_3OH$, $HCO^+$, and $N_2H^+$ show $h_Q$ profiles that quickly rises from low column densities to maximum emission efficiency at a column density of several $10^{22}$ cm$^{-2}$, and slowly decrease towards higher column density values.

The profiles of HCN, HNC, CN, and CCH show more complex double-peaked profiles. These lines show an initially high value of $h_Q$ at the lowest values of the column density ($\sim 10^{20-21}$ cm$^{-2}$), that decreases with increasing column density to reach a minimum value at $\sim 10^{22}$ cm$^{-2}$. The $h_Q$ then rises again to a maximum value of at column density around few $10^{22}$ cm$^{-2}$. We highlight again here, however, that values of the column density below the threshold value of $N_{H_2}^{\rm thresh} \sim 10^{21}$ cm$^{-2}$ suffer from significant fore- and background contamination (see Section 3.1). Therefore, these multiple peaked profiles for the emission efficiency should be taken with caution.

## 4 ANALYSIS

### 4.1 Fraction of the emission over various column densities

In an attempt to determine the typical extinction (column density) that is traced by a given molecular line transition, Kauffmann et al. (2017) assess the cumulative fraction of the line intensity as a function of extinction (column density); $W_Q(N_{H_2} < N'_{H_2})$. Here we follow the same method and determine the cumulative distribution of the integrated intensity for each line. We note that, as we integrate all the intensity along the line of sight without isolating sources (e.g. W49A) in velocity, we do not have a single distance that could be reasonably applied to each pixel. Therefore, instead of producing the cumulative function of the luminosity as done in Kauffmann et al. (2017), we produce the cumulative function of the integrated intensity values across the mapped region. These results are displayed in Fig. 9, where we split the observed molecular lines between four panels for plotting clarity.

Shown as shaded regions in Fig. 9 is the uncertainty on each cumulative distribution, which is determined as the deviation in the distributions when adding synthetic noise to both the molecular line and column density maps. To produce these noise curves, we first include synthetic noise in the form of a Gaussian distributions centred on zero with widths of $\sigma_{W_Q}$ (Section 2.1.3) and $\Delta N_{H_2}$ to
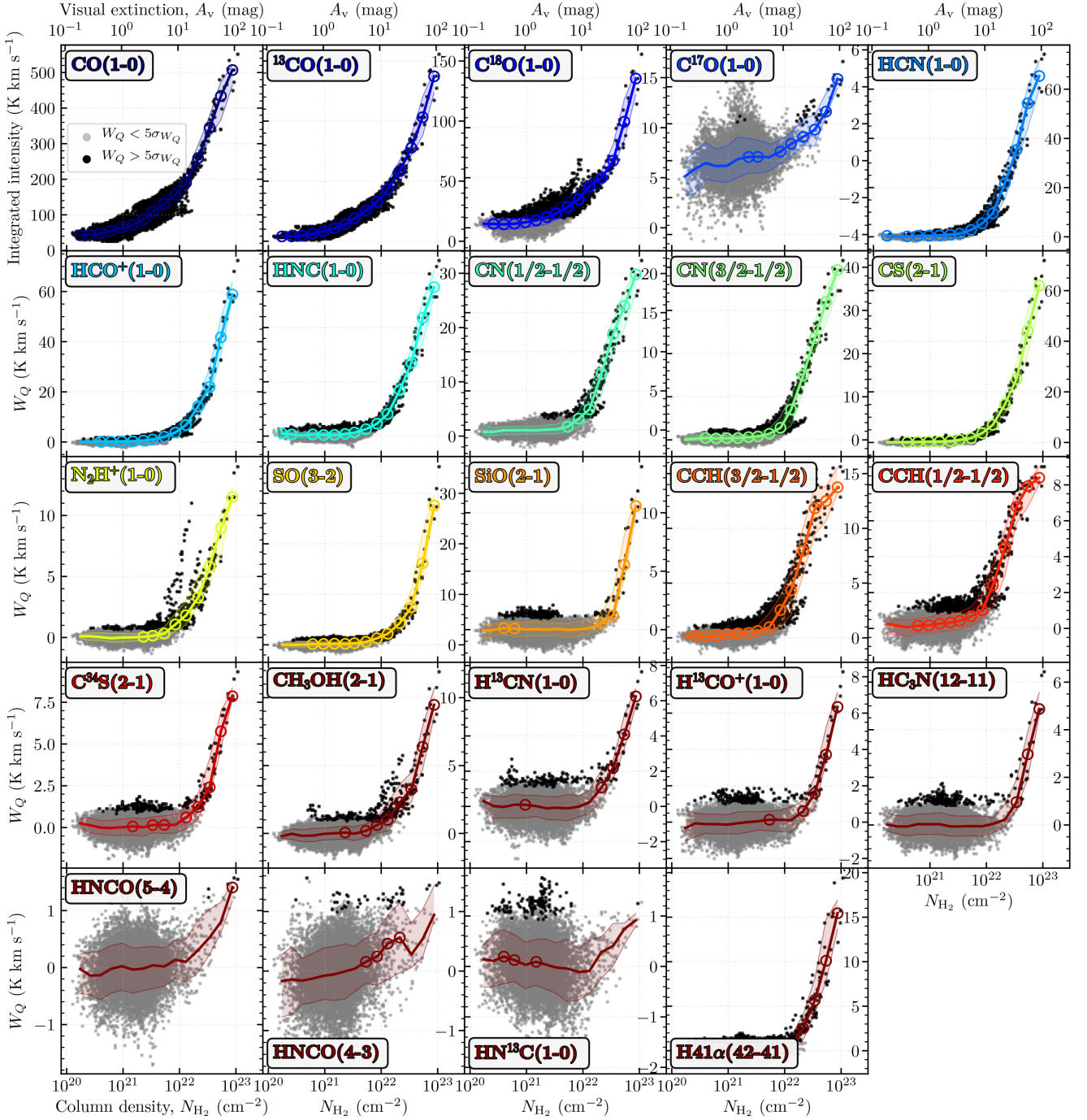
**Figure 7.** The integrated intensity of each molecular transition as a function of the molecular hydrogen column density derived from the mid-infrared *Herschel* observations. The upper *x*-axis of the top row of the panels shows the visual extinction, $A_v$ (see Section 2.2). The black and grey circular points show the data with an integrated intensity above and below $5\sigma_{W_Q}$, respectively (as shown in the box in the top left panel). The solid coloured lines show the median values of the integrated intensity within equally spaced logarithmic column density (extinction) bins. The circles show the bins that have a median value above $3\sigma_{\rm bin}$, where $\sigma_{\rm bin}$ represents the propagated uncertainty from all the points contained within the bin (see equation 2). The shaded region shows the standard deviation of all the points within each bin.

each of the integrated intensity maps and the column density map, respectively. This was repeated 1000 times to produce a sample of noisy maps. The shaded curves in Fig. 9 show the 5, 25, 75, and 95 percentiles of the resultant ensemble of cumulative distributions. Here the brightest lines (e.g. the CO lines) show a small uncertainty range (shaded region), as they are not significantly affected by the addition of $\sigma_{W_Q}$. To highlight the rigorous detection of these brighter

lines (e.g. $N_2H^+$), the inset panel shows the marginal increase in uncertainty with the addition of $10\sigma_{W_Q}$ noise. The weaker lines (e.g. $HC_3N$), on the other hand, show a larger variation due to the additional uncertainty, as the peak signal within these maps is closer to the $\sigma_{W_Q}$ level (compare integrated intensity uncertainty values given in Table 3 to integrated intensity statistics given in Table B1).
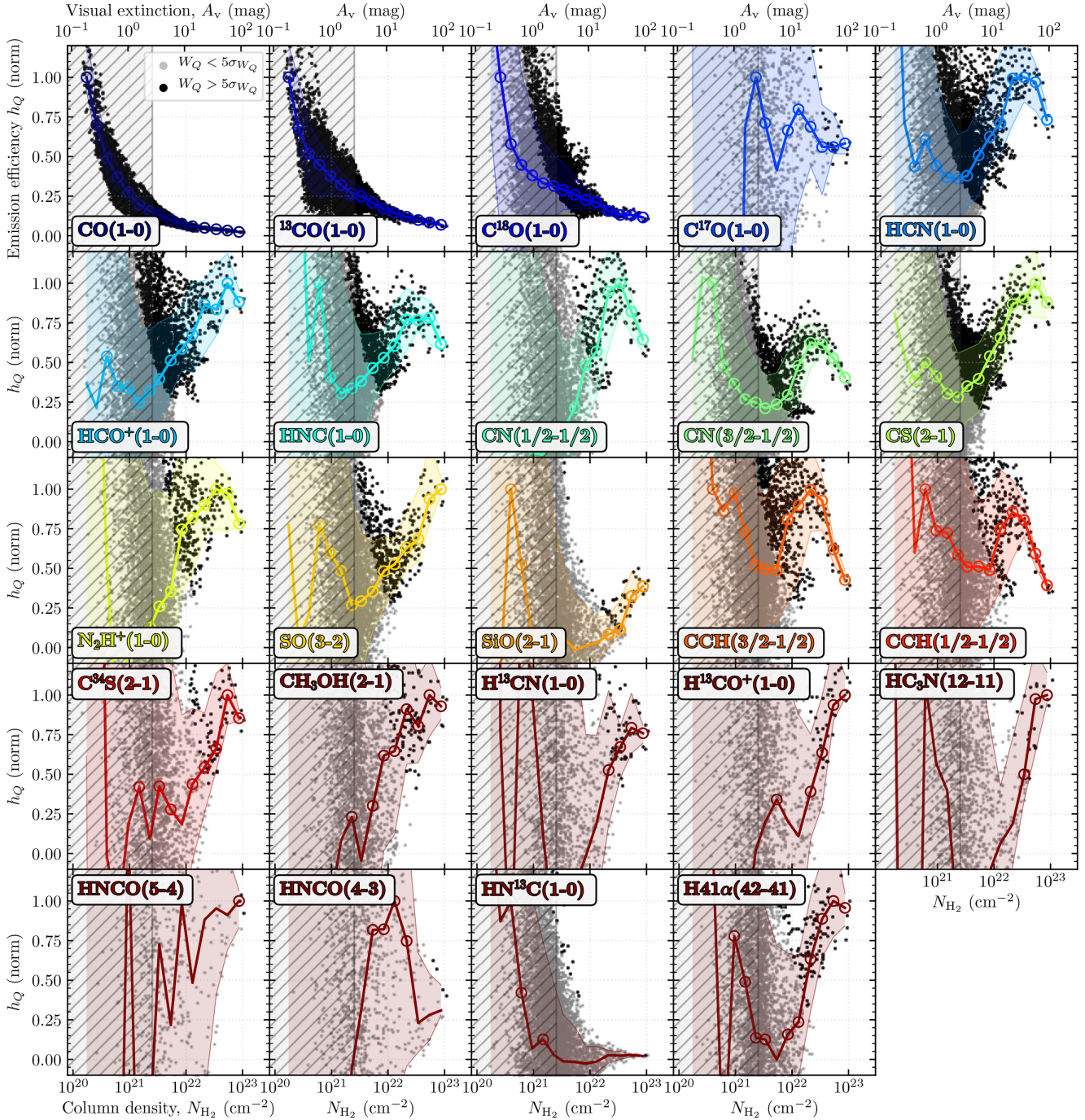
**Figure 8.** Same as Fig. 7 except that the *y*-axis shows the ratio of the integrated intensity to molecular hydrogen column density [$h_Q = W(Q)/N_{H_2}$]. We use this ratio as a proxy for the line emissivity, or the efficiency of an emitting transition per $H_2$ molecule. The *y*-axis has been normalized such that the maximum value of the logarithmically spaced bins is equal to unity. The shaded region highlights values of the column density below $N_{H_2}^{thresh} = 2.5 \times 10^{21}$ cm$^{-2}$ ($A_v^{thresh} = 2.7$ mag) that suffer from increased (>30 per cent) uncertainties due to foreground and background contamination (see Section 3.1). Values of the $h_Q$ below this threshold value should be taken with caution.

We also plot a dotted black line in Fig. 9 that shows the cumulative distribution of the pixels, and hence would represent the distribution of a line with a completely smooth integrated intensity map (i.e. a map containing pixels with only a single value). We plot the cumulative distribution of the column density values as a dashed black line in Fig. 9. Using these, we can assess how much of the molecular line intensity is contained below a given column density; or in the case of W49, how concen-

trated the emission is towards the column density peak in the W49A star-forming region. At any given column density, if the cumulative distribution profile resides between the smooth (dotted) and column density (dashed) profiles its remaining luminosity should be thought of as still being relatively evenly distributed across the mapped region (i.e. has a smoother integrated intensity map), whilst those to the right of the column density distribution should be thought of as having luminosities that are more
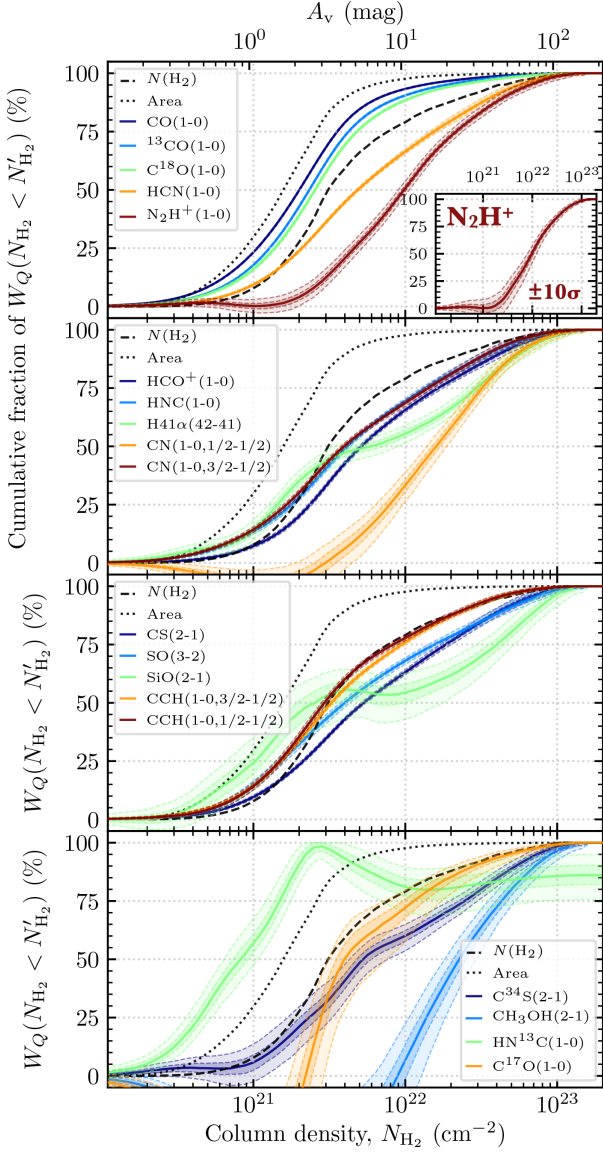
**Figure 9.** Cumulative fraction of the integrated intensity below a given column density for each molecular line. The solid line shows the observed value, and the shaded regions show the associated 5, 25, 75, and 95 percentile uncertainty limits (see Section 4.1). The black dotted line shows the cumulative distribution of the number of pixels, which can be thought of as a completely smooth distribution (i.e. a map containing only a single value). The dashed black line shows the cumulative distribution of the column density values. The lines have been split over four panels for plotting clarity. The inset panel shows the addition of a $10\,\sigma_{W_Q}$ noise to the $N_2H^+$ map.

strongly concentrated (i.e. has a more peaked integrated intensity map).

Fig. 9 shows that the cumulative distributions vary significantly among the different lines. We find that the CO lines are typically less concentrated than the column density (mass) distribution, whereas the remaining lines are typically more concentrated.

In Fig. 10, we show the fraction of the total intensity, $W_Q^{frac}$, from each of the molecular lines within four column density (extinction) masks, which have been chosen for direct comparison with Pety et al. (2017, their fig. 7). These authors suggest that within the Orion molecular cloud these ranges of column density (or extinction) can be used to represent diffuse (1 mag < $A_V$ <
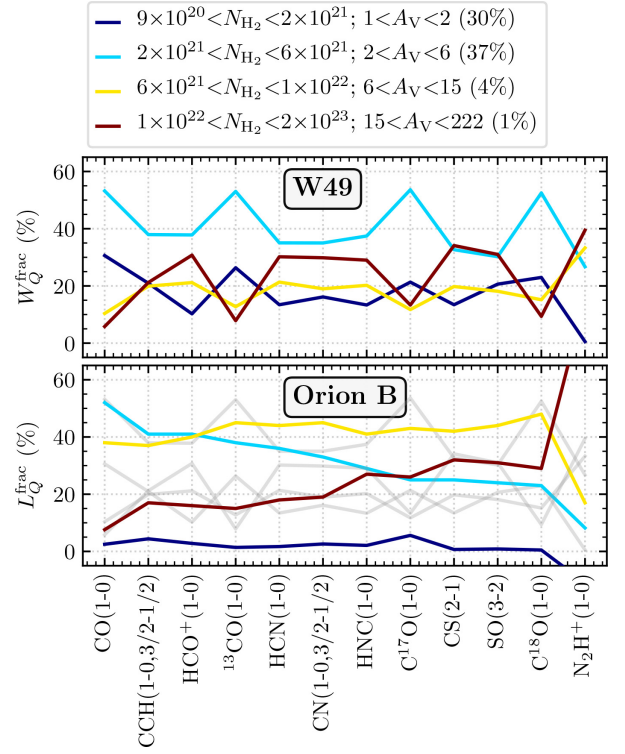


**Figure 10.** The fraction of the total intensity, $W_Q^{frac}$, from each molecular line within four column density masks, as shown in the legend above the plot (see Section 4.1). Also shown in the legend above the plot is the fraction of the total area covered by each of the column density (extinction) masks. The upper panel shows the intensity fraction determined across the W49 massive star-forming region studied in this work, and the lower panel shows the luminosity fraction, $L_Q^{frac}$, from the local star-forming region Orion B taken from Pety et al. (2017, as shown in their fig. 7). For comparison, the W49 results have also been plotted on the lower panel as grey lines.

2 mag), and translucent (2 mag < $A_V$ < 6 mag) gas, the environment of filaments (6 mag < $A_V$ < 15 mag), and dense gas (15 mag < $A_V$ < 222 mag).[7] We find that the CO lines have the majority of their intensity ($W_Q^{frac} \sim 80$ per cent) within the lowest two column density bins, which combined includes the majority of the mapped area (~70 per cent). The HCN, HCO$^+$, HNC, CS, and CN lines show a stark decrease in the intensity within the lowest column density bin ($W_{^{12}CO}^{frac} \sim 30$ per cent, $W_{HCN}^{frac} \sim 10$ per cent), and an increase within the highest column density bin ($W_{^{12}CO}^{frac} \sim 5$ per cent, $W_{HCN}^{frac} \sim 30$ per cent). The $N_2H^+$ shows an even higher fraction of emission towards the two highest bins, with $W_{N_2H^+}^{frac} \sim 40$ per cent of the total emission within the highest column density bin that accounts for only around 1 per cent of the total mapped area.

These results are compared to the Orion B star-forming region (Pety et al. 2017), which is shown in the lower panel of Fig. 10 (also with the W49 results in grey). We see that there are several notable differences between the two regions. First, there appears to be a lower fraction of the intensity within the dense gas for the CO lines in W49 when compared to Orion. This is expected given that

---

[7]The 1 mag < $A_V$ <2 mag and 2 mag < $A_V$ <6 mag extinction bins contain emission from below our extinction threshold of $A_v^{thresh} = 2.7$ mag (see Section 3.1). Fig. 5 shows that across $1< A_V$ <2.7 mag $N_{H_2(^{13}CO)}$ deviates from $N_{H_2}$ by 30–40 per cent, which we suggest sets an upper limit to the uncertainty on $W_Q^{frac}$ due to foreground and background contamination within this extinction range.

we observe a much larger spatial coverage within W49 ($\sim 10^4$ pc$^2$) than the Orion observations ($\sim 10$ pc$^2$), and hence recover a larger fraction of diffuse gas (relative to the dense gas). Secondly, we see that lines such as CN, HNC, HCO$^+$, and HCN have a marginally higher fraction of the intensity within dense gas. Lastly, we observe a significant difference in the fraction of dense gas for the N$_2$H$^+$ line, which has more than 80 per cent of the luminosity within the high-density regime in Orion compared to around 40 per cent across the mapped region. Watanabe et al. (2017), and Nishimura et al. (2017) conducted similar analyses of the 3 mm lines within the W51 and W3(OH) regions within the Galaxy. These authors also find results broadly in agreement with those found here. For example, Nishimura et al. (2017) classify their observed distributions into the three types: (i) species concentrated just around the centre of the star-forming core in W3(OH), such as CH$_3$OH, SO, and HC3N, (ii) species loosely concentrated around the centre, such as HCN, HCO$^+$, HNC, and CS, and (iii) species extended widely, such as CCH, C$^{18}$O, and $^{13}$CO.

### 4.2 Line emissivities as a function of both molecular hydrogen column density and dust temperature

So far in this work, we have investigated the variation of the emission from each molecular line as a function of the column density (extinction) and the critical or effective excitation density. However, as Pety et al. (2017) showed within the Orion B region, the temperature is also a fundamental parameter, as it also is closely linked to both the excitation of molecular lines and the formation chemistry (i.e. abundances). In this section, we investigate the simultaneous effects of the molecular hydrogen column density and dust temperature on the total line emission and emission efficiency. We remind the reader that spatial distributions of both the molecular hydrogen column density and dust temperature are presented in Fig. 1, and a comparison of the W49 and a representative portion of the Galactic plane is given in Fig. 4 (see Section 3.1). It is worth noting here, however, that the dust temperature may not be equal to the kinetic temperature of the gas. Without a probe of the latter within our observations (e.g. NH$_3$ or H$_2$CO), we work under the assumption the dust and gas temperatures are coupled across the mapped region.

To conduct this analysis, we first take the integrated intensity ($W_Q$) and emission efficiency ($h_Q$) maps and identify the value of the column density and dust temperature for each position. We then create evenly spaced bins for the observed range of column density and dust temperature, into which we assign the $W_Q$ and $h_Q$ values based on their column densities and dust temperatures (i.e. similar to the binning process outlined in Section 3.3). Finally, we calculate the summed $W_Q$ and median $h_Q$ values within each column-density temperature bin, and calculate the uncertainty on this value using equation (1) (see Section 3.3). We note that both HNCO lines and the C$^{17}$O and HN$^{13}$C line have been omitted from this analysis due to their lack of significant column density and temperature bins.

Figs 11 and 12 show the parameter space spanned by the dust temperature and column density (extinction). The overlaid contours show the 1, 10, 20, 30 ... 80, and 90 per cent levels of the maximum summed $W_Q$ and median $h_Q$ for a given line across the significant bins within the parameter space ($5\sigma_{\rm bin}$; see equation 2). These plots highlight the environmental properties which are producing the majority of the molecular line emission, and the conditions which are most efficiently producing this emission, respectively.

Fig. 11 shows that many of the lines emit the majority of their emission from moderate temperatures ($\sim 20$ K), and moderate column densities ($\sim 10^{21-22}$ cm$^{-2}$) following the peak in the number of pixels within the parameter space (compare to Fig. 4). In addition, many of the lines also peak towards the highest column density and temperature region of the parameter space ($\sim 20$ K; $\sim 10^{23}$ cm$^{-2}$). This region of the parameter space contains significantly fewer pixels, emphasizing that this peak is due to the concentration of bright pixels towards the centre of the W49A massive star-forming region.

Comparing how the total emission (Fig. 11) and emission efficiency (Fig. 12) are distributed across the observed parameter space, we find the column density and temperature ranges producing the most emission may not do so most efficiently. For example, although CO and $^{13}$CO are producing most of the emission within the moderate density and temperature regime, they have a peak in emission efficiency at almost an order of magnitude lower column density. In addition, we find that many of the lines that showed two peaks in the emission distribution have a single peak in the emission efficiency distribution (e.g. HCO$^+$, HCN, HNC, CS). Interestingly, several of the lines show comparable distributions for the emission and emission efficiency (N$_2$H$^+$, SO, CCH), highlighting that for these lines the gas properties producing the most emission are also doing so most efficiently.

## 5 DISCUSSION

In this section, we discuss the results of this work in the context of our current understanding of the origins of molecular line emission from the interstellar medium. We note that an outline of the uncertainties present within our observations and the caveats that these imply to our interpretations is presented at the end of this section.

### 5.1 Are 'dense gas' molecular line tracers preferentially tracing the 'dense' gas?

When studying the array of molecular line transitions observable within molecular clouds, the critical density of a given transition is typically used as a proxy of the density from which it is emitting. As such the different molecular line transitions can be used to probe the physical conditions of the various density regimes present throughout the cloud. Of particular interest are the properties of the 'dense gas' most closely associated with star formation. Much of our understanding of star formation is based on the study of these dense gas properties. For example, the star-forming potential of the molecular clouds is directly determined by the physical conditions (e.g. level of turbulence) of the dense gas, upon which the majority of the current analytic models for star formation are calibrated. Moreover, the mass determined from these dense gas tracers is used to derive empirical scaling relations to the level of star formation, which are widely used to calibrate simulations and estimate the star formation rate within extragalactic observations. The fundamental assumption typically adopted by molecular line studies is that the emission from any dense gas tracer, as inferred from its critical density, is (a) emitting from gas at that density (i.e. $n \sim n_{\rm crit}$), and (b) this gas is directly linked to the star formation process (e.g. Gao & Solomon 2004b). In this work, we address the first of these assumptions, and ask the question: are 'dense gas' molecular line tracers preferentially tracing the **'dense'** gas?

In Fig. 13, we present integrated intensity maps for a selection of the observed molecular lines, which have been ordered by increasing critical density (all integrated intensity maps are shown in Fig. 3).
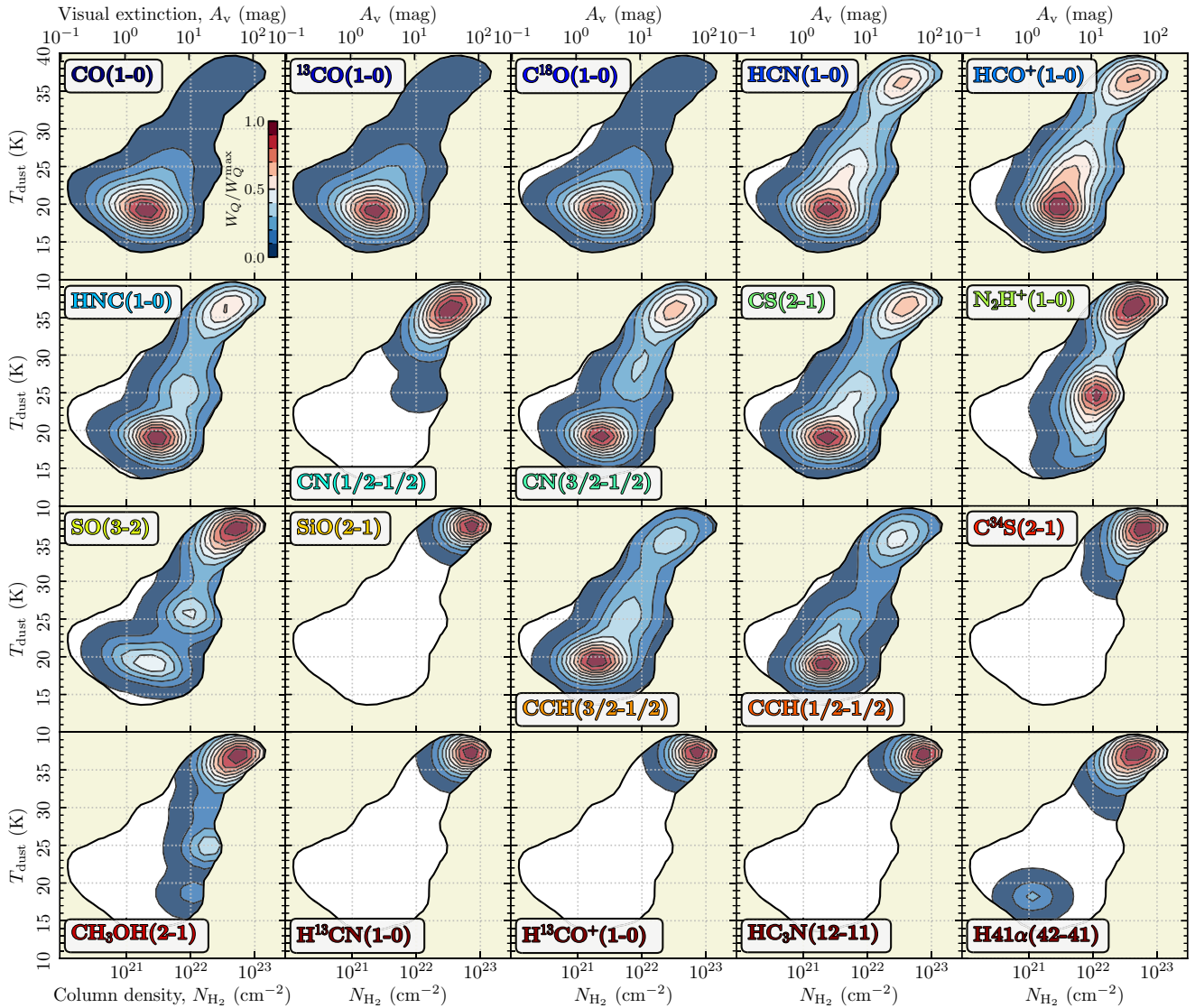
**Figure 11.** The distribution of integrated intensities as a function of the dust temperature and molecular hydrogen column density (extinction), where the integrated intensities have been binned along each axis to form a contour plot (cf. scatter plot shown in Fig. 4). The contours show the 1, 10, 20, 30 ... 80, and 90 per cent levels of the total integrated intensity for a given line across the significant bins within the parameter space ($5\sigma_{\rm bin}$; see equation 1). This plot highlights the properties of the region from which the majority of molecular line emission across the mapped region are originating.

This plot was made to be comparable to fig. 1 of Kauffmann et al. (2017), which presents the molecular line emission across the local Orion A star-forming region. Here, we can see broadly similar trends to what was found for Orion A. We find CO, $^{13}$CO, have extended emission across all the mapped region, whilst $^{17}$CO emits only towards the centre of W49A. Given that these molecular lines have comparable critical densities, this effect is likely due to the reduced abundance of the rarer isotopologues and insufficient integration time. Or, it could be that the rarer species are destroyed in the lower density material, reducing further their abundance (fractionation effects, e.g. Langer et al. 1984; Röllig & Ossenkopf 2013).

Moving beyond these CO lines in Fig. 13, we find little correspondence between the extent of the molecular line emission and the increase in critical density. For example, we see comparable spatial morphologies of the CCH, CN, and HCN lines, despite them having critical densities differing by almost two orders of magnitude. Moreover, we find that only $N_2H^+$ appears to also trace the column

density peak at the east of the mapped region (towards W49B) at comparable brightness to the peak towards the centre of the W49A; potentially highlighting this as a region of differing excitation or chemistry than W49. These results for all the observed lines are summarized in Fig. 6, from which we find that there is a significant scatter around the simple linear relation of the average integrated intensity across the mapped region and increasing critical density; several extreme outliers being $C^{18}O$ and $CH_3OH$. This scatter is increased when plotting the coverage across the mapped region (i.e. positions with an integrated intensity above five times the uncertainty) as a function of the critical density. Together, Figs 6 and 13 then clearly show that the brightness and coverage of a molecular line transition are not completely governed by the critical density; e.g. excitation effects or chemical abundance variations may also be important factors. This is in agreement with the results from the Orion molecular cloud (Kauffmann et al. 2017; Pety et al. 2017).
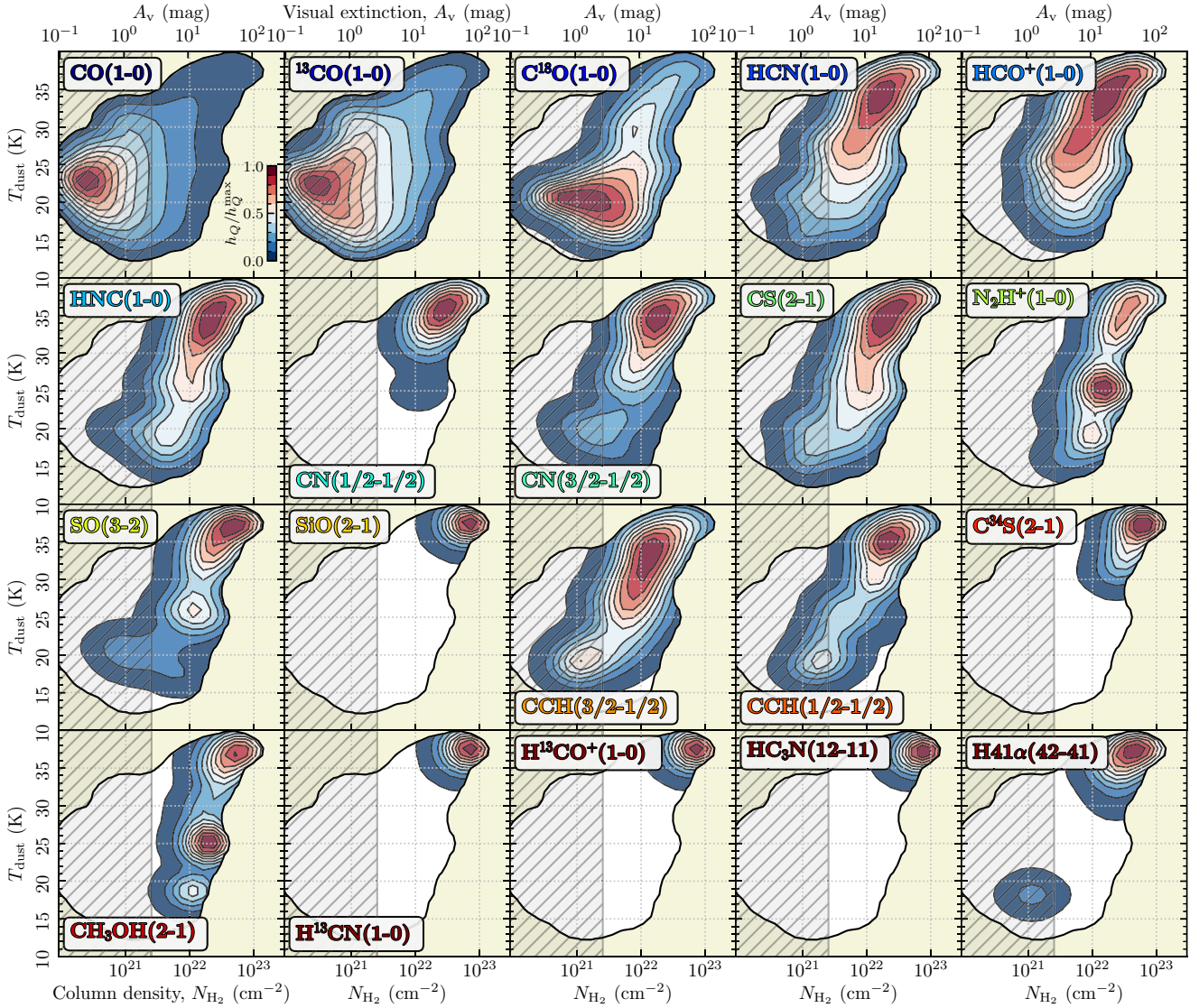
**Figure 12.** The distribution of emission efficiencies as a function of the dust temperature and molecular hydrogen column density (extinction), where the emission efficiencies have been binned along each axis to form a contour plot (cf. scatter plot shown in Fig. 4) The contours show the 1, 10, 20, 30 ... 80, and 90 per cent levels of the emission efficiency [$h_Q = W(Q)/N_{H_2}$] for a given line across the significant bins within the parameter space ($5\sigma_{bin}$; see equation 1). We highlight here values of the column density below $N_{H_2}^{thresh} = 2.5 \times 10^{21}$ cm$^{-2}$ ($A_v^{thresh} = 2.7$ mag) that suffer from increased (>30 per cent) uncertainties due to foreground and background contamination (see Section 3.1). Values of the $h_Q$ below this threshold value should be taken with caution. This plot highlights the properties from which the molecular lines are most efficiently emitting across the mapped region are originating.

In Section 3.3, we investigated the emissivity of the molecular lines. When doing so, we identified large deviations from a simple linear relation for the integrated intensity of a molecular line and the column density. To investigate this, we normalize the integrated intensity by the column density, which we defined as the emission efficiency ratio [$h_Q = W(Q)/N_{H_2}$; Kauffmann et al. 2017]. In Fig. 14, we present the emission efficiency ratio as a function of the column density for a selection of the observed molecular lines, which have plotted on a single axis for comparison (comparable to Fig. 8). This plot shows a variety of profiles for the different molecular lines with increasing column density. Also shown in the lower panel of Fig. 14 are the emission efficiency ratios determined within the Orion A star-forming region taken from Kauffmann et al. (2017, their fig. 2). We find a reasonable first-order agreement between W49 and Orion A. We see that $^{12}$CO has a maximum emission efficiency ratio at the lowest column density. C$^{18}$O, HCN, and N$_2$H$^+$ show

peak emission efficiency ratio values at increasingly higher column densities. There are, however, several second-order differences that should be mentioned. The first is that we do not observe a peaked emission efficiency ratio profile for $^{12}$CO within W49, as is seen within Orion (comparing blue curves in Fig. 14). Rather, we find a gradual decline of the $^{12}$CO emission efficiency ratio with increasing column density. The second is that for Orion the HCN emission efficiency ratio peaks at a lower column density than N$_2$H$^+$, whereas in W49 we find these are comparable, with N$_2$H$^+$ even potentially peaking at lower column density values. It is worth noting here that the spatial resolution of the W49 observations is ∼3 pc, whereas the beam of the Orion A observations is ∼0.05 pc. Therefore, it is encouraging that we are even broadly recovering the trends seen within Orion.

In Section 4.2, we built upon this analysis by assessing how the total intensity and emission efficiency ratio at a given column density
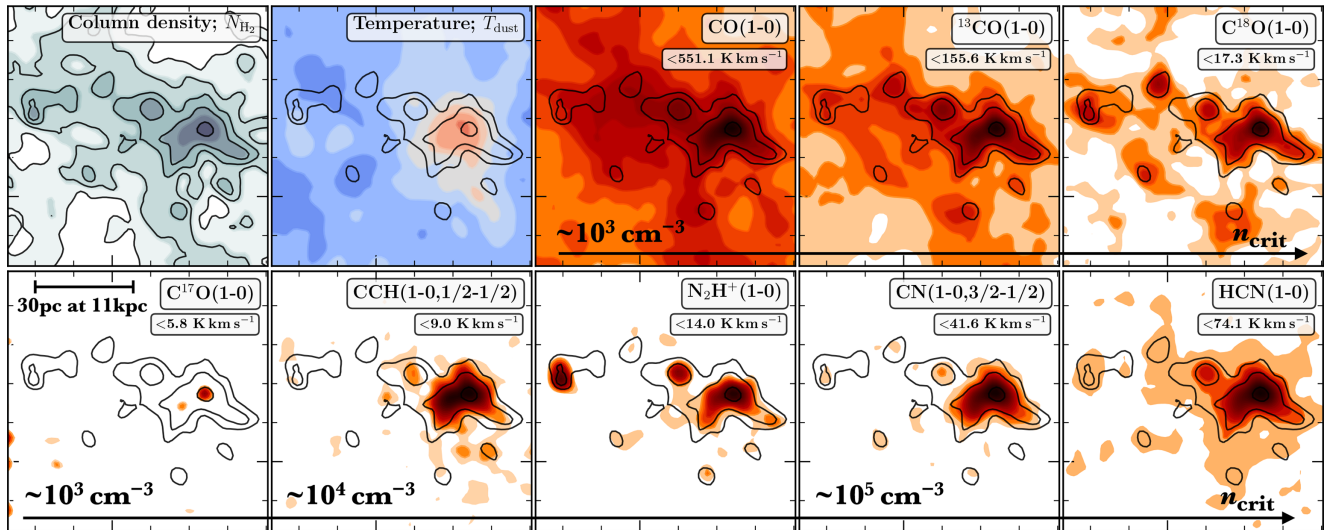
**Figure 13.** Maps of the integrated intensity for a selection of lines (see comparable figure for Orion A in Kauffmann et al. 2017). The upper left two panels show the column density and dust temperature determined from the *Herschel* observations. Overlaid on the upper left panel are contours $N_{H_2}$ in levels of 1 and $2.5 \times 10^{21}$ cm$^{-2}$, and in all panels levels of 5, 10, 50 $\times 10^{21}$ cm$^{-2}$. The filled contours for the molecular line emission are drawn at signal-to-noise ratios of 3, 5, 10, 30, 50, 100, 150, 200, and 300. All maps shown here have been smoothed to 60 arcsec ($\sim$3 pc at the source distance of 11 kpc), and the panels are ordered by increasing critical density. Labelled within each of the integrated intensity maps the peak value (see Table B1).

varies as a function of dust temperature. In Fig. 15, we present a simplified version of Figs 11 and 12. Here, we show the total intensity in the upper panel, and the emission efficiency ratio in the lower panel, as a function of column density and temperature for $^{12}$CO, C$^{18}$O, HCN, and N$_2$H$^+$ on a single axis to allow comparison. To highlight the highest intensity peaks and most efficiently emitting gas we only show contours above a 50 per cent level.

We first see from Fig. 15 that CO, C$^{18}$O, and HCN produce the majority of their total emission at moderate temperatures ($\sim$20 K), and moderate column densities ($\sim$10$^{21-22}$ cm$^{-2}$). This region of the parameter space contains the most number of pixels (compare to the Fig. 4), and hence shows that the majority of the emission from these lines is not strongly affected by smaller scale local intensity enhancements. The majority of N$_2$H$^+$ emission across the mapped region, on the other hand, comes from higher column densities and is split between moderate and high temperatures. These results may then have implications for understanding molecular line emission from unresolved measurements at or above the size scale of the mapped region ($\sim$100 pc at 11 kpc).

Comparing the total intensities to efficiency ratios in the lower panel of Fig. 15, however, we find very different distributions. We see that CO has an emissivity that peaks towards the lowest column densities, and at moderate dust temperatures. There are several well-studied mechanisms that could cause this effect. Optical depth effects are most likely the major cause of this for the observed range of column densities (10$^{20-23}$ cm$^{-2}$): CO for example, is thought to be optically thick in all but the lowest density gas within the Galaxy. An additional effect towards the high values of the column density would be CO depletion. In cold ($T_{dust} < 20$ K) and dense ($n_{H_2} > 10^4$ cm$^{-3}$) environments, CO molecules can 'freeze-out' on to the surfaces of dust grains forming icy mantles (e.g. Fontani et al. 2012; Giannetti et al. 2014), such that their observed gas-phase abundances are a factor of several smaller than expected in the case of no depletion (e.g. Kramer et al. 1999; Hernandez et al. 2011). We note, however, that we do not observe a strong dependence of the CO emission on the temperature at any given column density, and we would expect the freeze-out of CO to be strongly anticorrelated with increasing

temperature, which we do not observe. This could be due to the large spatial scale of the beam used to study W49 ($\sim$3 pc).

Secondly, Fig. 15 shows that C$^{18}$O has an emissivity that peaks at $N_{H_2} \sim 10^{21-22}$ cm$^{-2}$, which spans a temperature range of 15–25 K. As with the other CO isotopologues, above these densities, the C$^{18}$O emission saturates, which is likely affected by the optical depth. Interestingly, here we find that for a given column density the emission efficiency ratio appears to peak towards higher dust temperatures (e.g. see between $10^{21}$ and $2 \times 10^{21}$ cm$^{-2}$), which could tentatively point to the previously mentioned freeze-out of CO within the cold environments.

Thirdly, when analysing the emission efficiency ratios as a function of only the column density from Fig. 14, we concluded that both HCN and N$_2$H$^+$ peak at $N_{H_2} \sim 3 \times 10^{22}$ cm$^{-2}$; an order of magnitude above the C$^{18}$O peak. However, when we now compare with Fig. 15, we find that around this peak in column density, the distribution of HCN and N$_2$H$^+$ lines split into two discrete temperature peaks. We find that the N$_2$H$^+$ peaks towards the lowest observed temperatures at the column density within the observed parameter space ($\sim$20 K), whereas the HCN peaks towards moderate-to-high temperatures of around $\sim$30–25 K. This then highlights the usefulness of the dust temperature as an additional variable to further investigate the emissivity profiles within the regions even at a low spatial resolution ($\sim$3 pc for these W49 observations). It is not surprising that both HCN and N$_2$H$^+$ peak towards the densest gas as the abundance of N-bearing species is boosted within dense gas (e.g. Caselli et al. 1999). However, it is somewhat surprising that they peak so clearly at different temperatures. None the less, we propose that this could be explained by the following. The abundance of N$_2$H$^+$ is known to be particularly enhanced towards low temperatures, within regions where the CO molecules are frozen-out of the gas phase. This is due to CO being an effective destructor of the precursors in the formation pathway of N$_2$H$^+$; removing molecules such as NH$^+$ and H$_3^+$. On the other hand, Pety et al. (2017) showed that the molecules such as HCN (also HCO$^+$, CN, and HNC) are sensitive to the far-UV radiation produced from star formation, where star-forming
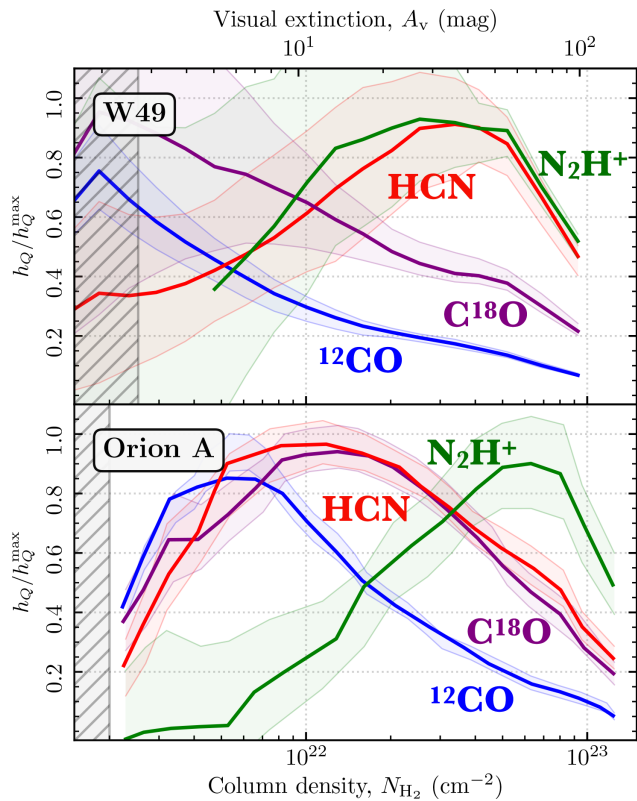
**Figure 14.** The ratio of the integrated intensity to the molecular hydrogen column density [or emission efficiency ratio; $h_Q = W(Q)/N_{H_2}$] as a function of column density for four the molecular lines (i.e. similar to Fig. 8). Shown in the upper panel are the data presented in this work for the W49 region. To allow a simple comparison, only bins that have a median value above $5\sigma_{bin}$ are displayed. The lower panel shows the emission efficiency ratios determined within the Orion A local star-forming region taken from Kauffmann et al. (2017, their fig. 2). We highlight by shading here column densities below the threshold limits for both sources, which are believed to suffer from foreground and background contamination (see Section 3.1). Values of the $h_Q$ below this threshold value should be taken with caution.

regions can be identified as having elevated dust temperatures. For example, the HCN (1−0) line is thought to be pumped by mid-infrared photons and/or the effects of X-ray Dominated Regions (e.g. Aalto et al. 2007), which can both be present within actively star-forming regions (see Fig. 1).

### 5.2 Focusing on the 'dense' gas within W49A

To better understand the line emission properties across the region we define a density model. To do so we focus on the central ∼40 pc (or 0.2°) of W49A,[8] where we are the most confident that the majority of the molecular line emission and column density is associated to the material at or close to the distance of the W49A star-forming region ($11.11^{+0.79}_{-0.69}$ kpc; Gwinn et al. 1992; Zhang et al. 2013). As the true three-dimensional source structure of this region is not well understood, and there is almost certainly a significant level of sub-structure present on scales much smaller than the ∼3 pc beam size used throughout this work, we choose to define the most basic

[8]We choose to define the central position of the W49A region as RA, Dec. (J2000) = $19^h10^m14^s$, $9°06°17''$, and take only the pixels within a box of 0.2° around this position for the analysis presented in this section.
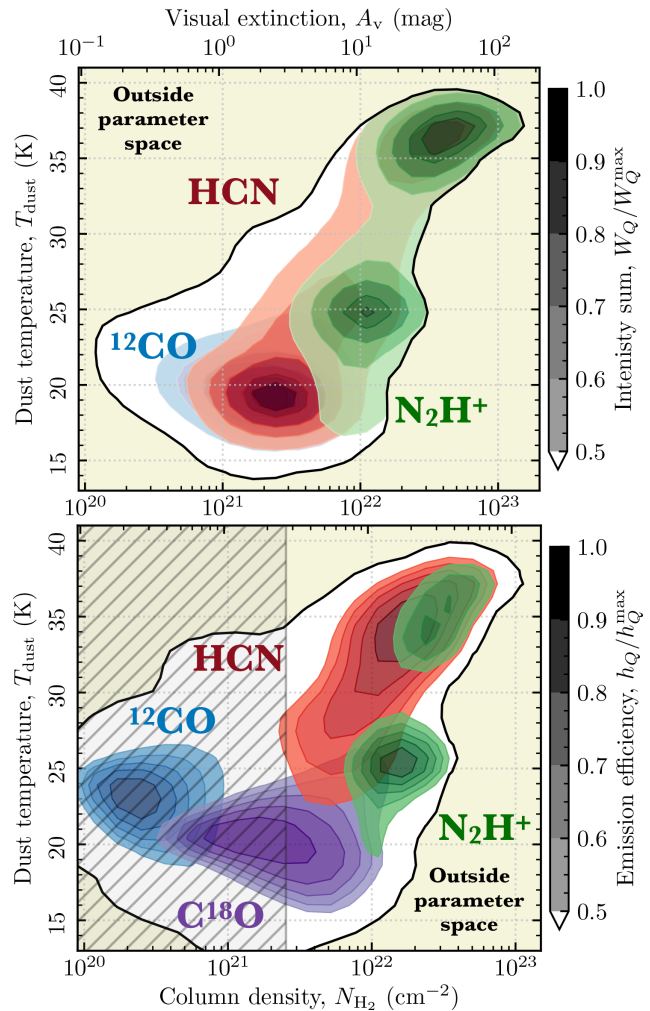
**Figure 15.** The total integrated intensity and emission efficiency [$h_Q = W(Q)/N_{H_2}$] as a function of column density and dust temperature for a selection of the molecular lines (see Figs 11 and 12). The contours show the 50, 60, 70 80, and 90 per cent levels of total intensity in the upper panel, and emission efficiency in the lower panel. The black contour shows the region of the parameter space probed by the observations presented in this work. In the lower panel, we highlight values of the column density below $N_{H_2}^{thresh} = 2.5 \times 10^{21}$ cm$^{-2}$ ($A_v^{thresh} = 2.7$ mag) that suffer from increased (>30 per cent) uncertainties due to foreground and background contamination (see Section 3.1). Values of the $h_Q$ below this threshold value should be taken with caution. This plot highlights the properties producing the most molecular line emission, and where the emission is being produced most efficiently.

possible density model. We propose that the number density using the spatial resolution of the *Herschel* observations can also be obtained by limiting the three-dimensional structure to a single spatial element along the line of sight: or in other words, by assuming that the measured column density at each position originates from a 20 arcsec × 20 arcsec × 20 arcsec cube (or ∼1 pc³ at ∼11 kpc). We determine an uncertainty on the model density by varying the cube size by (1) the linear size based on the uncertainty from the distance estimate, and (2) varying the depth of the box between 0.5 and 3 pc, which represent the average size of the sub-structures identified by Eden et al. (2018), and the size of the angular resolution of the molecular line observations (∼60 arcsec), respectively.

The lower panel of Fig. 16 shows the distribution of the number density as a function of the observed column density. We find
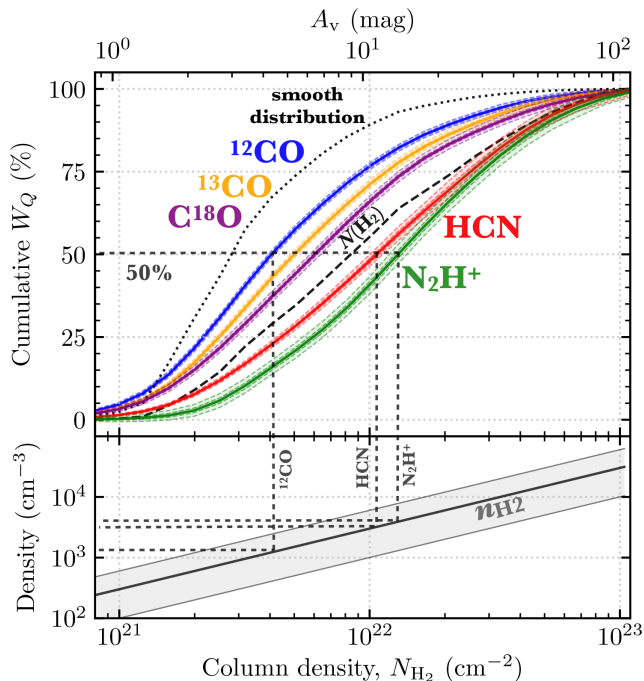
**Figure 16.** Cumulative fraction of the emission from each molecular line as a function of column density across the W49A region only. The lines and shaded regions within the upper panel have the same meaning as plotted for Fig. 9. The lower panel shows the model density as a function of the column density. The grey shaded region shows one standard deviation of the distribution of model densities for a given column density, whilst the solid line shows the median of this distribution. The $W_Q(N_{H_2} < N'_{H_2}) = 50$ per cent, $N^Q_{H_2,char}$, and $n^Q_{char}$ values for CO, HCN, and $N_2H^+$ are overlaid as black dashed lines.

minimum, median, and maximum values for the model density across the W49A region of $1.9 \times 10^2, 1.1 \times 10^3$, and $6.4 \times 10^4$ cm$^{-3}$, respectively. The upper panel of Fig. 16 shows the cumulative distributions for a selection of the molecular lines as a function of column density when limited to the W49A region. We find that this plot recovers many of the trends seen in the cumulative distribution function of the entire mapped region (cf. Fig. 9).

To then condense the cumulative distributions shown in Fig. 16 into a single value for each line, Kauffmann et al. (2017) define a characteristic column density, $N^Q_{H_2,char}$, as the column density that contains half the total line intensity; $W_Q(N_{H_2} < N'_{H_2}) = 50$ per cent. Highlighted in Fig. 16 is the $N^Q_{H_2,char}$ for CO, HCN, and $N_2H^+$, and the corresponding characteristic density: $n^Q_{char} = n(N^Q_{H_2,char})$. The characteristic column densities and densities for all lines are summarized within Table 4, which we can compare to the critical densities listed in Table 2. We find that the characteristic densities and critical densities for the low-density tracing CO lines are comparable ($n^Q_{char} \sim 10^3$ cm$^{-3}$). However, we find that the higher density tracing molecular lines, such as HCN and $N_2H^+$, have characteristic densities only marginally higher than the CO lines ($n^Q_{char} \sim 4 \times 10^3$ cm$^{-3}$), and hence significantly lower than their critical densities ($n_{crit} \sim 10^{5-6}$ cm$^{-3}$). A similar analysis was carried out for the Orion A molecular cloud by Kauffmann et al. (2017, see their fig. 3). Assuming a cylindrical density distribution, these authors determine a characteristic density for HCN of $n^Q_{char} \approx 900^{+1240}_{-550}$ cm$^{-3}$, which is comparable to within the uncertainty of the value calculated here ($3400 \pm 2800$ cm$^{-3}$).

**Table 4.** Properties determined from the cumulative distributions. Tabulated is the molecule, and its characteristic column density, $N^Q_{H_2,char}$, which is defined as the column density that contains half the total line intensity; $W_Q(N_{H_2} < N'_{H_2}) = 50$ per cent. The uncertainties show the range of $N^Q_{H_2,char}$ after adding a synthetic noise of $1\sigma_{W_Q}$ (see shaded region in Fig. 9). We provide the results of the analysis across the whole mapped region (i.e. using the cumulative distributions shown in Section 4.1), and when limited to W49A (i.e. using the cumulative distributions shown in Section 5.2). Also shown for W49A is the characteristic density, which we define as $n^Q_{char} = n(N^Q_{H_2,char})$ (see Section 5.2). This table has been ordered by increasing critical density (see Table 2). The full, machine-readable version of this table can be obtained from the supplementary online material.

| Molecule | $N^Q_{H_2,char}$ (whole map) ($\times 10^{21}$ cm$^{-2}$) | $N^Q_{H_2,char}$ (W49A) ($\times 10^{21}$ cm$^{-2}$) | $n^Q_{char}$ (W49A) ($\times 10^3$ cm$^{-3}$) |
|---|---|---|---|
| $C^{18}O(1-0)$ | $2.7 \pm 0.1$ | $6.7 \pm 0.2$ | $2.0 \pm 1.7$ |
| $^{13}CO(1-0)$ | $2.4 \pm 0.0$ | $5.3 \pm 0.2$ | $1.6 \pm 1.3$ |
| $C^{17}O(1-0)$ | $4.2 \pm 2.2$ | $9.0 \pm 4.0$ | $2.7 \pm 2.2$ |
| $CO(1-0)$ | $2.1 \pm 0.1$ | $4.5 \pm 0.1$ | $1.3 \pm 1.1$ |
| $N_2H^+(1-0)$ | $10.1 \pm 0.6$ | $14.2 \pm 0.8$ | $4.2 \pm 3.5$ |
| $HCN(1-0)$ | $5.0 \pm 0.1$ | $11.3 \pm 1.0$ | $3.4 \pm 2.8$ |
| ... | ... | ... | ... |

A potential explanation for the excess emission at characteristic densities below the critical density could be sub-thermal excitation (as previously mentioned for the extended map). Indeed, Evans (1989) highlighted that the subthermal emission from gas below the critical density can be significant for molecular lines within the sub-mm regime. Additionally, this could be partly a beam dilution effect. As the W49 region is known to contain significantly higher densities on smaller spatial scales that are not resolved by the observations presented in this work (e.g. Vastel et al. 2001; Eden et al. 2018). The line emission could then emit from these compact regions that are heavily beam diluted in the observations presented in this work. For lines such as HCN or $N_2H^+$, which have critical densities around three orders of magnitude higher than the observed characteristic density, we estimate that observations with a spatial resolution of around a factor of $\sqrt{10^3} = 30$ (i.e. 1 arcsec or 0.05 pc) would be required to investigate these beam dilution effects.

### 5.3 Column density and mass conversions factors

There are two commonly adopted conversion factors within the literature that are used to estimate the mass of the material within a source from molecular line emission. Such conversions are useful for distant objects – such as extragalactic sources – where other methods of determining the gas mass of a system are not possible (e.g. dust extinction/continuum measurements). In this section, we investigate how the typically assumed canonical conversion factors adopted within the literature hold within the W49 region.

#### 5.3.1 Column density conversion; X-factor ($X_Q$)

The first of these is typically referred to as the 'X-factor', and is used to convert the integrated intensity of a line to the molecular hydrogen column density. This can be defined as,

$$N_{H_2} = X_Q W_Q,  \qquad (3)$$

where $X_Q$ is the X-factor of a given molecular line transition in units of cm$^{-2}$ (K km s$^{-1}$)$^{-1}$. This conversion is typically used to estimate the total molecular gas mass from CO emission, which has

an assumed canonical conversion value of $X_{CO(1-0)} = 2 \times 10^{20}$ cm$^{-2}$ (K km s$^{-1}$)$^{-1}$ (Bolatto, Wolfire & Leroy 2013). Variations of the $X_{CO(1-0)}$ are, however, widely seen in both observations (see table 1 of Bolatto et al. 2013) and simulations (e.g. Shetty et al. 2011a,b; Clark & Glover 2015; Gong, Ostriker & Kim 2018).

We calculate the $X_{CO(1-0)}$ factor using the mean molecular hydrogen column density across the mapped region, and the mean integrated intensity. Treating the whole mapped region as a single pixel, for this analysis we do not apply any uncertainty threshold clipping to the molecular line data set (i.e. $X_{CO} = N_{H_2}^{mean}/W_{CO}^{mean}$; see Table E2 for the corresponding calculation for other lines). When doing so we find a mean $X_{CO(1-0)} = 0.3 \times 10^{20}$ cm$^{-2}$ (K km s$^{-1}$)$^{-1}$ across the mapped region (error on the mean shown as uncertainty), which is significantly below the value found within the literature (Bolatto et al. 2013). Pety et al. (2017) found a similarly low value of $X_{CO(1-0)}$ towards the Orion B star-forming region. These authors attributed this difference to the UV illumination of the gas by massive stars, where only diffuse gas or the UV shielded dense gas have $X_{CO(1-0)}$ close to the typically assumed value. Manual inspection of $X_{CO(1-0)}$ on a pixel-by-pixel basis show elevated value towards the highest column density and temperature positions; towards the W49A region where embedded high-mass stars are known to reside. It is then difficult to attribute the $X_{CO(1-0)}$ discrepancy observed across W49 to UV illumination.

### 5.3.2 Mass conversions; alpha-factor ($\alpha_Q$)

The second conversion factor we investigate here is between the gas mass and the luminosity of the molecular line, $\alpha_Q$. This can be defined as,

$$M = \alpha_Q L_Q, \qquad (4)$$

where $\alpha_Q$ is in units of M$_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$. This conversion is typically used for molecular lines that are thought to trace only the dense gas, and hence to calculate the mass of dense gas within the region (e.g. using HCN; Gao & Solomon 2004a). Gao & Solomon (2004b) outlined that the conversion for a virialized parcel of gas with a volume-averaged molecular hydrogen number density of $n_{H2} \sim 3 \times 10^4$, and optically thick HCN (1−0) emission with a brightness temperature of $T_{MB} = 35$ K, yields the value $\alpha_{HCN(1-0)} = 10$M$_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$. This calculation contains several fundamental assumptions for the HCN emission, which are not confirmed by the results presented in this work. For example, towards the centre of W49A we find peak main beam brightness temperatures for HCN of $T_{MB} \sim 4$ K, and characteristic number densities of $\sim 10^{3-4}$ cm$^{-3}$ (Section 5.2), which would then give a significantly higher than standard value of $\alpha_{HCN(1-0)} \sim 15$–50 M$_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$ [see Appendix E for the full derivation of $\alpha_{HCN(1-0)}$]. Indeed, higher values of the $\alpha_{HCN(1-0)}$ have been calculated for larger regions containing less efficiently emitting HCN (1-0). Kauffmann et al. (2017), for example, found $\alpha_{HCN(1-0)} = 20$ M$_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$ for the Orion A molecular cloud (also see e.g. Wu et al. 2010). Most recently, Evans et al. (2020) studied a sample of $\sim 20$ pc size molecular clumps within the Milky Way, and found $\alpha_{HCN(1-0)} = 20$ M$_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$ within the gas above a visual extinction of $A_v > 8$ mag, and $\alpha_{HCN(1-0)} = 6$ M$_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$ when integrating the HCN emission across the whole clumps.

We calculate $\alpha_Q$ in both the total luminosity of the bulk molecular gas, and limiting the analysis to only the luminosity from the 'dense' molecular gas following the $\alpha_Q$ definitions of Evans et al. (2020). To do so, we firstly convert the integrated intensity maps to luminosity

assuming a single distance of 11.11 kpc (Zhang et al. 2013). We then sum the luminosity across the whole mapped region without imposing an extinction threshold ($L_Q^{sum}$). In addition, we determine the luminosity within a $A_v > 8$ mag contour ($L_{Q,A_v>8\,mag}^{sum}$), which is typically referred to as a limit for dense gas (e.g. see Kauffmann et al. 2017). We impose the same extinction threshold for the molecular gas ($M_{A_v>8\,mag}^{sum}$). The mass conversion is then calculated as the mass of $A_v > 8$ mag gas divided by the total luminosity of the given line: $\alpha_Q = M_{A_v>8\,mag}^{sum}/L_Q^{sum}$. This definition is most appropriate for determining the dense gas mass from unresolved (extra-galactic) observations that measure the total HCN luminosity including emission below and $A_v$ of 8 mag. Additionally, we calculate $\alpha_Q$ using both the mass and luminosity above the $A_v > 8$ mag threshold: $\alpha_Q(A_v > 8\,mag) = M_{A_v>8\,mag}^{sum}/L_{Q,A_v>8\,mag}^{sum}$. This second method is appropriate for resolved (galactic) observations, where the HCN emission can be measured exclusively within the dense gas.

We find $\alpha_{HCN(1-0)} = 30 \pm 1$ M$_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$, and $\alpha_{HCN(1-0)}(A_v > 8\,mag) = 73 \pm 4$ M$_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$ (see Table E1 for additional conversion factors such as $\alpha_{CO(1-0)}$).[9] These $\alpha_{HCN(1-0)}$ values are both significantly higher than the typically adopted value [$\sim 10-20$ M$_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$], and more similar to the value determined analytically based on the observed source properties [15–50 M$_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$; Appendix E]. Fig. 17 shows a comparison of the measured $\alpha_{HCN(1-0)}$ to values available within the literature. The points have been coloured by extragalactic (blue; García-Burillo et al. 2012), high-mass (red; Evans et al. 2020; Nguyen-Luong et al. 2020), and local star-forming regions (green; Kauffmann et al. 2017; Shimajiri et al. 2017). We see that there is a large scatter in $\alpha_{HCN(1-0)}$ ranging $\sim 0.3$–300 M$_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$ across all sources. Within this scatter, we find the high-mass and extragalactic sources appear to be systematically lower compared to local star-forming regions. We propose that the lower $\alpha_{HCN(1-0)}$ within higher mass regions could be due to the contribution of HCN emission from lower density gas present within subsequently large maps (i.e. towards more distance higher mass star-forming regions). Also highlighted on Fig. 17 is the map size used to determine each $\alpha_{HCN(1-0)}$ value, and indeed we see that higher mass star-forming regions are studies across two to four orders of magnitude large areas compared to local star-forming regions. An alternative explanation, as suggested by Shimajiri et al. (2017), is that the higher levels of star formation and, hence, the radiation field could lower $\alpha_{HCN(1-0)}$.

The analysis presented here highlights that significant caution should be exercised when adopting mass conversion factors from the literature. For example, if we take the mapped region as a typical bright 100 pc × 100 pc pixel within an extragalactic observation, using the typically adopted dense gas conversion factor would cause an underestimation of the dense gas mass by a factor of up to $\sim 3$ with respect to the extragalactic conversion factor $\alpha_{HCN(1-0)}$ determined in this work. This would cause an underestimation of the dense gas fraction ($f_{dense} = M_{dense}/M_{tot}$), and in turn, any property that contains $f_{dense}$ (e.g. the dense gas star formation efficiency: SFE$_{dense} = f_{dense}M_{tot}$/SFR). This would then cause significant scatter in any comparison to the dense gas star formation efficiency relations (e.g. Lada et al. 2010).

---

[9]Table E2 provides $\alpha_Q$ values that have been determined without imposing extinction thresholds on either the mass or luminosity measurements.
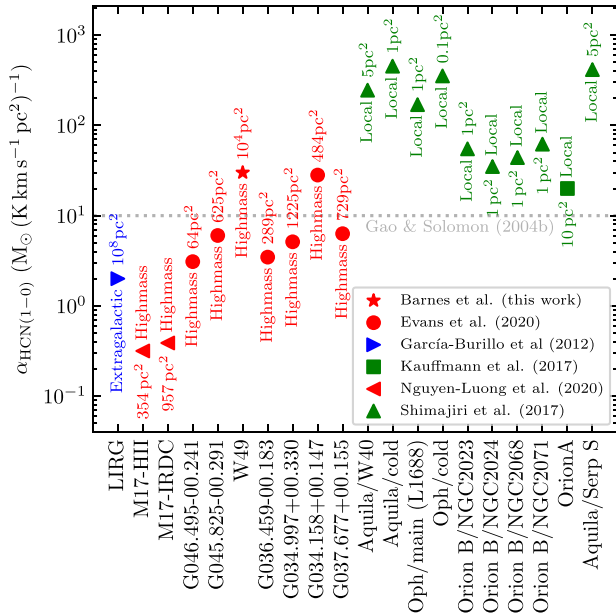
**Figure 17.** A comparison of $\alpha_{HCN(1-0)}$ relevant for unresolved star-forming regions measured across various sources taken from the literature (García-Burillo et al. 2012; Kauffmann et al. 2017; Shimajiri et al. 2017; Evans et al. 2020; Nguyen-Luong et al. 2020). The points have been coloured by the type of source studied by each work, which have been classified as extragalactic (blue), high-mass (red), and local star-forming regions (green). The horizontal dashed line shows the typically adopted value determined of $\alpha_{HCN(1-0)} = 10 M_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$ (Gao & Solomon 2004b). Note that $\alpha_{HCN(1-0)}$ has been calculated for all sources using the mass of gas above an extinction threshold of $A_v \gtrsim 8$ mag, divided by the total HCN luminosity across the region ($\alpha_Q(all) = M^{sum}_{A_v > 8\,mag}/L^{sum}_Q$). For each point, we label the size of the map used to determine both $M^{sum}_{A_v > 8\,mag}$ and $L^{sum}_Q$. This figure highlights that the contribution of HCN emission from lower density gas present within large maps (i.e. towards more distance higher mass star forming regions) could be a potential cause for lower than typically assumed $\alpha_{HCN(1-0)}$ values. Additionally, the higher levels of star formation and, hence, radiation field have also been suggested to lower $\alpha_{HCN(1-0)}$ (see Shimajiri et al. 2017).

## 5.4 Detection limits for extragalactic studies

In this section, we assume that the 100 pc × 100 pc region of W49 studied here is representative of most massive star-forming regions within the Galaxy. We then pose a question: if this region were a single pixel in an extragalactic observation, would we be able to detect the molecular lines based on the brightnesses observed in this work? Pety et al. (2017) performed a similar experiment for the Orion B star-forming region, yet this had a very limited spatial coverage of 5 pc × 5 pc, which would correspond to an angular resolution at extragalactic distances that is currently difficult to obtain with even the largest mm-telescopes (e.g. ALMA; 1 pc at 10 Mpc is ~0.02 arcsec). The observations presented here cover a much larger spatial area that corresponds to a projected angular size of ~2 arcsec at 10 Mpc, which is becoming much more routinely observed over large areas of extragalactic systems (e.g. Schinnerer et al. 2013; Leroy et al. 2016).

We estimate that with 10 h of Atacama Large Millimeter Array observing time, we could reach a $3\sigma_{W_Q} \sim 0.3$ K km s$^{-1}$ line sensitivity within a ~2 arcsec beam (100 pc at 10 Mpc) across

a 200 arcsec × 200 arcsec mosaic (or 10 kpc at 10 Mpc).[10] We compare this sensitivity to the mean integrated intensities across the mapped region listed in Table 3. We find that the CO ($W_Q^{mean} \sim 80$ K km s$^{-1}$) and $^{13}$CO (~10 K km s$^{-1}$) molecular lines are significantly ($>100\sigma_{W_Q}$) above this value, and therefore would be easily detectable within the proposed ALMA observations. Additionally, lines such as HCN (1.4 K km s$^{-1}$), C$^{18}$O (1.0 K km s$^{-1}$), and HCO$^+$ (0.9 K km s$^{-1}$) are above this sensitivity estimate and should, therefore, also be moderately ($10\sigma_{W_Q}$) detectable. To significantly detect e.g. N$_2$H$^+$ (0.17 K km s$^{-1}$), however, would require an increase in sensitivity of a factor of around two, which translates to a factor of around four more integration time (~46 hours).[11] It is worth noting that these estimates have been determined using the W49 region, which is one of the brightest regions of molecular line emission within the Milky Way. Therefore, these estimate may not be representative of the detection expected across the whole disc for the proposed set of extragalactic observations.

In this section, we have discussed the potential to detect weak molecular lines (e.g. N$_2$H$^+$) across the complete discs of nearby spiral galaxies thought to be comparable in their star formation properties to the Milky Way. However, it is worth highlighting that these lines have already been detected towards extragalactic sources using deep single-pointing or small mosaic observations (e.g. Sage & Ziurys 1995). However, these detections are typically towards galactic centres, or starburst systems that are known to have extended and elevated e.g. HCN or N$_2$H$^+$ emission relative to typical star-forming regions in disc galaxies (e.g. Meier & Turner 2005; Sakamoto et al. 2010). For example, Ginard et al. (2015) studied the central ~kpc of the starburst galaxy M82 at scales of ~50–100 pc, and found typical N$_2$H$^+$ (1-0) integrated intensities in the range of ~5–10 K km s$^{-1}$ (see their fig. 2). This is significantly higher than the mean N$_2$H$^+$ (1-0) integrated intensity measured across the whole mapped region of W49 (see Table 3). The integrated intensity range of ~5–10 K km s$^{-1}$ is limited to the central ~10 pc of W49A (see Fig. 3 and Table B1), which may be indicative of the 'mini-starburst' nature of W49A (e.g. Peng et al. 2010; Stock et al. 2014).

## 5.5 Caveats

At several points throughout the discussion, we mention the possible systematic uncertainties in the observations that may have affected our results. In this section, we summarize the most potentially significant of these caveats.

*Limited spatial resolution*: The observations presented throughout this work have been smoothed to a spatial resolution of ~3 pc at the assumed source distance (11 kpc; Zhang et al. 2013). Higher resolution observations have shown that much sub-structure exists within W49A on scales much smaller than this beam size, which would then not be resolved with these observations (Galván-Madrid et al. 2013). For example, using James Clerk Maxwell Telescope and

---

[10]This time estimate has been performed using the ALMA Cycle 8 observing tool, and accounts for only the (C43-3 configuration) 12 m array observations. For these estimates, we choose the rest frequency of the HCN(1–0) transition, a representative bandwidth of 282 kHz (or 0.96 km s$^{-1}$), and Nyquist sampling of 39 12 m array pointings that create a 200 arcsec × 200 arcsec mosaic. The quoted time includes the on-source time to reach $3\sigma_{W_Q} \sim 0.3$ K km s$^{-1}$ of around 7 h, and calibration and overhead time of around 4.12 h.

[11]The on-source only 12 m array observing time required to reach $3\sigma_{W_Q} \sim 0.15$ K km s$^{-1}$ would be 27.6 h.

higher resolution *Herschel* observations, Eden et al. (2018) identified ∼50 compact sources within the central ∼60 pc of W49A that have effective radii ranging from 10 to 20 arcsec (∼0.2 to 0.3 pc). These cores have masses of $>10^3$ M$_\odot$, which when assuming a spherical geometry approximately have column densities of $>10^{23}$ cm$^{-2}$. Column density values of $10^{23}$ cm$^{-2}$ are only observed towards the centre of W49A within the ∼60 arcsec smoothed *Herschel* maps (see Fig. 1). The dense region identified by Eden et al. (2018) would then be beam diluted within our smoothed maps, along with any molecular line emission originating from these regions. It is plausible that this would cause molecular line emission predominately originating from isolated compact, high (column) dense regions to appear to originate from low column densities when diluted. This is still, however, speculative and it is difficult to determine exactly how this would affect the results presented in this work.

We should note, however, that it is promising that the studies of the Orion star-forming regions find similar results to this work; e.g. emission efficiency ratio as a function of column density (section 8; Kauffmann et al. 2017; Pety et al. 2017). We previously mentioned that it is surprising that we even recover these trends, given that the spatial resolution of the Orion studies is ∼0.1 pc; an order of magnitude higher than this work (∼3 pc). This might then highlight that limited spatial resolution is not such an issue, and it will be interesting to investigate if this holds for other distant sources within the LEGO survey.

*Large spatial coverage*: One of the fundamental improvements of this work, and the LEGO survey as a whole, over previous attempts to understand line emissivities of weak molecular line transitions is the large spatial coverage at high sensitivity. In the case of W49A, we have a total mapped area of ∼100 pc × 100 pc at the source distance (11 kpc; Zhang et al. 2013). This has allowed us to analyse the molecular line emission across a very broad range of column densities. In this first work, we have considered the mapped region as a whole, and have not considered the individual environments contained within the map. These environments could, for example, differ considerably in their chemistry, and therefore have very different abundances of the various molecular species. Given the large distance to W49A, it is almost certain that some emission within the mapped area will not be physically associated and have very different chemical properties. For example, the W49B supernova remnant is within the mapped region (see Fig. 1). However, it is not currently clear if W49B is in any way linked to W49A; with some surveys placing them up to several kilo parsecs away from each other (Radhakrishnan et al. 1972; Chen et al. 2013; Zhu, Tian & Zuo 2014). One may assume that the physical and chemical properties of any molecular clouds interacting with this supernova could be different from those within actively star-forming region W49A.

*Multiple sources along line of sight*: This final caveat is somewhat linked to both the previous two, and concerns the fact that we do not separate the distinct sources along each line of sight within our map. We proposed that this may cause scatter in our results via two plausible scenarios: (a) if two (or more) dense clouds are present along with a single line of sight, (b) if one dense cloud is present with a large amount of foreground material. The first of these scenarios is most likely the case towards the centre of the mapped region for the majority of molecular lines. W49 is known to be a very kinematically complex region (Galván-Madrid et al. 2010), and indeed manual inspection of the spectra show that multiple velocity components are present within this region. Away from the centre of W49A, the CO lines suffer the widespread overlap of velocity components at ∼10, ∼40, and ∼60 km s$^{-1}$, which are all extended across the mapped region. The remaining lines do not suffer from such large-scale spatial

overlap; only peaking towards W49A at ∼10 km s$^{-1}$ (albeit with several distinct components around this velocity), and then to the east of the mapped region at ∼60 km s$^{-1}$. An in-depth investigation of this velocity structure will be presented in future work. The second of the above-mentioned scenarios would cause an issue if there was a single dense cloud and a significant amount of diffuse foreground emission. This foreground material would contribute to the total column density along the line-of-sight, and to the more easily excited lines (e.g. CO). This material may not, however, emit strongly from e.g. HCN, which would cause an over (under) estimation in $\alpha_{HCN}$ ($h_{HCN}$) value.

## 6 CONCLUSIONS

Molecular line emission has long been used to probe the properties of molecular clouds within the Milky Way, and now with the advent of the current generation of (sub)mm-telescopes, they are also quickly becoming major tools for the study of extragalactic systems. It is typically assumed that emission from carbon monoxide (CO) and its isotopologues allows us to probe the bulk gas properties in galaxies, whilst molecular species such as HCN and HCO$^+$ (and N$_2$H$^+$ in future) potentially allow us to selectively characterize the densest gas. This is based on fundamental assumptions of the abundance and critical densities of these molecules across the density regimes within molecular clouds. It has, however, become clear such assumptions may not hold for all environments within individual, resolved star-forming regions (e.g. Kauffmann et al. 2017; Pety et al. 2017; Evans et al. 2020). For example, the so-called 'dense gas tracers' are assumed to emit at or above their critical densities (i.e. density when the collisional de-excitation rate is equal to the rate of spontaneous emission of photons for a molecular transition), however, additional mechanisms could play a major role in the total luminosity of the molecular line measured over a large region (e.g. Evans 1999; Shirley 2015; Goldsmith & Kauffmann 2017). Before we can exploit extragalactic line emission data to its full potential, we first need to develop a detailed understanding of these – and many other – emission lines in Milky Way molecular clouds. With this in mind, the 'Line Emission to assess Galaxy Observations' project (LEGO) aims at developing a first comprehensive picture of how the 3 mm band emission lines in Milky Way molecular clouds depend on factors like cloud density, star formation feedback, galactic environment, and metallicity.

In this paper, we present a study of emissivity of a selection of commonly observed molecular lines within the 3 mm band across a large mapped region containing the W49A massive star-forming source (∼100 pc × 100 pc at 11 kpc). The main results are summarized below.

(i) We detect bright, extended emission from the CO and $^{13}$CO (1−0) lines, which have significant emission – i.e. pixels with integrated intensities higher than three times their associated uncertainty – across the entire mapped region (Section 4.1). The C$^{18}$O (1−0) is the next most extended line, which has significant emission over half the mapped region (spatial coverage of ∼35 per cent). We find that the higher density tracing lines such as HCN, HCO$^+$ (1−0) cover just under a third of the mapped region (spatial coverage of ∼30 per cent), and, therefore, are fairly extended yet much less than CO and $^{13}$CO (1−0). Lastly, the lines such as N$_2$H$^+$ (1−0) are the most spatially compact (spatial coverage of < 10 per cent). These results show the critical density of the observed molecular lines cannot be used as trivial predictors for how extended the emission is across a star-forming region (Fig. 6). Such a result is not surprising

and could be easily explained by the lower abundance and, therefore, optical depth, of rarer molecules. In other words, it would not be fair to suggest that molecular line X should be more extended or brighter than molecular line Y across a given parcel of molecular gas solely because of the lower critical density of molecular line Y. Rather, the relative abundances of the molecules X and Y should be taken into account. As if the molecular Y is significantly more abundant than molecule X, molecular line Y may be brighter and more extended than molecular line X regardless of its lower critical density.

(ii) It is desirable to distil the emission characteristics of each molecular line into a single number. To do so, following the analysis of Kauffmann et al. (2017), we define the characteristic column density at which half the line intensity is observed [$N^Q_{H_2, char} = N_{H_2}(W_Q = 50$ per cent); Section 4.1]. We find that the dense gas tracers (e.g. HCN and $N_2H^+$) have higher $N^Q_{H_2, char}$ relative to the CO lines. We also outline a simple number density model for the central region of W49 (Section 5.2), which we then use to define the characteristic density for each line [$n^Q_{char} = n(N^Q_{H_2, char})$]. We find that the CO lines have comparable characteristic densities and critical densities ($n^Q_{char} \sim 10^3$ cm$^{-3}$). However, the higher density tracing molecular lines, such as HCN and $N_2H^+$, have characteristic densities only marginally higher than the CO lines ($n^Q_{char} \sim 4 \times 10^3$ cm$^{-3}$), and hence significantly lower than their critical densities ($n_{crit} \sim 10^{5-6}$ cm$^{-3}$). A potential explanation for the excess emission at characteristic densities below the critical density could be sub-thermal excitation, which can be significant for molecular lines within the sub-mm regime (e.g. Evans 1989). This confirms that the critical density of a line cannot be used as a probe of the physical volume density in an absolute sense. That said, the use of emission from dense gas tracers, such as HCN and $N_2H^+$, relative to CO emission may still be a useful tool to probe the proportion of relatively denser gas. Future works from the LEGO survey will aim at expanding the dynamic range in $n^Q_{char}$ between CO and the dense gas tracers.

(iii) The ratio of the integrated intensity over the column density can be thought as a proxy for the line emissivity, or the efficiency of an emitting transition per $H_2$ molecule (also previously called the emission efficiency). One may assume that the emissivity is a function of number density, and, therefore, a function of column density, which would be the case if the molecular abundance or excitation is directly linked to the density (Kauffmann et al. 2017). We find, however, significant variations in the emissivity of the various molecular line transitions (Section 3.3). The CO and $^{13}$CO transitions show an overall decrease with increasing column density (see Fig. 8). The C$^{18}$O transition shows a peak at $\sim 10^{21}$ cm$^{-2}$, and decreases towards higher column density values. We find that the remaining lines of HCN, HCO$^+$, HNC, CN, CS, CH$_3$OH, and $N_2H^+$ have similar peaked profiles, albeit towards higher values of the column density (few $10^{22}$ cm$^{-2}$). These trends are broadly in agreement with the results found towards the Orion star-forming regions (e.g. Kauffmann et al. 2017; Pety et al. 2017), which is somewhat surprising given the observations presented here have over an order of magnitude larger spatial resolution ($\sim 2$–3 pc at the source distance of 11 kpc).

(iv) Direct comparison to the Orion A results (Kauffmann et al. 2017), however, highlight several relative differences between the emission efficiency profiles between the two regions (see Fig. 14). For example, within Orion the HCN emission efficiency peaks at a lower column density than $N_2H^+$, whereas in W49 we find these are comparable, with $N_2H^+$ even potentially peaking at lower column density values. To further investigate this, we assess how the emission efficiency at a given column density varies as a function of dust temperature (Section 4.2). We find that around the single column

density emission efficiency peak, the distribution for HCN and $N_2H^+$ lines split into two discrete emission efficiency temperature peaks. We find that the $N_2H^+$ peaks towards the lowest observed temperatures at the highest column density within the observed parameter space ($\sim 20$ K), whereas the HCN peaks towards moderate-to-high temperatures of around $\sim 30$–25 K. We believe this could be caused by the chemical formation pathways of these molecules that favour different temperature regimes, or the enhancement in the emission due to transition specific excitation mechanisms [e.g. the mid-IR pumping of HCN (1–0)].

(v) We also investigate several commonly used conversion factors between molecular line emission and molecular gas mass and column density across the W49 region (Section 5.3). First, we focus on the conversion factor between the integrated intensity of CO and column density ($N_{H_2} = X_Q W_Q$; Section 5.3). We calculate an average value across the mapped region of

$$X_{CO(1-0)} = 0.297 \pm 0.006 \times 10^{20} \text{ cm}^{-2} \text{ (K km s}^{-1})^{-1},$$

which is significantly lower than the canonical value within the literature [$\sim 2 \times 10^{20}$ cm$^{-2}$ (K km s$^{-1}$)$^{-1}$; Bolatto et al. 2013]. Secondly, we investigated the conversion between the total luminosity of HCN and gas mass above a visual extinction of $A_v > 8$ mag within the region (i.e. $\alpha_Q = M^{sum}_{A_v > 8 mag}/L^{sum}_Q$). We find

$$\alpha_{HCN(1-0)} = 30 \pm 1 \, M_\odot \text{ (K km s}^{-1} \text{ pc}^2)^{-1},$$

which is significantly higher than the typically adopted value within the literature [10 $M_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$; Gao & Solomon 2004a,b]. If we assume that the 100 pc $\times$ 100 pc region of W49 studied here is a typical massive star-forming complex within the Milky Way, or within an extragalactic system, these results have serious implications for our understanding of star formation. For example, if we take our mapped region as a single 100 pc $\times$ 100 pc pixel in an extragalactic survey, and use the typically assumed conversion factors as opposed to the conversion factors presented in this work. We would overestimate the total molecular column density and molecular gas mass by a factor of seven, and underestimate the dense gas mass by a factor of three. This would then cause over an order of magnitude underestimation of a dense gas mass fraction ($f_{dense} = M_{dense}/M_{tot}$), and in turn, cause a large systematic offset in any property that contains $f_{dense}$ (e.g. the dense gas mass star formation efficiency: $SFE_{dense} = f_{dense}M_{tot}/SFR$). It is worth noting that here we take the $\alpha_{HCN(1-0)}$ measured over the mapped region as an example single pixel in an extragalactic observation, and, in reality, the statistics achieved with large maps of extragalactic systems *may* cause deviations in $\alpha_{HCN(1-0)}$ to be averaged-out over various mass and evolutionary stage star-forming regions. None the less, this example should highlight the need for further resolved measurements of mass conversion factors to precisely measure such properties as the dense gas star formation efficiency, which are needed to accurately constrain current star formation theories.

## DATA AVAILABILITY

The full machine-readable format of several tables within this article are available in the online supplementary material. The data underlying this article will be shared on reasonable request to the corresponding author.

## REFERENCES

Aalto S., Spaans M., Wiedner M. C., Hüttemeister S., 2007, A&A, 464, 193

Anantharamaiah K. R., 1985, JA&A, 6, 177

Barnes A. T., Longmore S. N., Battersby C., Bally J., Kruijssen J. M. D., Henshaw J. D., Walker D. L., 2017, MNRAS, 469, 2263

Bigiel F. et al., 2016, ApJ, 822, L26

Boger G. I., Sternberg A., 2005, ApJ, 632, 302

Bohlin R. C., Savage B. D., Drake J. F., 1978, ApJ, 224, 132

Bolatto A. D., Wolfire M., Leroy A. K., 2013, ARA&A, 51, 207

Carey S. J. et al., 2009, PASP, 121, 76

Carter M. et al., 2012, A&A, 538, A89

Caselli P., Walmsley C. M., Tafalla M., Dore L., Myers P. C., 1999, ApJ, 523, L165

Chen Y., Jiang B., Zhou P., Su Y., Zhou X., Li H., Zhang X., 2013, in Ray A., McCray R. A., eds, Proc. IAU Symp. 296, Supernova Environmental Impacts. Cambridge Univ. Press, Cambridge, p. 170

Churchwell E. et al., 2009, PASP, 121, 213

Clark P. C., Glover S. C. O., 2015, MNRAS, 452, 2057

Dame T. M., 2011, preprint (arXiv:1101.1499)

De Pree C. G., Mehringer D. M., Goss W. M., 1997, ApJ, 482, 307

De Pree C. G., Wilner D. J., Goss W. M., Welch W. J., McGrath E., 2000, ApJ, 540, 308

Dickman R. L., Snell R. L., Schloerb F. P., 1986, ApJ, 309, 326

Eden D. J. et al., 2018, MNRAS, 477, 3369

Endres C. P., Schlemmer S., Schilke P., Stutzki J., Müller H. S. P., 2016, J. Mol. Spectrosc., 327, 95

Evans N. J. II, 1999, ARA&A, 37, 311

Evans Neal J. I., 1989, Rev. Mex. Astron. Astrofís., 18, 21

Evans N. J., Kim K.-T., Wu J., Chao Z., Heyer M., Liu T., Nguyen-Luong Q., Kauffmann J., 2020, ApJ, 894, 103

Federrath C., Roman-Duval J., Klessen R. S., Schmidt W., Mac Low M.-M., 2010, A&A, 512, A81

Federrath C. et al., 2016, ApJ, 832, 143

Fitzpatrick E. L., 1999, PASP, 111, 63

Fontani F., Giannetti A., Beltrán M. T., Dodson R., Rioja M., Brand J., Caselli P., Cesaroni R., 2012, MNRAS, 423, 2342

Galván-Madrid R., Zhang Q., Keto E., Ho P. T. P., Zapata L. A., Rodríguez L. F., Pineda J. E., Vázquez-Semadeni E., 2010, ApJ, 725, 17

Galván-Madrid R. et al., 2013, ApJ, 779, 121

Gao Y., Solomon P. M., 2004a, ApJS, 152, 63

Gao Y., Solomon P. M., 2004b, ApJ, 606, 271

García-Burillo S., Usero A., Alonso-Herrero A., Graciá-Carpio J., Pereira-Santaella M., Colina L., Planesas P., Arribas S., 2012, A&A, 539, A8

Giannetti A. et al., 2014, A&A, 570, A65

Ginard D. et al., 2015, A&A, 578, A49

Goldsmith P. F., Kauffmann J., 2017, ApJ, 841, 25

Gong M., Ostriker E. C., Kim C.-G., 2018, ApJ, 858, 16

Guzmán A. E., Sanhueza P., Contreras Y., Smith H. A., Jackson J. M., Hoq S., Rathborne J. M., 2015, ApJ, 815, 130

Gwinn C. R., Moran J. M., Reid M. J., 1992, ApJ, 393, 149

Heiderman A., Evans Neal J. I., Allen L. E., Huard T., Heyer M., 2010, ApJ, 723, 1019

Hennebelle P., Chabrier G., 2013, ApJ, 770, 150

Hernandez A. K., Tan J. C., Caselli P., Butler M. J., Jiménez-Serra I., Fontani F., Barnes P., 2011, ApJ, 738, 11

Homeier N. L., Alves J., 2005, A&A, 430, 481

Jiménez-Donaire M. J. et al., 2019, ApJ, 880, 127

Kauffmann J., Goldsmith P. F., Melnick G., Tolls V., Guzman A., Menten K. M., 2017, A&A, 605, L5

Kramer C., Alves J., Lada C. J., Lada E. A., Sievers A., Ungerechts H., Walmsley C. M., 1999, A&A, 342, 257

Krumholz M. R., McKee C. F., 2005, ApJ, 630, 250

Lacy J. H., Sneden C., Kim H., Jaffe D. T., 2017, ApJ, 838, 66

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

Lampton M., Margon B., Bowyer S., 1976, ApJ, 208, 177

Langer W. D., Graedel T. E., Frerking M. A., Armentrout P. B., 1984, ApJ, 277, 581

Larson R. B., 1981, MNRAS, 194, 809

Leroy A. K. et al., 2016, ApJ, 831, 16

Leurini S., Schilke P., Menten K., Flower D., Pottage J., Xu L.-H., 2004, A&A, 422, 573

Lovas F. J., 2004, J. Phys. Chem. Ref. Data, 33, 177

Meier D. S., Turner J. L., 2005, ApJ, 618, 259

Menten K., Walmsley C., Henkel C., Wilson T., 1988, A&A, 198, 253

Molinari S. et al., 2010, PASP, 122, 314

Molinari S. et al., 2011, ApJ, 735, L33

Nagy Z., van der Tak F. F. S., Fuller G. A., Spaans M., Plume R., 2012, A&A, 542, A6

Nagy Z., van der Tak F. F. S., Fuller G. A., Plume R., 2015, A&A, 577, A127

Nguyen-Luong Q. et al., 2020, ApJ, 891, 66

Nishimura Y., Watanabe Y., Harada N., Shimonishi T., Sakai N., Aikawa Y., Kawamura A., Yamamoto S., 2017, ApJ, 848, 17

Orkisz J. H. et al., 2017, A&A, 599, A99

Padoan P., Nordlund Å., 2011, ApJ, 730, 40

Peng T.-C., Wyrowski F., van der Tak F., Menten K., Walmsley C., 2010, A&A, 520, A84

Pety J. et al., 2017, A&A, 599, A98

Radhakrishnan V., Goss W. M., Murray J. D., Brooks J. W., 1972, ApJS, 24, 49

Roberts H., van der Tak F. F. S., Fuller G. A., Plume R., Bayet E., 2011, A&A, 525, A107

Röllig M., Ossenkopf V., 2013, A&A, 550, A56

Rugel M. R. et al., 2019, A&A, 622, A48

Sage L. J., Ziurys L. M., 1995, ApJ, 447, 625

Sakamoto K., Aalto S., Evans A. S., Wiedner M. C., Wilner D. J., 2010, ApJ, 725, L228

Schinnerer E. et al., 2013, ApJ, 779, 42

Schöier F. L., van der Tak F. F. S., van Dishoeck E. F., Black J. H., 2005, A&A, 432, 369

Shetty R., Glover S. C., Dullemond C. P., Klessen R. S., 2011a, MNRAS, 412, 1686

Shetty R., Glover S. C., Dullemond C. P., Ostriker E. C., Harris A. I., Klessen R. S., 2011b, MNRAS, 415, 3253

Shimajiri Y. et al., 2017, A&A, 604, A74

Shirley Y. L., 2015, PASP, 127, 299

Solomon P. M., Rivolo A. R., Barrett J., Yahil A., 1987, ApJ, 319, 730

Stock D., Peeters E., Choi W.-Y., Shannon M., 2014, ApJ, 791, 99

van der Tak F. F. S., Black J. H., Schöier F. L., Jansen D. J., van Dishoeck E. F., 2007, A&A, 468, 627

Vastel C., Spaans M., Ceccarelli C., Tielens A. G. G. M., Caux E., 2001, A&A, 376, 1064

Watanabe Y., Sakai N., Sorai K., Yamamoto S., 2014, ApJ, 788, 4

Watanabe Y. et al., 2015, ApJ, 809, 162

Watanabe Y., Nishimura Y., Harada N., Sakai N., Shimonishi T., Aikawa Y., Kawamura A., Yamamoto S., 2017, ApJ, 845, 116

Welch W. J., Dreher J. W., Jackson J. M., Terebey S., Vogel S. N., 1987, Science, 238, 1550

Westerhout G., 1958, Bull. Astron. Inst. Neth., 14, 215

Wu J., Evans Neal J. I., Shirley Y. L., Knez C., 2010, ApJS, 188, 313

Zhang B., Reid M. J., Menten K. M., Zheng X. W., Brunthaler A., Dame T. M., Xu Y., 2013, ApJ, 775, 79

Zhu H., Tian W. W., Zuo P., 2014, ApJ, 793, 95

## SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

**online.zip**

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

## APPENDIX A: LEGO MOLECULAR LINE DATA BASE INFORMATION

Tables 2 and A1 record the information on molecular emission lines the LEGO survey uses to observe and characterize emission lines. Line frequencies are from Lovas (2004), while energy levels, Einstein coefficients and downward collisional rate coefficients are taken from the Leiden Atomic and Molecular Database (Schöier et al. 2005; LAMDA, accessed in November 2019). Section 2.1.2 provides further information on how Tables 2 and A1 were constructed.

## APPENDIX B: INTEGRATED INTENSITY MAP STATISTICS

In this section, we present statistics determined from the molecular line integrated intensity maps (see Section 2.1.3). Shown in Table B1 is the minimum value of all pixels across each map (i.e. without taking a significance threshold). We see that CO and $^{13}$CO have positive minimum values, which shows that they are detected across all positions within the map (also see Fig. 3). The other molecular lines, however, are negative, showing that they are detected above a threshold intensity. This is given as the minimum value above $W_Q > 3\sigma_{W_Q}$ also given in Table B1. In addition, we also give the minimum, 5, 16, 50, 84, and 95 percentile ranges when taking only the significant positions. Note that Table B1 has been ordered by decreasing $A_{cov}$ values (see Table 3). We can see that there is little correlation to the line brightness and its coverage across the region (e.g. see $W_Q^{max}$). Some lines are, therefore, seen as extended as relative faint [e.g. CCH (1-0, 1/2-1/2)], whilst other are seen as compact and bright [e.g. SiO (2-1); see Fig. 7].

**Table A1.** Details on the emission lines covered by the LEGO survey, and processed by our pipeline. Section 2.1.2 explains how lines examined in this study were selected. Each of the selected lines was given a LEGO reference code in the reduction pipeline, which are presented in the first column of the table. The rest frequency recorded in this table, $\nu_{rest}$, refers to the specific full transition (Lovas 2004). These transitions are sometimes part of larger groups, as explained in Section 2.1.2 (see the final column of this table). In those cases, $\nu_{min}$ and $\nu_{max}$ state the minimum and maximum frequencies of the lines considered to form a group. The full, machine-readable version of this table can be obtained from the supplementary online material.

| LEGO ID | Species | Full transition | Frequencies (GHz) | | | Notes |
|---|---|---|---|---|---|---|
| | | | $\nu_{rest}$ | $\nu_{min}$ | $\nu_{max}$ | |
| H13CN_86 | H$^{13}$CN | $J = 1-0, F = 2-1$ | 86.3401764 | 86.3387367 | 86.3422551 | Group of 3 transitions |
| H13CO+_87 | H$^{13}$CO$^+$ | $J = 1-0$ | 86.7542880 | ... | ... | ... |
| SiO_87 | SiO | $J = 2-1$ | 86.8469950 | ... | ... | ... |
| N13C_87 | HN$^{13}$C | $J = 1-0, F = 2-1$ | 87.0908590 | 87.090735 | 87.090942 | Group of 3 transitions |
| CCH_87a | CCH | $N = 1-0, J = 3/2-1/2, F = 2-1$ | 87.3169250 | 87.284156 | 87.328624 | Group of 3 transitions |
| ... | ... | ... | ... | ... | ... | ... |

**Table B1.** Integrated intensity statistics across the mapped region (i.e. that covered with both vertical and horizontal on-the-fly scans). The columns show the molecule name, the minimum values of all pixels within the map, and the minimum, 5, 16, 50, 84, and 95 percentile ranges for the intensities at positions with $W_Q > 3\sigma_{W_Q}$. The full, machine-readable version of this table can be obtained from the supplementary online material.

| Line | $W_Q^{\min}$ (all) | $W_Q^{\min}$ | $W_Q^{5\%}$ | $W_Q^{16\%}$ | $W_Q^{\text{median}}$ (K km s$^{-1}$) | $W_Q^{84\%}$ | $W_Q^{95\%}$ | $W_Q^{\max}$ |
|---|---|---|---|---|---|---|---|---|
| CO (1-0) | 25.19 | 25.19 | 41.49 | 54.19 | 75.69 | 113.52 | 147.71 | 551.06 |
| $^{13}$CO (1-0) | 0.49 | 0.86 | 3.57 | 6.19 | 11.89 | 19.98 | 28.71 | 155.55 |
| C$^{18}$O (1-0) | −1.36 | 0.60 | 0.77 | 0.91 | 1.34 | 2.29 | 3.51 | 17.31 |
| HCN (1-0) | −1.84 | 0.54 | 0.71 | 0.82 | 1.26 | 2.69 | 6.91 | 74.13 |
| HCO$^+$ (1-0) | −2.07 | 0.58 | 0.72 | 0.82 | 1.25 | 2.74 | 6.39 | 72.50 |
| N$_2$H$^+$ (1-0) | −1.26 | 0.39 | 0.45 | 0.49 | 0.69 | 1.45 | 4.40 | 14.05 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |

## APPENDIX C: RANDOM UNCERTAINTIES ON THE H$_2$ COLUMN DENSITY AND DUST TEMPERATURE MAPS

We follow the weighted least squared fitting procedure outlined in Guzmán et al. (2015) to calculate the molecular hydrogen column densities and dust temperatures from the *Herschel* Hi-Gal observations (see Section 2). This involves fitting the *Herschel* filter fluxes to the modified blackbody model, and adjusting the free parameters (surface density and dust temperature) in order to minimize the weighted least squared difference:

$$\chi^2 = \sum_i \frac{(\text{data}_i - \text{model}_i)}{\sigma_i^2}, \qquad (C1)$$

where $i$ runs through each filter, $\sigma_i^2 = (0.1 * \text{data}_i)^2 + \sigma_{i,\text{N}}^2$ and $\sigma_{i,\text{N}}$ is given by the fifth column of table 1 in Guzmán et al. (2015).

To get the random uncertainty associated with the fit at each position, we follow the method outlined by Lampton, Margon &

Bowyer (1976). We test how 'sensitive' is the minimization function to variations in the input parameters. In practice, this means finding the $\chi^2 = \chi^2_{\min} + 2.3$ contour (2.3 is adequate to get 1$\sigma$ confidence intervals with two free parameters). This contour (usually close to an ellipse) defines an uncertainty region, whose projections on to the parameter axes define the 1$\sigma$ confidence intervals (see fig. 3 in Guzmán et al. 2015). Fig. C1 shows the size of these confidence intervals towards the positive and negative side of the best-fitting parameter. The parameters shown here are the dust temperature, and the logarithm of the H$_2$ column density.

We investigate the fractional uncertainty across the mapped region by taking the ratios of the positive and negative uncertainty at each
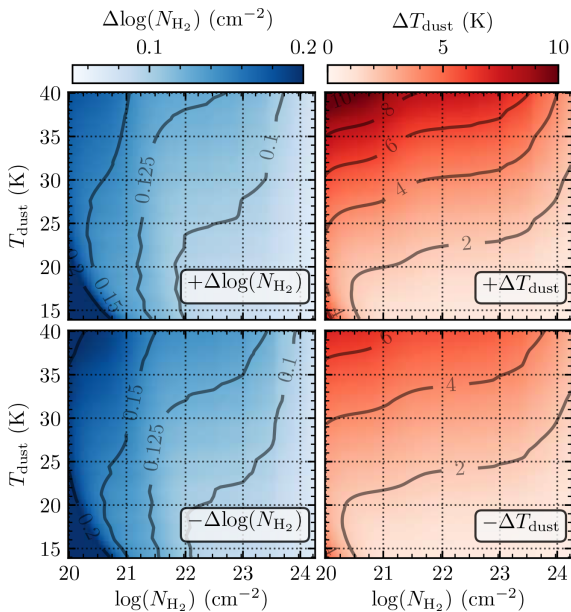


**Figure C1.** The random uncertainty associated with the H$_2$ column densities and dust temperatures. Shown from left to right in colour scale is the positive and negative uncertainties on the (log) H$_2$ column density, and dust temperature as a function of the H$_2$ column density and dust temperature. Overlaid and labelled on each panel are contours of the background colour scale for reference.
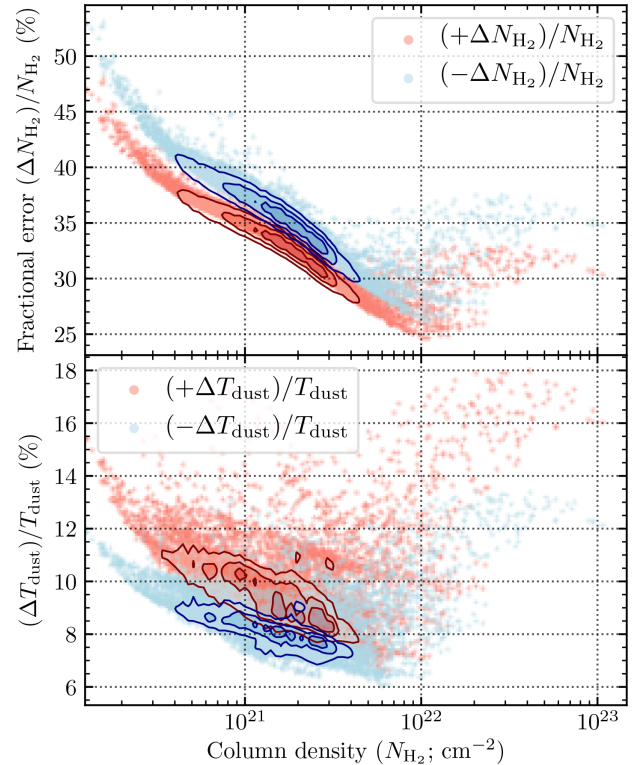


**Figure C2.** The fractional random uncertainty associated with the H$_2$ column density (upper panel) and dust temperature (lower panel) measurements across the mapped region. The fractional uncertainty is defined at each position as the ratio of the uncertainty over the measured value, and plotted as a function of the H$_2$ column density. The blue and red points show positive and negative uncertainties, respectively. The overlaid contours show the density of plotted points in increasing levels of 25, 50, and 75 per cent.

position over the corresponding measured value at the same position. The distribution of the $H_2$ column density and dust temperature fractional uncertainties are plotted in Fig. C2. We find that the mean fractional uncertainties produced by random noise on the column density and dust temperature maps are $\frac{\Delta N_{H_2}}{N_{H_2}} = ^{+33}_{-35}$ per cent and $\frac{\Delta T_{dust}}{T_{dust}} = ^{+9}_{-8}$ per cent.

## APPENDIX D: LINE OF SIGHT COMPONENTS IN SPECTRAL CUBES

In this section, we present a brief inspection of the velocity structure observed within the molecular line data cubes. To do so, we integrate the line intensity within discrete velocity windows across the velocity range where the significant emission is seen in the average spectra (Fig. 2). Fig. D1 displays the result when using velocity windows of $-20$ to $80$ km s$^{-1}$ in steps of $20$ km s$^{-1}$ for the CO (1−0), C$^{18}$O (1−0), and HCN (1−0) maps. We note that to produce these maps we use cubes that have all spectral channels below $5\,\sigma_{rms}$ masked (see Table 2), rather than using the masked cubes produced by the CO masking routine (see Section 2.1.3). These plots highlight the

three broad velocity components that are present across the mapped region, which peak at velocities of $\sim10$, $\sim40$, $\sim60$ km s$^{-1}$. The brightest of these is the $\sim10$ km s$^{-1}$ component, which is seen in all the displayed lines and appears to be spatially correspondent with the W49A star-forming region (see Fig. 1).

## APPENDIX E: UNDERSTANDING THE MASS CONVERSION FACTORS

In this section, we review the calculation of the mass conversion factor discussed in Section 5.3; namely $\alpha_Q = L_Q/M$. The determination of a constant value for this conversion is fundamentally based on two assumptions: (1) the region is in virial equilibrium, i.e. not undergoing collapse or expansion, (2) the observed emission is optically thick.

A parcel of gas is defined as being in virial equilibrium when twice the internal kinetic energy equals the potential energy, which following Solomon et al. (1987) can be expressed as,
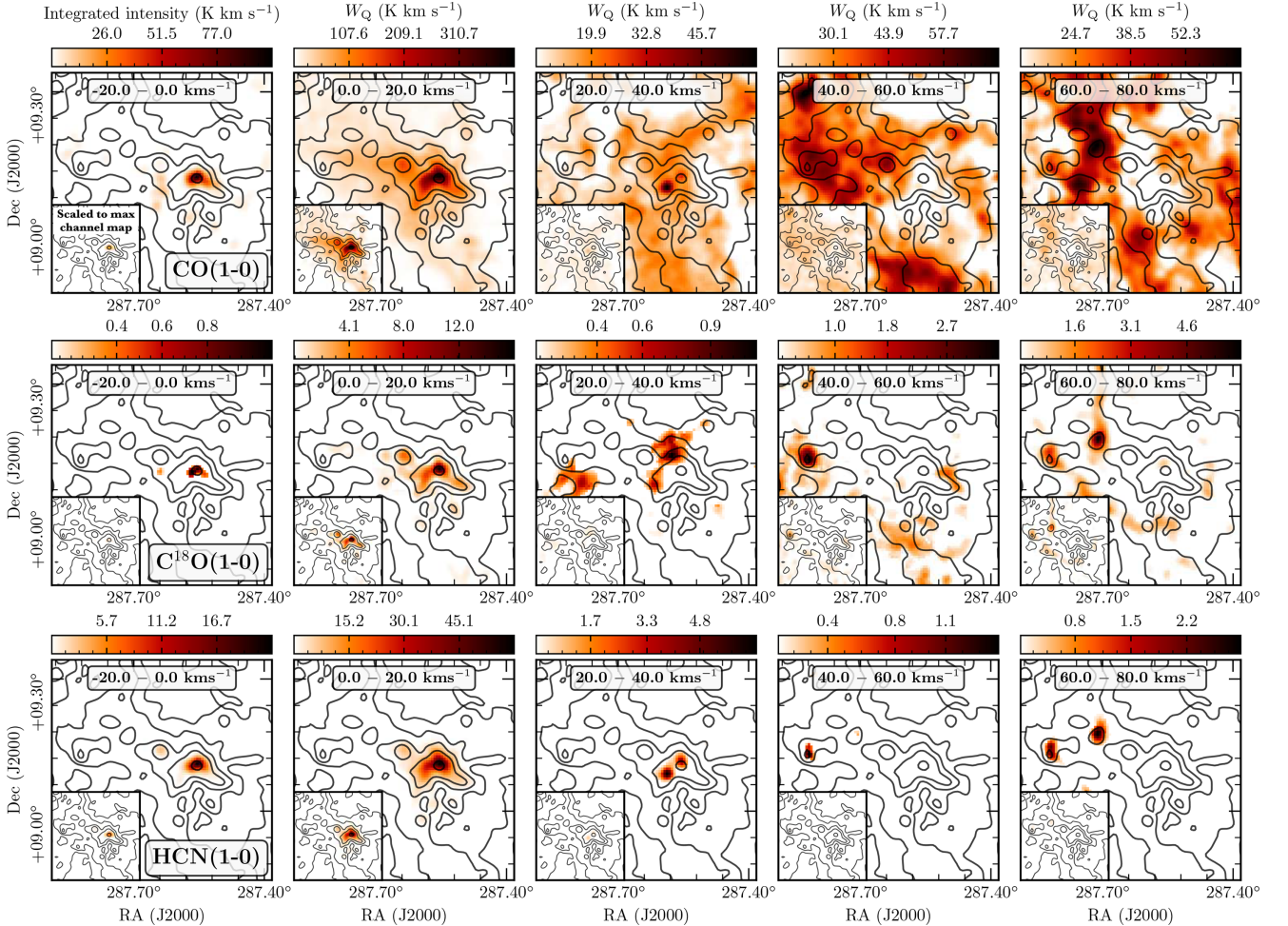
$$M_{vir} \propto R\sigma^2, \tag{E1}$$

**Figure D1.** Channel maps for CO (1−0), C$^{18}$O (1−0), and HCN (1−0) across the mapped region of W49. The panels from left to right show the intensity integrated in velocity windows of $-20$ to $80$ km s$^{-1}$ in steps of $20$ km s$^{-1}$ (see Section 2.1.3). The maps shown in the main plot of each panel have a colour scale scaled to the values within that velocity window, whilst the maps shown in the lower left of each panel have the identical colour scale which has been scaled to the window with the highest integrated intensity. The black contours are of the *Herschel* derived molecular hydrogen column density, in levels of 1, 2.5, 5, 10, 50, 100 $\times 10^{22}$ cm$^{-2}$.

**Table E1.** Calculated conversion factors for the various molecular line transitions (see Section 5.3). We highlight the two most commonly adopted conversion factors: $X_{CO(1-0)}$ and $\alpha_{HCN(1-0)}$. The values shown in the upper half of the table have been calculated using the column density and mass of $A_v > 8$ mag gas divided by the mean integrated intensity and total luminosity of the given line across the region (e.g. $\alpha_Q = M^{sum}_{A_v>8\,mag}/L^{sum}_Q$), whilst those within the lower half have been calculated using both the column density and mass, and integrated intensity and luminosity above the $A_v > 8$ mag threshold ($\alpha_Q(A_v > 8\,mag) = M^{sum}_{A_v>8\,mag}/L^{sum}_{Q,A_v>8\,mag}$). The full, machine-readable version of this table can be obtained from the supplementary online material.

| Line | $\log(W^{mean}_Q)$ (K km s$^{-1}$) | $\log(L^{sum}_Q)$ (K km s$^{-1}$ pc$^2$) | $X_Q$ [cm$^{-2}$ (K km s$^{-1}$)$^{-1}$] ($\times 10^{20}$) | $\alpha_Q$ [M$_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$] |
|---|---|---|---|---|
| | $W^{mean}_Q$; $L^{sum}_Q$; $N^{mean}_{H_2,A_v>8\,mag} = 1.7 \times 10^{21}$ cm$^{-2}$; $M^{sum}_{A_v>8\,mag} = 4.0 \times 10^5$ M$_\odot$ | | | |
| CO(1-0) | $1.928 \pm 0.003$ | $5.911 \pm 0.002$ | $2.02 \pm 0.094$ | $0.491 \pm 0.005$ |
| HCN(1-0) | $0.15 \pm 0.02$ | $4.13 \pm 0.014$ | $124 \pm 6.9$ | $30.1 \pm 0.98$ |
| ... | $W^{mean}_{Q,A_v>8\,mag}$; $L^{sum}_{Q,A_v>8\,mag}$; $N^{mean}_{H_2,A_v>8\,mag} = 1.7 \times 10^{22}$ cm$^{-2}$; $M^{sum}_{A_v>8\,mag} = 4.0 \times 10^5$ M$_\odot$ | ... | ... | ... |
| CO(1-0) | $2.32 \pm 0.01$ | $4.9 \pm 0.01$ | $0.82 \pm 0.042$ | $5.2 \pm 0.13$ |
| HCN(1-0) | $1.18 \pm 0.03$ | $3.75 \pm 0.023$ | $11.5 \pm 0.8$ | $73 \pm 3.9$ |
| ... | ... | ... | ... | ... |

**Table E2.** Calculated conversion factors for the various molecular line transitions (identical to Table E1). The values shown have been calculated without imposing any extinction threshold on the gas column density and gas mass, and the integrated intensity and luminosity of the given line (such that e.g. $\alpha_Q = M^{sum}/L^{sum}_Q$). The full, machine-readable version of this table can be obtained from the supplementary online material.

| Line | $\log(W^{mean}_Q)$ (K km s$^{-1}$) | $\log(L^{sum}_Q)$ (K km s$^{-1}$ pc$^2$) | $X_Q$ [cm$^{-2}$ (K km s$^{-1}$)$^{-1}$] ($\times 10^{20}$) | $\alpha_Q$ [M$_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$] |
|---|---|---|---|---|
| | $W^{mean}_Q$; $L^{sum}_Q$; $N^{mean}_{H_2} = 2.5 \times 10^{21}$ cm$^{-2}$; $M^{sum} = 5.4 \times 10^5$ M$_\odot$ | | | |
| CO(1-0) | $1.928 \pm 0.003$ | $5.911 \pm 0.002$ | $0.297 \pm 0.0057$ | $0.67 \pm 0.013$ |
| HCN(1-0) | $0.15 \pm 0.02$ | $4.13 \pm 0.014$ | $18.2 \pm 0.67$ | $41 \pm 1.5$ |
| ... | ... | ... | ... | ... |

where $M_{vir}$ is the virial mass, $R$ the radius, and $\sigma$ the observed velocity dispersion. The virial mass can also be given as a product of its density, $\rho$, and volume,

$$M_{vir} \propto \rho R^3. \tag{E2}$$

Empirically, molecular clouds are observed to follow a size–line width relation (Larson 1981),

$$\sigma \propto R^{0.5}. \tag{E3}$$

Equating equations (E1) and (E2), and substituting equation (E3) gives ($R \propto \sigma^2$),

$$\sigma \propto \rho^{-0.5}. \tag{E4}$$

Equating equations (E3) and (E4) gives,

$$R \propto \rho^{-1}, \tag{E5}$$

and substituting this into equation (E2),

$$M_{vir} \propto \rho^{-2}. \tag{E6}$$

The luminosity of a molecular line transition, $L$, can be approximated as the product of the integrated intensity ($W = T_{MB}\sqrt{2}\sigma$), and the area over which the intensity has been measured ($A = \pi R^2$),

$$L = AW \propto T_{MB}\sigma R^2 \tag{E7}$$

Substituting equations (E4) and (E5) in the above gives,

$$L \propto T_{MB}\rho^{-2.5}. \tag{E8}$$

Finally, taking the ratio of the virial mass in equation (E6), and the luminosity in equation (E8) yields,

$$\alpha = ML^{-1} \propto \rho^{-0.5} (T_{MB}\rho^{-2.5})^{-1} \propto \frac{\rho^{0.5}}{T_{MB}}, \tag{E9}$$

which in terms of molecular hydrogen number density, $n_{H_2}$, can be approximated as,

$$\alpha \approx 2.1 \frac{n_{H_2}^{0.5}}{T_{MB}}. \tag{E10}$$

Gao & Solomon (2004a,b) approximate that a virialized core with a volume averaged molecular hydrogen number density of $n_{H_2} \sim 3 \times 10^4$ cm$^{-3}$, and optically think HCN (1−0) emission with a brightness temperature of $T_{MB} = 35$ K, gives $\alpha_{HCN(1-0)} = 10$ M$_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$. Alternatively, in this work we measure a peak main beam brightness temperatures for HCN of $T_{MB} \sim 4$ K, and characteristic number densities of $\sim 10^{3-4}$ cm$^{-3}$ (Section 5.2), which would then give a significantly higher than standard values of $\alpha_{HCN(1-0)} \sim 15$ to $-50$ M$_\odot$ (K km s$^{-1}$ pc$^2$)$^{-1}$ [see Appendix E for the full derivation of $\alpha_{HCN(1-0)}$].

The column density conversion $X_Q = W_Q/N_{H_2}$ follows a very similar calculation, and can be found in full in e.g. Dickman, Snell & Schloerb (1986).

[1] *Argelander-Institut für Astronomie (AIfA), Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany*

[2] *Haystack Observatory, Massachusetts Institute of Technology, Westford, MA 01886, USA*

[3] *Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany*

[4] *National Astronomical Observatory of Japan, National Institutes of Natural Sciences, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan*

[5] *Instituto de Radioastronomía Milimétrica (IRAM), Avenida Divina Pastora, 7, 18012, Granada, Spain*

[6] *Max-Planck-Institute for Extraterrestrial Physics (MPE), Giessenbachstr 1, D-85748 Garching, Germany*

[7] *Laboratoire d'astrophysique de Bordeaux, Univ. Bordeaux, CNRS, B18N, allée Geoffroy Saint-Hilaire, F-33615 Pessac, France*

[8] *Department of Space, Earth and Environment, Chalmers University of Technology, Onsala Observatory, SE-439 92 Onsala, Sweden*

[9] *Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Stop C1400, Austin, TX 78712-1205, USA*

[10] *Institut für Theoretische Astrophysik, Zentrum für Astronomie der Universität Heidelberg, Albert-Ueberle-Strasse 2, D-69120 Heidelberg, Germany*

[11] *Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA*

[12] *Institut de Radioastronomie Millimé, 300 rue de la Piscine, Domaine Universitaire, F-38406 Saint Martin d'Hè, France*

[13] *Institute of Astronomy, The University of Tokyo, 2-21-1, Osawa, Mitaka, Tokyo 181-0015, Japan*

[14] *National Astronomical Observatory of Japan, 2-21-1, Osawa, Mitaka, Tokyo 181-8588, Japan*

[15] *Department of Physics and Astronomy, UCL, Gower St., London WC1E 6BT, UK*

[16] *Materials Science and Engineering, College of Engineering, Shibaura Institute of Technology, 3-7-5 Toyosu, Koto-ku, Tokyo 135-8548, Japan*

[17] *Physics and Astronomy, University of Exeter, Stocker Road, Exeter EX4 4QL, UK*

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