

# Direct observations of the atomic-molecular phase transition in the Milky Way’s nuclear wind

Karlie A. Noon<sup>1</sup>,<sup>1</sup>★ Mark R. Krumholz<sup>1,2</sup>, Enrico M. Di Teodoro<sup>3</sup>, Naomi M. McClure-Griffiths,<sup>1</sup> Felix J. Lockman<sup>4</sup> and Lucia Armillotta<sup>5</sup>

<sup>1</sup>Research School of Astronomy and Astrophysics, The Australian National University, Canberra, Australian Capital Territory 2611, Australia

<sup>2</sup>ARC Centre of Excellence for Astronomy in Three Dimensions (ASTRO3D), Canberra, ACT 2611, Australia

<sup>3</sup>Dipartimento di Fisica e Astronomia, Università degli Studi di Firenze, via Giovanni Sansone 1, I-50019 Sesto Fiorentino, Firenze, Italy

<sup>4</sup>National Radio Astronomy Observatory, Green Bank, WV 24944, USA

<sup>5</sup>Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA

Accepted 2023 June 18. Received 2023 June 8; in original form 2023 April 12

## ABSTRACT

Hundreds of high-velocity atomic gas clouds exist above and below the Galactic Centre, with some containing a molecular component. However, the origin of these clouds in the Milky Way’s wind is unclear. This paper presents new high-resolution MeerKAT observations of three atomic gas clouds and studies the relationship between the atomic and molecular phases at  $\sim 1$  pc scales. The clouds’ atomic hydrogen column densities,  $N_{\text{H I}}$ , are less than a few  $\times 10^{20} \text{ cm}^{-2}$ , but the two clouds closest to the Galactic Centre none the less have detectable CO emission. This implies the presence of  $\text{H}_2$  at levels of  $N_{\text{H I}}$  at least a factor of ten lower than in the typical Galactic interstellar medium. For the cloud closest to the Galactic Centre, detectable CO coexists across the entire range of H I column densities. In contrast, for the intermediate cloud, detectable CO is heavily biased towards the highest values of  $N_{\text{H I}}$ . The cloud most distant from the Galactic Centre has no detectable CO at similar  $N_{\text{H I}}$  values. Moreover, we find that the two clouds with detectable CO are too molecule-rich to be in chemical equilibrium, given the depths of their atomic shielding layers, which suggests a scenario whereby these clouds consist of pre-existing molecular gas from the disc that the Galactic wind has swept up, and that is dissociating into atomic hydrogen as it flows away from the Galaxy. We estimate that entrained molecular material of this type has a  $\sim \text{few} - 10$  Myr lifetime before photodissociating.

**Key words:** Galaxy: centre – ISM: clouds – ISM: kinematics and dynamics – ISM: molecules – radio lines: ISM.

## 1 INTRODUCTION

A large-scale multiphase bipolar outflow emanates from the Milky Way’s (MW) Galactic Centre (GC), driven either by past nuclear activity or by past periods of intense star formation taking place in the central molecular zone (e.g. Bland-Hawthorn & Cohen 2003; Su, Slatyer & Finkbeiner 2010; Crocker et al. 2015; Yang, Ruszkowski & Zweibel 2022). This nuclear wind includes components visible at all observable wavelengths ranging from the very hot X-ray emitting phase at  $T \sim 10^6\text{--}8$  K (e.g. Koyama et al. 1989; Snowden et al. 1997; Su et al. 2010), to the radio-bright cold atomic and molecular gas phase at just  $\sim 10\text{--}100$  K (e.g. Lockman 1984; Sofue & Handa 1984; McClure-Griffiths et al. 2013; Di Teodoro et al. 2020; Cashman et al. 2021). Gamma-ray observations by Su et al. (2010), Dobler et al. (2010), and Carretti et al. (2013) reveal that the wind also includes a relativistic component, called the Fermi Bubbles, that extends up to 10 kpc above and below the GC. The various components of the wind carry energy, baryons, and metals from the interstellar medium to the Galaxy’s circumgalactic medium (CGM).

The MW’s wind is interesting in its own right, but it also plays an important role in the study of galactic winds (GWs) more broadly. Since it is at a distance of only  $\sim 8$  kpc, it serves as the closest

laboratory to study the physics of the GWs responsible for carrying large amounts of star-forming material out of galactic discs, regulating star formation, and significantly impacting galactic evolution. The mechanisms that drive the winds, their physical properties, the spatial and mass distribution of their multiphase components, and their impacts on the host galaxy are still poorly understood (Veilleux et al. 2020).

One problem for which the MW wind can provide special insight is the origin of cold atomic and molecular gas in outflows. Such gas has been seen in numerous extragalactic observations (e.g. Morganti, Oosterloo & Tsvetanov 1998; Seaquist & Clark 2001; Greve 2004; Leroy et al. 2015; Martini et al. 2018), and conceivably dominates the total wind mass flux (e.g. Yuan, Krumholz & Martin 2022), but its presence is difficult to understand since simulations suggest that the material driving hot winds should rapidly shock-heat cold gas to high temperatures (e.g. Schneider & Robertson 2017). A large number of models have been proposed to explain the survival of cool gas (e.g. McCourt et al. 2015; Huang et al. 2020; Schneider et al. 2020; Kanjilal, Dutta & Sharma 2021), but these have proven difficult to test in observations due to the limited spatial resolution available in extragalactic data, which precludes resolving individual clouds entrained in the wind. By contrast, modern radio telescopes can resolve  $\sim 1$  pc spatial scales in the MW wind, making it possible to study the bulk properties and morphology of individual cold clouds suspended in it.

★ E-mail: [karlie.noon@anu.edu.au](mailto:karlie.noon@anu.edu.au)

A second, closely related problem where the MW again offers unique advantages is in understanding how the phase structure of winds evolves with distance from the host galaxy. Large amounts of atomic and molecular gas are entrained in galactic outflows, yet galaxy circumgalactic media are mostly comprised of ionized gas (Werk et al. 2014). This implies that a phase transition must occur, and there are claims that this transition has been observed in the nearby starburst galaxy M82 (Leroy et al. 2015; Martini et al. 2018). However, the interpretation of the data as showing evidence for a phase transition remains controversial due to uncertainties about how the velocities of different chemical components of the wind evolve with height (Yuan et al. 2022). Similarly, there are theoretical model predictions for how GW phases vary with galactocentric distance (e.g. Fielding & Bryan 2022), but these have not yet been subjected to strong observational tests. Resolved observations of individual clouds at a range of galactocentric distances would be ideal for this purpose, but at present they are not available in external galaxies.

By contrast, a number of clouds suitable for studies of this type are known in the MW. Lockman (1984) and Lockman & McClure-Griffiths (2016) noted a large void of neutral hydrogen gas above and below the disc in the inner 2 kpc of the MW, potentially indicative of large-scale winds sweeping the region free of cold gas (Bregman 1980). Within this void region, McClure-Griffiths et al. (2013) observed 86 anomalous atomic hydrogen (H I) clouds above and below the GC with velocity distributions consistent with being entrained in an outflowing wind. Since then, Di Teodoro et al. (2018) and Lockman, Di Teodoro & McClure-Griffiths (2020), combined with McClure-Griffiths et al. (2013), have detected a total population of approximately 200 cold, dense H I clouds that have kinematics consistent with being entrained in a nuclear outflow from the GC. From this population, Di Teodoro et al. (2020) observed that two clouds were associated with small clumps of cold molecular matter. This represents the first observations of molecular matter entrained in the MW's nuclear wind.

Three plausible scenarios have been proposed to explain the origin and survival of these molecular structures, each carrying with it different assumptions about where cold gas comes from and how gas phases change as the wind flows. One scenario is that the nuclear wind has blown chunks of molecular clouds from the disc that, as they flow outward, dissociate (e.g. Espada et al. 2010; Saito et al. 2022a, b) into H I (and possibly thereafter ionize) due to the surrounding radiation and gas. In a second scenario, the hot wind has entrained atomic gas from the disc and compressed it into molecular form due to the high pressure. Finally, a third scenario is that the wind was originally composed mostly of hot, ionized material that has subsequently condensed and cooled into atoms and molecules due to thermal instabilities (e.g. Thompson et al. 2016; Schneider, Robertson & Thompson 2018; Maccagni et al. 2021). In the first of these scenarios, there is also a possibility that the wind has entrained atomic hydrogen along with the molecular matter, as atomic and molecular gas have long been observed to be correlated (e.g. Simonson 1973; Wannier, Lichten & Morris 1983; Andersson, Roger & Wannier 1992), and on theoretical grounds, all molecular clouds in the disc are expected to be protected by a dusty atomic envelope (Krumholz, McKee & Tumlinson 2008, 2009; McKee & Krumholz 2010). Resolved observations of individual clouds can differentiate between these scenarios, which in turn will provide useful insight into how wind phase structures evolve as gas is blown away from the disc into the CGM. The insight we gain from the MW can help us understand winds more generally.

To further investigate the phase structure of material entrained in the MW's wind, this paper will focus on three bright H I clouds

identified in prior surveys (McClure-Griffiths et al. 2013; Di Teodoro et al. 2018). Two of these clouds have been observed in the  $^{12}\text{CO}(2\rightarrow 1)$  emission line (Di Teodoro et al. 2020), while a third was observed but not detected, yielding informative upper limits. Low- $J$  CO lines are an accessible and widely used proxy for molecular gas, allowing us to explore the chemical states of the clouds. We complement these data with new, high-resolution H I interferometric observations from MeerKAT, allowing us for the first time to study the phase structure of MW wind clouds with high, matched resolution for the atomic and molecular phases.

The remainder of this paper is structured as follows: First, in Section 2 we present both the new MeerKAT 21 cm observations and CO data, and describe how we extract column densities and masses for the atomic and molecular components. Section 3 compares the H I and CO data, both from a morphological and quantitative point of view, and analyses the chemical state of the gas. We interpret these observations in light of the scenarios described above in Section 4, and summarize our findings in Section 5.

## 2 OBSERVATIONS AND DATA REDUCTION

In this section, we outline the observations of the three target clouds that are the focus of this paper. Hereinafter, we will refer to these sources as C1, C2, and C3. We describe the H I observations and processing in Section 2.1 and the corresponding procedures for the CO data in Section 2.2. We summarize some features of the observations in Table 1.

### 2.1 H I observations

We carried out high-resolution observations of the 21 cm H I emission line in December 2020 using the MeerKAT radio interferometer (see Jonas & MeerKAT Team 2016 for a description of the telescope). We observed the three objects over a total integration time of 3 h using a single pointing for each cloud centred at Galactic coordinates  $(l, b) = (-1.86^\circ, -3.84^\circ)$ ,  $(l, b) = (-2.42^\circ, 5.56^\circ)$ , and  $(l, b) = (13.30^\circ, 7.00^\circ)$  for C1, C2, and C3, respectively. For these observations, 59 of the 64 MeerKAT antennas were available, giving 1770 baselines. The shortest and longest baselines were 29 m and 8 km, respectively. The observations have a bandwidth of 856–712 MHz across 32 768 channels at a spectral resolution of 26.12 kHz, corresponding to a velocity resolution of  $5.5 \text{ km s}^{-1}$  at the H I frequency of 1420.405 MHz.

The MeerKAT pipeline calibrated the raw data. The pipeline includes flagging of radio frequency interference, bandpass calibration, and receiver gain fluctuations calibration. For an in-depth guide to the MeerKAT calibration pipeline, see Wang et al. (2021). The South African Radio Astronomy Observatory's archive<sup>1</sup> stores the calibrated  $u-v$  data set. To image these data, we use the Common Astronomy Software Applications (CASA; McMullin et al. 2007), version 5.4.1-32. We first split the full  $u-v$  data set into three separate data sets, one for each source. Using the *uvcontsub* task in CASA, we generate a model of the continuum sources by fitting emission-free channels to a line. The *uvcontsub* function subtracts this model from the  $u-v$  data set. We use the *tclean* task to image each of the three sources into separate data cubes. We use a linear interpolator, a standard gridded with a cell size of 5 arcsec, a common restoring beam, a multiscale deconvolver with scales = [0, 6, 10, 20, 40, 80] pixels and natural weighting. To mask the images, we use

<sup>1</sup><https://archive.sarao.ac.za/>

**Table 1.** Measured and derived properties of observed clouds. Galactic longitudes  $l$  and latitudes  $b$  record the centres of the fields of view.  $V_{\text{LSR}}$  is the H I central local standard of rest velocity,  $N_{\text{H,H I,peak}}$  and  $N_{\text{H,H}_2,\text{peak}}$  are the peak derived column density of the H I and H<sub>2</sub> data, respectively, given in units H nuclei per cm<sup>2</sup> – thus, the quantity we report for the H<sub>2</sub> data is the number of H nuclei per unit area we infer, not the number of H<sub>2</sub> molecules, so the H I and H<sub>2</sub> numbers are directly comparable.  $M_{\text{at}}$  is the derived H I mass from equation (3), and  $M_{\text{mol}}$  is the corresponding derived H<sub>2</sub> mass from equation (7).

| Cloud ID | $l$<br>(°) | $b$<br>(°) | $V_{\text{LSR}}$<br>(km s <sup>-1</sup> ) | $N_{\text{H,H I,peak}}$<br>(cm <sup>-2</sup> ) | $N_{\text{H,H}_2,\text{peak}}$<br>(cm <sup>-2</sup> ) | $M_{\text{at}}$<br>(M <sub>⊙</sub> ) | $M_{\text{mol}}$<br>(M <sub>⊙</sub> ) |
|----------|------------|------------|-------------------------------------------|------------------------------------------------|-------------------------------------------------------|--------------------------------------|---------------------------------------|
| C1       | −1.85      | −3.87      | 165                                       | $1.8 \times 10^{20}$                           | $1.38 \times 10^{21}$                                 | 342                                  | 394                                   |
| C2       | −2.43      | 5.55       | 264                                       | $2.3 \times 10^{20}$                           | $1.00 \times 10^{21}$                                 | 721                                  | 571                                   |
| C3       | 13.25      | 7.0        | −99                                       | $1.5 \times 10^{20}$                           | –                                                     | 3429                                 | –                                     |

an automultithresh mask with a side lobe threshold of 3, a noise threshold of 5, a low-noise threshold of 1.5, and a negative threshold of 7. Cleaning ceases after reaching the interaction limit of 10 000 (as was the case for C1) or when the cleaning procedure reaches a stopping threshold which is given by the noise level of the dirty cube (as was the case for C2 and C3).

At 1420 MHz, the restored beam is 24.9 arcsec × 21.7 arcsec with position angle  $-2.8^\circ$ . The primary beam full width at half-maximum is 67 arcmin × 67 arcmin. The rms noise ( $\sigma_{\text{chn}}$ ) is 304, 275, and 365 mK for C1, C2, and C3, respectively, per 5.5 km s<sup>-1</sup> channel. To correct for the missing short-spacing baselines, we combine the interferometric data with Green Bank Telescope (GBT) single-dish observations from Di Teodoro et al. (2018) using CASA’s *feather* task with a single-dish flux scaling factor of 1. See Appendix A for details on our comparison between the MeerKAT and GBT data.

We derive neutral hydrogen column densities from the velocity-integrated brightness temperatures we measure with MeerKAT as:

$$N_{\text{H I}} = 1.823 \times 10^{18} \int T_{\text{b}} dv \text{ cm}^{-2}, \quad (1)$$

where  $N_{\text{H I}}$  is the column density and  $T_{\text{b}}$  is the brightness temperature. The constant  $1.823 \times 10^{18}$  assumes that the gas is optically thin, which is usually a good assumption for relatively low-column density gas (Dickey & Lockman 1990). Using the PYTHON package SPECTRALCUBE (Ginsburg et al. 2015), we evaluate the velocity-integrated brightness temperature,  $\int T_{\text{b}} dv$ . We show the resulting H I column density maps within a given velocity range for each cloud in the left-hand column of Fig. 1.

Given  $N_{\text{H I}}$ , the mass surface density of gas traced by H I is

$$\Sigma_{\text{at}} = \left( \frac{N_{\text{H I}}}{9.18 \times 10^{19} \text{ cm}^{-2}} \right) M_{\odot} \text{ pc}^{-2} \quad (2)$$

including helium; the numerical factor in the denominator assumes a hydrogen mass fraction  $X = 0.74$ . The corresponding total mass for atomic gas is:

$$M_{\text{at}} = 1.36 m_{\text{H}} D^2 \int N_{\text{H I}} d\Omega, \quad (3)$$

where the factor  $1.36 = 1/X$  is to account for helium,  $m_{\text{H}}$  is the mass of the hydrogen atom,  $D$  is the distance to the clouds, and the integral is over the solid angle,  $\Omega$ , subtended by the observations. We assume  $D = 8.2$  kpc (GRAVITY Collaboration 2019; see Appendix B for masses derived using kinematic distance estimates as described in Di Teodoro et al. 2020). We calculate the mass of C1 between velocity channels 148–181 km s<sup>-1</sup> within an area defined by  $(l, b) = (-1.65^\circ \text{ to } -2.05^\circ, -3.71^\circ \text{ to } -4.10^\circ)$ . All detections  $\geq 4\sigma_{\text{rms}}$  are included, where  $\sigma_{\text{rms}}$  is the rms noise of the 0th-moment map. For C2, we consider all detections  $\geq 4\sigma_{\text{rms}}$  between velocity channels 236–291 km s<sup>-1</sup> within an area defined

by  $(l, b) = (-2.16^\circ \text{ to } -2.66^\circ, 5.37^\circ \text{ to } 5.87^\circ)$ . For C3, we consider detections  $\geq 4\sigma_{\text{rms}}$  between velocity channels  $-121$  to  $-77$  km s<sup>-1</sup> within an area defined by  $(l, b) = (12.87^\circ \text{ to } 13.78^\circ, 6.50^\circ \text{ to } 7.40^\circ)$ . To determine the velocity range of the clouds, we include all channels where there are either H I or CO detections. As no CO is detected in C3, the width was determined solely from the H I emission.

We calculate  $\sigma_{\text{rms}}$  values individually for each cloud. To do so, we first calculate the channel standard deviation ( $\sigma_{\text{chn}}$ ) across a number of emission-free channels ( $N_{\text{chn}}$ ) and then calculate the  $\sigma_{\text{rms}}$  for the 0th-moment map as

$$\sigma_{\text{rms}} = \sigma_{\text{chn}} \Delta_{\text{chw}} \sqrt{N_{\text{chn}}/1.2}, \quad (4)$$

where  $\Delta_{\text{chw}}$  is the channel width of the observations (5.5 km s<sup>-1</sup>), and the factor 1.2 is the velocity spread approximation accounting for the fact that the channels are not independent. We then mask all pixels less than  $4\sigma_{\text{rms}}$  in the 0th-moment map.

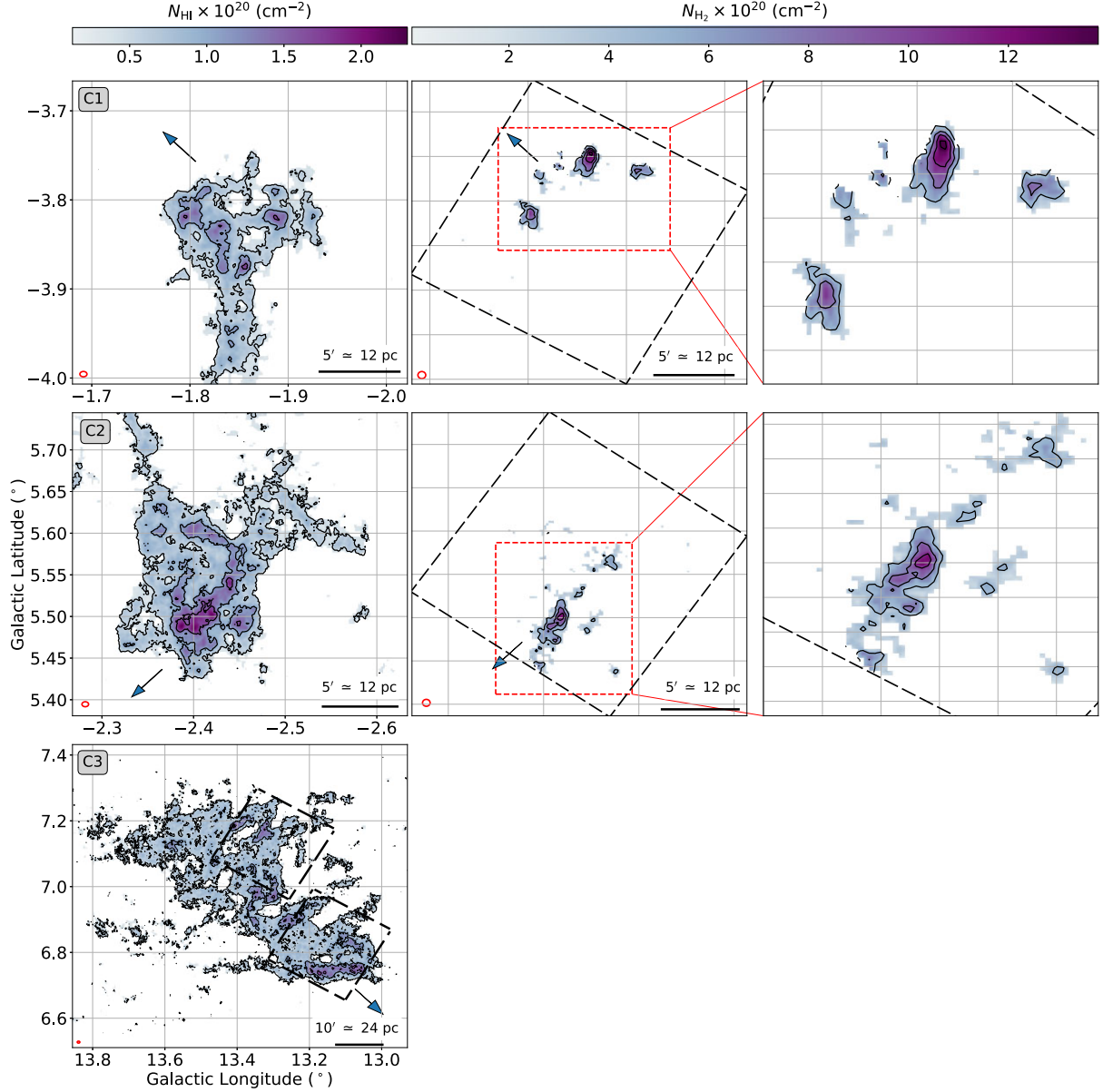
## 2.2 CO observations

We compare the MeerKAT H I data to observations of the <sup>12</sup>CO(2→1) emission line at 230.538 GHz carried out with the Atacama Pathfinder EXperiment (APEX). The CO data for C1 and C2 are taken from Di Teodoro et al. (2020), and we refer to this work for data acquisition and reduction details. C1 and C2 observations have an angular resolution of 28 arcsec with a channel width of 0.25 km s<sup>-1</sup>, covering a 15 arcmin × 15 arcmin field centred on each cloud. The rms noise ( $\sigma_{\text{chn,CO}}$ ) is 65 and 55 mK for C1 and C2, respectively, per 0.25 km s<sup>-1</sup> channel. C3 was observed as a part of a new observational programme with APEX (Di Teodoro et al., in preparation), using the new nFLASH heterodyne receiver and covering a bandwidth of 8 GHz with 61 kHz spectral channels (corresponding to 0.08 km s<sup>-1</sup>). Two 15 arcmin × 15 arcmin on-the-fly maps were observed for C3, centred on the densest regions of the H I MeerKAT emission (see Fig. 1 for an outline of the observed regions). Data reduction and imaging for C3 followed the standard procedure described in Di Teodoro et al. (2020) for C1 and C2. The final rms noise for C3 is 82 mK in a 0.25 km s<sup>-1</sup> channel with a spatial resolution of 28 arcsec. Unlike in C1 and C2, CO emission was not detected across any of the two fields of C3, not even after spectral smoothing to channel widths of 1 and 2 km s<sup>-1</sup>. Thus, C3, CO data just provide upper limits to the amount of molecular gas in the cloud.

We derive the hydrogen column densities as traced by molecular hydrogen (H<sub>2</sub>) (see Fig. 1) from the CO observations as:

$$N_{\text{H}_2} = 2 \times 10^{20} \int T_{\text{b}} dv \text{ cm}^{-2}, \quad (5)$$

where all values are the same as in equation (1) except for the  $2 \times 10^{20}$  constant, which is the CO-to-H<sub>2</sub> conversion factor,  $X_{\text{CO}}$ .



**Figure 1.** Column density maps derived from the MeerKAT  $\sim 23$  arcsec angular resolution H I observations (left) and the APEX  $\sim 28$  arcsec  $^{12}\text{CO}(2\rightarrow 1)$  observations (middle and right) following the procedure outlined in Section 2. All panels show the inferred number density of H nuclei per unit area, so the H I and CO maps can be compared directly. The top panels show C1, the middle panels show C2, and the bottom panel shows C3; C3 was not detected in CO. Red circles in the bottom left-hand corner of each map show the beam size of the observations. The black horizontal line in the bottom right-hand corner of each map shows the physical scale at an assumed distance of 8.2 kpc, while arrows point toward the projected position of the GC. The black dashed squares in the middle panels of C1 and C2 and in the map of C3 outline the field of view of the CO observations. The red dashed boxes in the middle panels of C1 and C2 outline the zoomed-in area of the right-hand panels. Contour levels for H I are  $[0.5, 1.05, 1.6] \times 10^{20}$  H nuclei  $\text{cm}^{-2}$  and contour levels for CO are  $[4.625, 7.75, 10.875, 14] \times 10^{20}$  H nuclei  $\text{cm}^{-2}$ . MeerKAT H I data have a  $4\sigma_{\text{rms}}$  mask, CO data have a  $5\sigma_{\text{rms,CO}}$  mask.

This conversion factor is unknown within the Galaxy’s nuclear wind, but Di Teodoro et al. (2020) use radiative-transfer and non-LTE excitation modelling with the DESPOTIC code (Krumholz 2014) to conclude that  $X_{\text{CO}} = 2 \times 10^{20} \text{ cm}^{-2} (\text{K km s}^{-1})^{-1}$  is likely a lower limit for the nuclear wind. Given  $N_{\text{H}_2}$ , we calculate the total mass surface density (including helium) as traced by  $\text{H}_2$  as

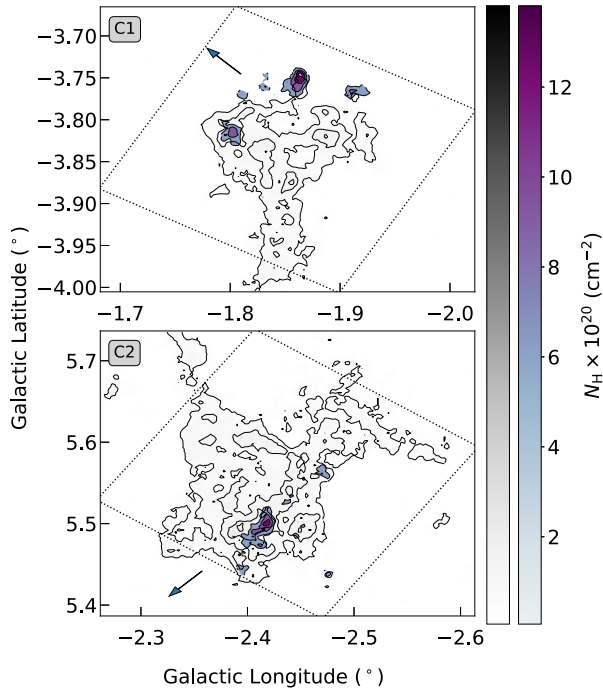
$$\Sigma_{\text{mol}} = \left( \frac{N_{\text{H}_2}}{4.59 \times 10^{19} \text{ cm}^{-2}} \right) M_{\odot} \text{ pc}^{-2}, \quad (6)$$

where the numerical factor again assumes  $X = 0.74$ , and the corresponding total mass as

$$M_{\text{mol}} = 1.36 m_{\text{H}_2} D^2 \int N_{\text{H}_2} d\Omega, \quad (7)$$

where  $m_{\text{H}_2}$  is the mass of molecular hydrogen. Note that, when comparing the H I and  $\text{H}_2$  data, to avoid confusion, we will always use either mass surface density or surface densities of H nuclei,  $N_{\text{H}} = 2N_{\text{H}_2}$ .

All CO maps were observed by dividing the  $15 \text{ arcmin} \times 15 \text{ arcmin}$  fields into nine smaller  $5 \text{ arcmin} \times 5 \text{ arcmin}$  submaps. Each submap



**Figure 2.** Same as Fig. 1, but now showing the CO data (blue-purple colours) superimposed on the H I (greyscale). The top panel shows C1, whereas the bottom panel shows C2. The contour levels are the same as those in Fig. 1. Here,  $N_H = 2N_{H_2}$ .

was observed under different weather conditions and with slightly different exposure times. As a result, the rms noise varies across each field.

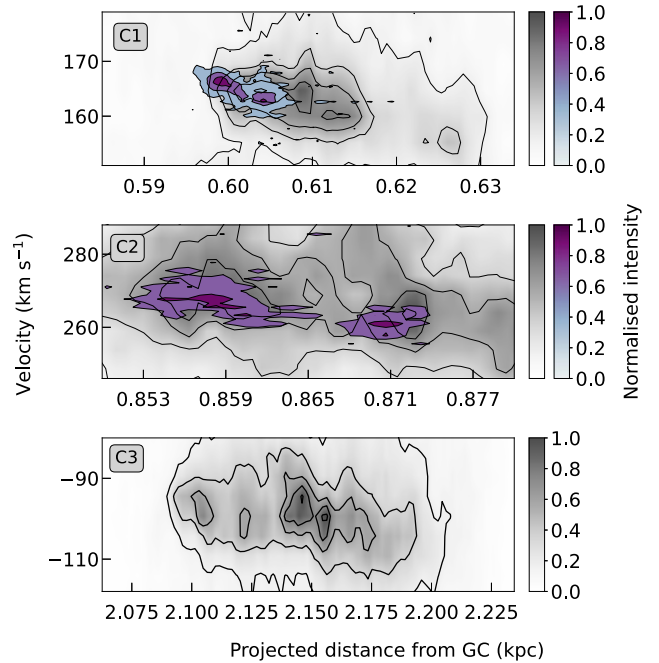
To account for this, we generate noise maps of each cube to produce an appropriate mask for the CO data. First, we calculated the rms noise of each voxel that does not contain cloud emission, producing a noise map, which is then multiplied by the channel width and the square root of the number of channels in the 0th-moment map. To create a signal-to-noise map (S/N), we divide the original moment map by the noise map. We then mask all pixels less than 5 in the S/N cube and use the masked S/N map to mask the original 0th-moment map. Note, the CO channels are independent of each other as they have been regridded from 0.08 to 0.25 km s<sup>-1</sup>.

### 3 ANALYSIS

We now analyse the H I and CO data, starting with a discussion of their relative morphologies in Section 3.1 and then proceeding to a quantitative analysis of the molecular fraction of the clouds in Section 3.2. When comparing the two data sets, we always bin the H I data to the slightly coarser spatial resolution of the CO data, but we note that, even before binning, the CO and H I data have very similar angular resolutions (28 arcsec versus 24.9 arcsec × 21.7 arcsec), so binning effects are minimal.

#### 3.1 Morphological comparison

Fig. 1 shows the column density maps of the clouds in H I and CO separately, and Fig. 2 shows the CO superimposed on the H I for C1 and C2; in the latter figure, we have rebinned the H I data to the same spatial grid as the CO. For C1, the CO emission consists of five fragmented clumps aligned almost perpendicular to the direction of



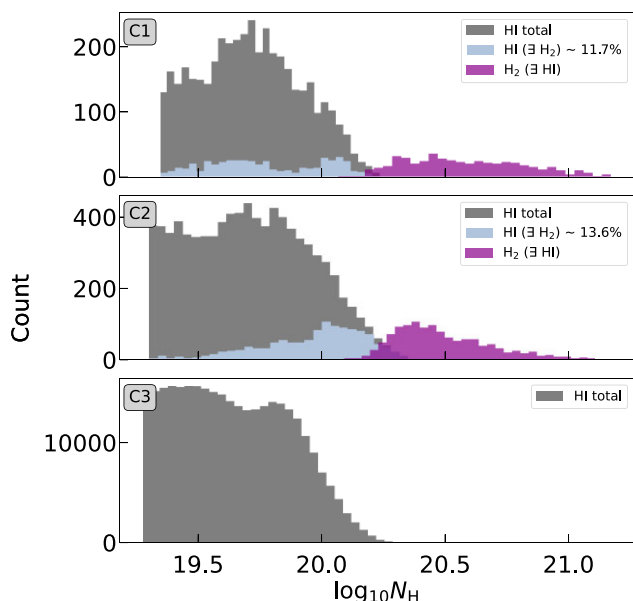
**Figure 3.** PV maps of the CO (blue-purple colours) and H I (greyscale) observations where velocity is the  $V_{LSR}$ . The top panel shows C1, the middle panel shows C2, and the bottom panel shows C3. The position is measured by projected distance from the GC, computed as discussed in Section 3.1, and we compute the intensity at a given PV point by integrating the observed position-position-velocity (PPV) cube along the direction transverse to the GC. Both H I and CO data are normalized to have a maximum of unity in the PV diagram. H I contour levels are [0.15, 0.35, 0.55, 0.75, 0.95], CO contour levels are [0.2, 0.5, 0.8, 1].

the GC. The H I trails behind the compact CO clumps as the distance from the GC increases. The H I produces a tail-like morphology, while the CO appears to be concentrated near the head of the cloud. For C2, the CO is central to the H I, aligned parallel to the direction of the GC. The CO and H I peaks nearly coincide spatially, and both species trail off in column density as the distance from the GC increases. The CO appears to be at the core of the cloud, surrounded by the H I.

We show all three clouds' position-velocity (PV) maps in Fig. 3 where velocity is the  $V_{LSR}$ . This figure measures positions in terms of projected distances from the GC, which we calculate assuming that both the cloud and the GC are at the same distance  $D = 8.2$  kpc from the Sun. To construct the PV map, we integrate all emissions in the direction transverse to the GC and normalize all intensities to a common scale so that we can easily compare the H I and CO data.

For C1, the CO lies closer to the GC than the H I, consistent with what we see in the column density maps, whereas for C2, the CO coincides spatially and spectrally with the H I. Further, Fig. 3 reveals that the velocity structures between the H I and CO are similar for C1 and C2. In both, there is a clear gradient whereby as the projected distance from the GC increases, there is a  $\sim 5$ – $10$  km s<sup>-1</sup> decrease in velocity with a  $\sim -1$  km s<sup>-1</sup> pc<sup>-1</sup> gradient. For C3, while there is no CO to compare with the H I, the H I shows a velocity gradient similar to that seen in C1 and C2.

We further investigate the cloud morphologies by comparing the probability distribution functions (PDFs) of H I and H<sub>2</sub> in Fig. 4, distinguishing between the total H I distribution and the distribution considering only those pixels for which we also detect CO. Here, we notice a significant difference between the clouds: while C1 and



**Figure 4.** Histograms of pixel counts for total H I column densities in grey, H I columns limited to pixels where there is also a detection in CO in light blue, and H<sub>2</sub> columns where there is also a detection in H I in purple. The fraction of H I-detected pixels, where there is also a CO detection, is indicated in the legend for the C1 and C2 panels. Here, H<sub>2</sub> is in terms of H nuclei,  $N_{\text{H}} = 2N_{\text{H}_2}$ .

C2 have similar column density PDFs overall, for C1 sight lines where there are detections in both H I and CO cover most of the H I column density range; except at the highest H I columns, there is little correlation between the H I column density in a pixel and the probability that it will harbour detectable CO emission. For C2, by contrast, H I pixels where we detect CO are much more heavily biased towards the highest H I columns, consistent with the visual impression in Fig. 2. For C3, there are no CO detections, and the H I column density PDF is shifted to lower values on average than those found in C1 and C2. None the less, C3 contains significant numbers of pixels at H I column densities similar to those at which C1 and C2 show significant CO detection fractions. Thus, the overall conclusion to draw from Fig. 4 is that the H I–H<sub>2</sub> relationship is substantially different in all three clouds.

### 3.2 H I-to-H<sub>2</sub> ratio and molecular fraction

We next investigate the molecular fractions [ $M_{\text{mol}}/(M_{\text{at}} + M_{\text{mol}})$ ] of the clouds by showing this fraction as a function of position in Fig. 5, and by plotting  $\Sigma_{\text{at}}$  versus  $\Sigma_{\text{at}} + \Sigma_{\text{mol}}$  pixel-by-pixel in Fig. 6.

Surface densities are derived from the masked maps as described in Section 2.1 and Section 2.2. The left-hand panels of Fig. 5 show the molecular fraction only when there is a detection in both H I and CO, while the right-hand panels are the same map but present upper limits for pixels where there is a non-detection in CO but a detection in H I. We calculate the upper limit of CO using the position-dependent masking level computed as described in Section 2.2. We see that for the two clouds containing a CO component, the molecular fraction is high ( $\sim 1$ ) over much of the area where there is a detection in both H I and CO. When considering the upper limits for CO, 99 per cent of C1 and C2 may have a molecular fraction  $\geq 50$  per cent.

Similarly, in Fig. 6 the purple triangles show only pixels when there is a detection in both CO and H I, while horizontal grey lines

show pixels where we have an H I detection but only an upper limit for CO, with the limits for H<sub>2</sub> derived from the CO noise maps; the grey lines show a range from an H<sub>2</sub> column of 0 (i.e.  $\Sigma_{\text{at}} + \Sigma_{\text{mol}} = \Sigma_{\text{at}}$ ) to a total column density derived by adding  $\Sigma_{\text{at}}$  the H<sub>2</sub> column corresponding to the level below which CO is masked; this is  $5\sigma_{\text{rms,CO}}$ . For C3, since there is no observed CO emission, we only show upper limits in the bottom panel of Fig. 6.

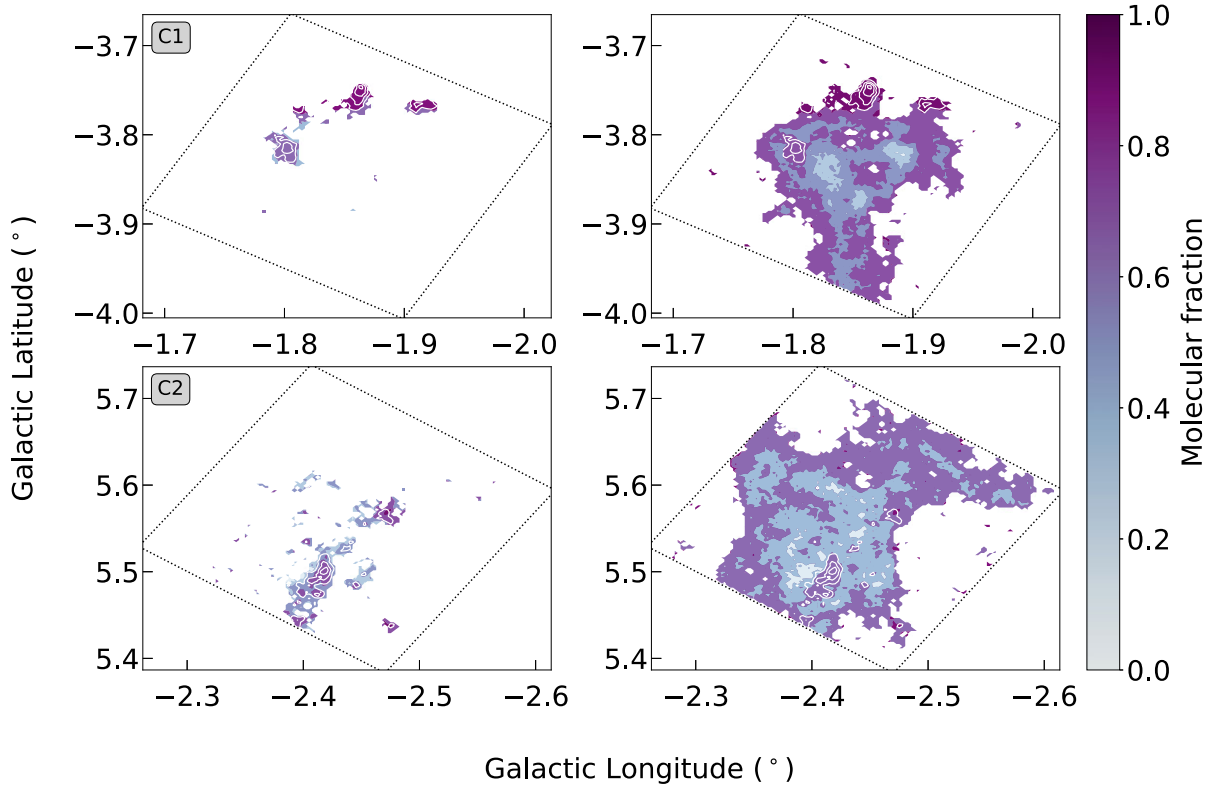
For comparison, we also show  $\Sigma_{\text{at}}$  versus  $\Sigma_{\text{at}} + \Sigma_{\text{mol}}$  for five dark and star-forming regions in the Perseus molecular cloud in the solar neighbourhood observed by Lee et al. (2015) (light blue squares), results from FUSE survey of high-latitude, interstellar H<sub>2</sub> as presented in Gillmon et al. (2006) (pink circles), results from the Copernicus satellite interstellar Lyman  $\alpha$  lines as presented in Bohlin et al. (1978) (dark blue crosses), along with the theoretical relation describing the equilibrium H I–H<sub>2</sub> phase balance (Krumholz et al. 2008, 2009; McKee & Krumholz 2010) (black line) for varying metallicities ( $Z = 0.5, 1$ , and  $2Z_{\odot}$ ). We refer to this equilibrium line as the KMT model. We also show orange lines indicating where the sightlines have molecular fractions ( $f_{\text{H}_2}$ ) of 70 per cent, 90 per cent, and 97 per cent.

For solar metallicity ( $Z = Z_{\odot}$ ), the KMT model predicts that H I converts to H<sub>2</sub>, thereby setting a maximum H I column, at  $\Sigma_{\text{at}} = 8.8 \text{ M}_{\odot} \text{ pc}^{-2}$ . The most important thing to note from this figure is that, for C1 and C2, all  $\Sigma_{\text{at}}$  values are in the range  $\sim 0.3\text{--}3 \text{ M}_{\odot} \text{ pc}^{-2}$ , well below the  $\sim 10 \text{ M}_{\odot} \text{ pc}^{-2}$  level that the KMT model predicts is required to shield molecular gas against photodissociation for solar metallicity, and which is observed in the Perseus cloud, the high-latitude sightlines, and the interstellar Lyman  $\alpha$  lines. Despite this, we detect appreciable columns of H<sub>2</sub>, with H<sub>2</sub> dominating the total column on at least some lines of sight.

Thus, the fraction of H<sub>2</sub> relative to H I in these clouds is very different from the atomic to molecular fraction observed in the Galactic plane, where  $\sim 1 \text{ M}_{\odot} \text{ pc}^{-2}$  columns of H I are essentially never observed to host significant quantities of H<sub>2</sub>, as shown in Fig. 6. The presence of molecular material with low columns of H I implies that  $N_{\text{HI}} \approx 20^{21} \text{ cm}^{-2}$  is not the sole determinant of the presence of H<sub>2</sub> as expected for gas in chemical equilibrium (Krumholz, McKee & Tumlinson 2009) and as observed (Liszt & Gerin 2023) in the Galaxy.

We emphasize that this result is *not* dependent on our (highly uncertain) choice of  $X_{\text{CO}}$ . Using a different value of  $X_{\text{CO}}$  would shift the cloud of observed points left or right in Fig. 6, and thus could not bring the data close to the locus occupied by the Perseus, Gillmon et al. (2006), or Bohlin et al. (1978) points. Moreover, given that many of the points have  $f_{\text{H}_2} > 90$  per cent, shifting them far enough to the left so that they lie close to the KMT model line (indicating a purely atomic composition) and the H I-dominated FUSE and Copernicus points would require a value of  $X_{\text{CO}}$  more than an order of magnitude smaller than our fiducial estimate, and close to the absolute limit for optically thin CO emission  $X_{\text{CO}} \approx 10^{19} \text{ cm}^{-2}/(\text{K km s}^{-1})$  (Bolatto, Wolfire & Leroy 2013).

Nor could the result be explained by GC gas having super-solar metallicity. Metallicity clearly does affect the H I column at which the transition to H<sub>2</sub> occurs as demonstrated by the Bohlin et al. (1978) points which have slightly subsolar metallicities, accounting for the offset from the  $Z = Z_{\odot}$  line. However, the KMT model predicts that, in order to form H<sub>2</sub> at H I column densities of  $\sim 1 \text{ M}_{\odot} \text{ pc}^{-2}$  as we observe, the dust abundance would have to be  $\approx 10 \times$  the solar neighbourhood value, implausibly high, and even in the most metal-rich external galaxies the H I to H<sub>2</sub> transition is observed to begin only at H I surface densities  $\sim 5 \text{ M}_{\odot} \text{ pc}^{-2}$  (Wong et al. 2013; Schruba, Bialy & Sternberg 2018). For a realistic GC metallicity of  $Z \lesssim 2Z_{\odot}$ ,



**Figure 5.** Molecular fraction maps with C1 in the upper panels and C2 in the lower panels. The left-hand panels show the molecular fraction only when there is a detection in both H I and CO, while the right-hand panels also include upper limits on the molecular fraction derived from upper limits on the CO in those pixels where CO is not detected but H I is; the CO upper limits correspond to our masking level of five times the (position-dependent) rms noise level in each pixel. White contours show the CO data at the same contour levels as those used in Fig. 1. The black dotted box is the CO observation field of view.

the observed points still lie far from the equilibrium KMT model prediction.

## 4 DISCUSSION

In this discussion, we investigate the likely origins of both atomic and molecular gas entrained in the MW’s nuclear wind by considering the morphologies of the two observed species in Section 4.1 and the chemical states of each cloud in Section 4.2. We finish this section by considering the lifetime of a cloud in the GWs in Section 4.3.

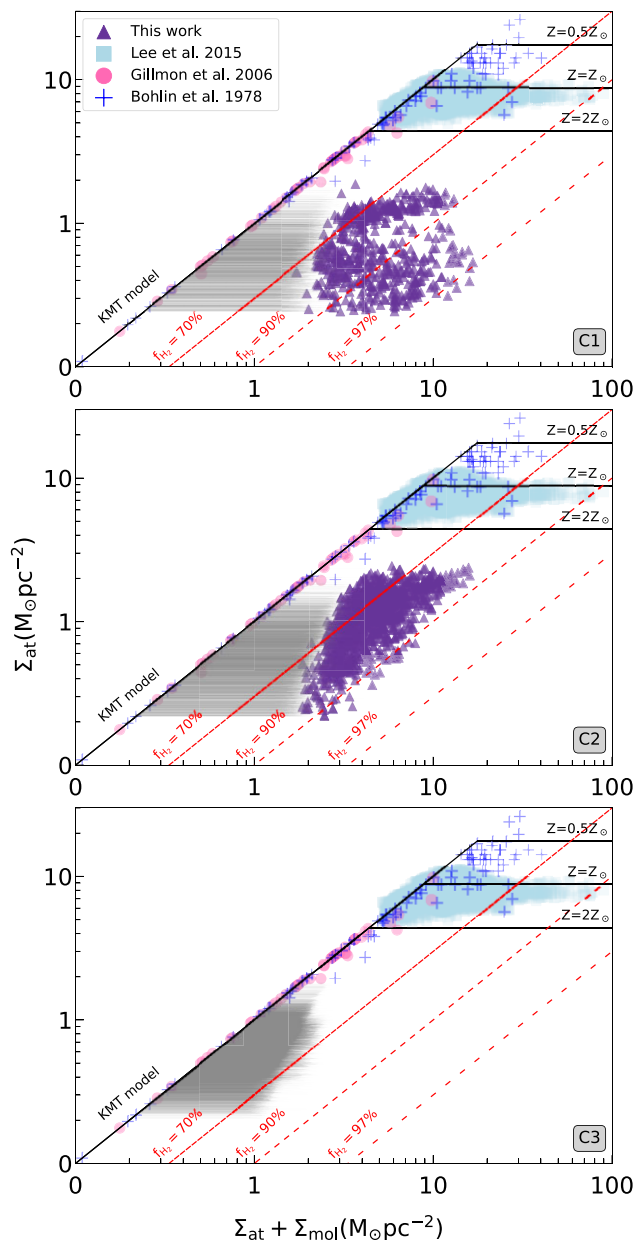
### 4.1 Morphologies

The two clouds where we detect both H I and CO show contrasting morphologies, potentially suggesting we observe clouds of different origins or, as suggested in Di Teodoro et al. (2020), different evolutionary stages of the same process. If the molecular gas is disassociating into atomic hydrogen, we could expect to see the H I surround the molecular gas as we see in molecular clouds in the disc (e.g. Stanimirović et al. 2014). However, it is also conceivable that in the presence of a strong wind, the H I could trail behind the H<sub>2</sub> due to stripping effects – that is, H I may ‘boil’ off the surface of the molecular cloud but then be swept back by the ram pressure of the unseen hot gas that drives the wind. We see both of these

morphologies across C1 and C2 (see Figs 2 and 4).<sup>2</sup> In the early stages of cloud evaporation, H I would be a surface phenomenon unrelated to the depth of the cloud, as appears to be the case for C1 in Fig. 4. Later on, when the disassociation front has broken up much of the H<sub>2</sub>, fresh H I would continue to be produced from the remaining dense H<sub>2</sub> regions, as appears to be the case for C2 in Fig. 4. Finally, in the late stages of cloud evaporation, all H<sub>2</sub> would have dissociated into H I, as appears to be the case in C3.

Alternatively, if the H I is compressing into molecular gas, we expect this to occur at the central, densest regions of the cloud. As the molecular gas is positioned along the inner regions of C2 and towards the GC with H I trailing behind in C1, either scenario (or a combination of the two) could be taking place. Given the lack of molecular matter in C3 and its greater distance from the GC compared to C1 and C2 (Fig. 3), it is possible it once had a molecular component, but it has since all dissociated into H I. Lastly, when considering the third scenario presented in Section 1, it is difficult to predict the expected geometry for a cool cloud that has condensed out of the rapidly cooling hot phase of GWs. Thus, a cloud’s morphology alone cannot disentangle these scenarios.

<sup>2</sup>The difference between the two clouds may also be a result of projection effects: Di Teodoro et al. (2018) present a kinematic model of the wind that predicts that C1 lies almost directly transverse to the GC from our perspective, while C2 is a similar projected distance from the GC but is several kpc closer to us in the radial direction. Thus, the radial vector from each cloud to the GC lies nearly in the plane of the sky for C1, but is significantly tilted relative to the plane of the sky for C2.



**Figure 6.**  $\Sigma_{\text{at}}$  versus  $\Sigma_{\text{at}} + \Sigma_{\text{mol}}$  for C1 (top), C2 (middle), and C3 (bottom). Purple triangles are detections in both H I and CO, while the horizontal grey lines are the upper and lower limits of  $\Sigma_{\text{at}} + \Sigma_{\text{mol}}$  derived by taking the minimum  $\text{H}_2$  column to be 0 and the maximum to be the  $\text{H}_2$  column corresponding to the CO detection threshold; this is  $5\sigma_{\text{rms,CO}}$ . Black solid lines show the H I– $\text{H}_2$  relation predicted by the Krumholz–McKee–Tumlinson (KMT) model (Krumholz et al. 2008, 2009; McKee & Krumholz 2010) for half-solar, solar, and double solar metallicities; the diagonal portion of this line has slope unity and corresponds to a purely H I composition, with no  $\text{H}_2$ . Red, dashed lines show different molecular fractions ranging from  $f_{\text{H}_2} = 70$  per cent–97 per cent. Light blue squares show measurements from Lee et al. (2015) for five dark and star-forming regions in the Perseus molecular cloud in the MW plane, pink circles show results from the Far Ultraviolet Spectroscopic Explorer (FUSE) survey of high-latitude  $\text{H}_2$  regions from Gillmon et al. (2006) and dark blue crosses show results from Copernicus observations (Bohlin, Savage & Drake 1978).

## 4.2 Chemical state

The chemical state of the clouds is more informative. As discussed in Section 3.2 and shown in Fig. 5, there are significant portions of C1 and C2 by both mass and area where the molecular fraction is close to 1, and considering our sensitivity limits, there may be additional undetected regions of moderately molecular ( $f_{\text{H}_2} \sim 0.3\text{--}0.5$ ) material. More importantly, as shown in Fig. 6, this high  $\text{H}_2$  fraction is greater than we would expect given the low H I column densities, based on either theoretical models or on the measured H I– $\text{H}_2$  relationship in the Galactic plane. While high molecular fractions in CO-detected pixels are not surprising since low levels of CO column densities are either undetectable or have been masked, it is none the less significant that we detect CO at all, indicating the presence of considerable columns of  $\text{H}_2$  at levels of H I where normally there would not be any (e.g. Krumholz et al. 2009; Liszt & Gerin 2023).

This overabundance in  $\text{H}_2$  makes a compression-driven  $\text{H}_2$  formation scenario unlikely. While it is easy to understand how the  $\text{H}_2$  chemistry could be out of equilibrium, given the relatively long equilibration time-scales (a point to which we return momentarily), if the system were transitioning from H I-dominated to beginning to form  $\text{H}_2$ , we would expect the system to be out of equilibrium in the direction of too little  $\text{H}_2$ , not too much as we observe. Quantitatively, for evolutionary times  $t$  smaller than the equilibrium time-scale  $t_{\text{eq}}$ , we expect the molecular fraction  $\Sigma_{\text{mol}}/(\Sigma_{\text{at}} + \Sigma_{\text{mol}})$  to be intermediate between the initial state which exists at  $t \ll t_{\text{eq}}$  and the equilibrium state achieved at  $t \gg t_{\text{eq}}$ . Thus, if the initial state contains no significant amount of  $\text{H}_2$ ,  $\Sigma_{\text{mol}}/(\Sigma_{\text{at}} + \Sigma_{\text{mol}}) \approx 0$ , and the final equilibrium has a non-zero  $\text{H}_2$  fraction, as posited in the compression scenario, then we would expect that at all times the value of  $\Sigma_{\text{mol}}/(\Sigma_{\text{at}} + \Sigma_{\text{mol}})$  would be intermediate between zero and the equilibrium value; at no point would we expect to find the system in a state where it *exceeds* the equilibrium value, which is the state that we actually observe. The same argument implies that it is also unlikely that the clouds are a result of condensation of cool material out of a hot phase.

By contrast, the observed H I surface densities and overabundance of  $\text{H}_2$  are consistent with our first scenario, where a stripping of H I from the envelope of a pre-existing molecular cloud leaves behind too little H I to shield the  $\text{H}_2$  from the Galaxy’s dissociating radiation field, and the  $\text{H}_2$  is in the process of converting to H I. This scenario also explains the non-detection of CO emission from C3: this cloud is significantly further from the GC in projection than C1 or C2, and thus has likely had more time to undergo photodissociation than the other two clouds. The photodissociation process has been completed for this cloud, leaving behind only H I. Consistent with this picture, in Fig. 6 we also observe that the cloud closest to the disc (C1) has a wider spread of points in the direction below and to the right of the equilibrium line, implying a higher  $\text{H}_2$  abundance, while C2 (which is further from the disc) has points that, while still indicating significant amounts of  $\text{H}_2$ , lie closer to the line. In other words, as clouds move further from the disc, the points on the  $\Sigma_{\text{at}}$  versus  $\Sigma_{\text{at}} + \Sigma_{\text{mol}}$  plot migrate closer towards the theoretical equilibrium line as the clouds’ molecular components dissociate.

## 4.3 The lifetime of molecular clouds in the Milky Way’s nuclear wind

In the dissociation scenario, we can estimate the lifetimes of molecular clouds in the wind from theory and check that the resulting

predictions are consistent with the time-scales inferred from cloud kinematics. For the theoretical estimate, we follow the general approach of the KMT model (Krumholz et al. 2008, 2009; McKee & Krumholz 2010)<sup>3</sup>: we first note that the flux of dissociating Lyman–Werner band photons on the surface of a cloud exposed to the full, unshielded interstellar radiation field (ISRF) is  $F^* = 2.1 \times 10^7 \chi$  photons  $\text{cm}^{-2} \text{s}^{-1}$ , where  $\chi$  is the ISRF strength normalized to the solar neighbourhood value. If this flux falls on an optically thick cloud of fully molecular gas, the very large ratio of the  $\text{H}_2$  opacity to the dust opacity ensures that almost all the photons will be absorbed by  $\text{H}_2$  molecules, and a fraction  $f_{\text{diss}} \approx 0.1$  of these absorptions will result in the dissociation of the molecule. This will convert  $\text{H}_2$ -dominated material to  $\text{H I}$ -dominated at a rate (expressed as mass converted per unit area per unit time)

$$\dot{\Sigma}_{\text{mol} \rightarrow \text{at}} = 2 f_{\text{diss}} F^* \frac{m_{\text{H}}}{X} = 1.4 \chi \text{ M}_{\odot} \text{ pc}^{-2} \text{ Myr}^{-1}, \quad (8)$$

where the factor of two accounts for there being two H nuclei per  $\text{H}_2$  molecule. Thus, the characteristic time required to dissociate a molecular cloud of initial surface density  $\Sigma_{\text{H}_2}$  is

$$t_{\text{diss}} = \frac{\Sigma_{\text{mol}}}{\dot{\Sigma}_{\text{mol} \rightarrow \text{at}}} = 7.0 \left( \frac{\Sigma_{\text{mol}} / \chi}{10 \text{ M}_{\odot} \text{ pc}^{-2}} \right) \text{ Myr}. \quad (9)$$

This estimate carries significant uncertainties – most obviously the strength of the ISRF  $\chi$ , but also that, as dissociation proceeds and a shielding layer of  $\text{H I}$ -dominated material builds up, the conversion rate will fall because dissociations will be offset by  $\text{H}_2$  formation in the shielding layer and because potentially dissociating photons will be absorbed by dust grains while they transit the shielding zone. The actual dissociation time will depend on the rate at which hydrodynamic processes strip away this shielding layer, exposing more unshielded  $\text{H}_2$ ; in the absence of such stripping, an  $\text{H}_2$  cloud with a surface density  $\Sigma_{\text{mol}} \gtrsim 10 \text{ M}_{\odot} \text{ pc}^{-2}$  will eventually produce a shielding layer that will fully balance dissociation, and the remaining  $\text{H}_2$  will survive indefinitely. None the less, this rough estimate suggests that the characteristic survival time of  $\text{H}_2$ -dominated clouds whose shielding layers are being removed is a few to  $\sim 10$  Myr.

This time-scale lines up well with the kinematically estimated ages of our observed clouds. Applying the model of Lockman et al. (2020) to our three clouds yields kinematic age estimates of  $\approx 3.5$ ,  $\approx 7$ , and  $\approx 8.3$  Myr with corresponding distances from the GC of 0.8, 1.8, and 2.1 kpc for C1, C2, and C3, respectively. These ages are entirely consistent with our theoretical expectation that molecular clouds injected into the wind without adequate shielding should retain significant  $\text{H}_2$  fractions for times of a few to 10 Myr, and favour a picture where C1 and C2 are out of equilibrium and are partway through the process of converting to  $\text{H I}$ , while C3, which is somewhat older, is further along the path towards an equilibrium chemical configuration with negligible  $\text{H}_2$ . Thus, observations as well

as theory support a  $\sim$ few – 10 Myr lifetime for molecular clouds in the wind.

## 5 CONCLUSIONS

We present new high-resolution MeerKAT observations of three atomic clouds entrained in the MW GW. We combine these with APEX CO observations at closely matched resolutions to characterize the clouds' molecular content. We use these data to compare the atomic and molecular morphologies, and to determine spatially resolved molecular fractions. Of the three clouds, the one closest to the GC in projection (C1) has its molecular component positioned towards the GC and the atomic component trailing behind, the intermediate-distance cloud (C2) has its molecular component positioned within the body of the  $\text{H I}$  near the  $\text{H I}$  column density maximum, and the cloud farthest from the GC (C3) contains no detected molecular component at all. All three clouds show a velocity gradient of  $\sim -1 \text{ km s}^{-1} \text{ pc}^{-1}$  with distance from the GC. We find molecular fractions by mass in C1 and C2 that are quite high,  $\gtrsim 50$  per cent and  $\gtrsim 40$  per cent, respectively. In many pixels where we detect  $\text{H}_2$ , molecular material forms a significant fraction of the total.

Most importantly, we find that the  $\text{H I}$ – $\text{H}_2$  chemistry is out of equilibrium, and that, as a result, the GC clouds lie in a portion of  $\text{H I}$ – $\text{H}_2$  parameter space that is unpopulated in the Galactic disc. The  $\text{H I}$  surface density  $\Sigma_{\text{at}}$  is  $0.3$ – $3 \text{ M}_{\odot} \text{ pc}^{-2}$  despite the presence of substantial quantities of  $\text{H}_2$ , much lower than  $\sim 10 \text{ M}_{\odot} \text{ pc}^{-2}$   $\text{H I}$  columns observed around molecular clouds in the disc (e.g. Lee et al. 2015; Liszt & Gerin 2023), and suggested to be necessary to shield  $\text{H}_2$  by the KMT model (Krumholz et al. 2008, 2009; McKee & Krumholz 2010). This implies that there is insufficient  $\text{H I}$  column density to shield the molecular gas against dissociation. As a result of this disequilibrium, we conclude that the three clouds likely originate from molecular clouds in the Galactic disc that were entrained in the Galactic wind and are photodissociated into atomic hydrogen by the radiation field in the winds. This picture is also consistent with the non-detection of C3 in CO, since this is the dynamically oldest of the three clouds and has had the longest to photodissociate. Comparing the kinematic ages with theoretical predictions for the  $\text{H}_2$  dissociation rate, we conclude that the data are consistent with a picture where the wind consists primarily of molecular gas entrained into the outflow, which has a characteristic survival time of  $\sim 10$  Myr after it is entrained before it dissociates into  $\text{H I}$ .

## ACKNOWLEDGEMENTS

The MeerKAT telescope is operated by the South African Radio Astronomy Observatory, which is a facility of the National Research Foundation, an agency of the Department of Science and Innovation. The authors acknowledge and thank Min-Young Lee and Snežana Stanimirović for providing the data from Lee et al. (2015) in electronic form. MRK acknowledges support from the Australian Research Council through Laureate Fellowship FL220100020. EMD was supported by the European Research Council (ERC) under grant agreement no. 10104075. NM-G is the recipient of an Australian Research Council Australian Laureate Fellowship (project number FL210100039) funded by the Australian Government. The MeerKAT telescope is operated by the South African Radio Astronomy Observatory, which is a facility of the National Research Foundation, an agency of the Department of Science and Innovation. The Green Bank Observatory and the Green Bank Telescope are facilities

<sup>3</sup>We note that in addition to the photodissociation process we consider here, in principle,  $\text{H}_2$  could be collisionally dissociated as well; this could occur if either turbulent mixing or conduction raise the temperature in the  $\text{H}_2$  to several thousand K, hot enough for collisional dissociation to be significant. However, this seems unlikely to contribute to the  $\text{H I}$ – $\text{H}_2$  balance directly because the similarity between the  $\text{H}_2$  binding energy (4.5 eV) and the  $\text{H I}$  ionization energy (13.6 eV) means that there is only a very narrow range of temperature ( $\approx 0.5$ – $1.5 \times 10^4$  K), where collisions will dissociate  $\text{H}_2$  but not also ionize  $\text{H I}$ . Below this range, collisional dissociation is unimportant, and above it, collisions will rapidly convert both  $\text{H}_2$  and  $\text{H I}$  to  $\text{H}^+$ . Thus, for collisions to be an effective means of converting  $\text{H}_2$  to  $\text{H I}$  requires fine-tuning of the temperature.

of the U.S. National Science Foundation, operated by Associated Universities, Inc.

## DATA AVAILABILITY

The data underlying this article along with a general implementation of the code used to process the data cubes and produce all of the figures in this paper are available via Zenodo at <https://doi.org/10.5281/zenodo.8060960>.

## REFERENCES

- Andersson B. G., Roger R. S., Wannier P. G., 1992, *A&A*, 260, 355
- Bland-Hawthorn J., Cohen M., 2003, *ApJ*, 582, 246
- 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
- Bregman J. N., 1980, *ApJ*, 237, 280
- Carretti E. et al., 2013, *Nature*, 493, 66
- Cashman F. H. et al., 2021, *ApJ*, 923, L11
- Crocker R. M., Bicknell G. V., Taylor A. M., Carretti E., 2015, *ApJ*, 808, 107
- Di Teodoro E. M., McClure-Griffiths N. M., Lockman F. J., Denbo S. R., Endsley R., Ford H. A., Harrington K., 2018, *ApJ*, 855, 33
- Di Teodoro E. M., McClure-Griffiths N. M., Lockman F. J., Armillotta L., 2020, *Nature*, 584, 364
- Dickey J. M., Lockman F. J., 1990, *ARA&A*, 28, 215
- Dobler G., Finkbeiner D. P., Cholis I., Slatyer T., Weiner N., 2010, *ApJ*, 717, 825
- Espada D. et al., 2010, *ApJ*, 720, 666
- Fielding D. B., Bryan G. L., 2022, *ApJ*, 924, 82
- GRAVITY Collaboration, 2019, *A&A*, 625, L10
- Gillmon K., Shull J. M., Tumlinson J., Danforth C., 2006, *ApJ*, 636, 891
- Ginsburg A. et al., 2015, in Iono D., Tatematsu K., Wootten A., Testi L., eds, ASP Conf. Ser. Vol. 499, *Revolution in Astronomy with ALMA: The Third Year*. Astron. Soc. Pac., San Francisco, p. 363
- Greve A., 2004, *A&A*, 416, 67
- Huang S., Katz N., Scannapieco E., Cottle J., Davé R., Weinberg D. H., Peebles M. S., Brüggemann M., 2020, *MNRAS*, 497, 2586
- Jonas J., 2016, *Proc. Sci.*, Vol. 277, *MeerKAT Science: On the Pathway to the SKA*. Sissa, Trieste, p. 1
- Kanjilal V., Dutta A., Sharma P., 2021, *MNRAS*, 501, 1143
- Koyama K., Awaki H., Kunieda H., Takano S., Tawara Y., 1989, *Nature*, 339, 603
- Krumholz M. R., 2014, *MNRAS*, 437, 1662
- Krumholz M. R., McKee C. F., Tumlinson J., 2008, *ApJ*, 689, 865
- Krumholz M. R., McKee C. F., Tumlinson J., 2009, *ApJ*, 693, 216
- Lee M.-Y., Stanimirović S., Murray C. E., Heiles C., Miller J., 2015, *ApJ*, 809, 56
- Leroy A. K. et al., 2015, *ApJ*, 814, 83
- Liszt H., Gerin M., 2023, *ApJ*, 943, 172
- Lockman F. J., 1984, *ApJ*, 283, 90
- Lockman F. J., McClure-Griffiths N. M., 2016, *ApJ*, 826, 215
- Lockman F. J., Di Teodoro E. M., McClure-Griffiths N. M., 2020, *ApJ*, 888, 51
- Maccagni F. M. et al., 2021, *A&A*, 656, A45
- McClure-Griffiths N. M., Green J. A., Hill A. S., Lockman F. J., Dickey J. M., Gaensler B. M., Green A. J., 2013, *ApJ*, 770, L4
- McCourt M., O’Leary R. M., Madigan A.-M., Quataert E., 2015, *MNRAS*, 449, 2
- McKee C. F., Krumholz M. R., 2010, *ApJ*, 709, 308
- McMullin J. P., Waters B., Schiebel D., Young W., Golap K., 2007, in Shaw R. A., Hill F., Bell D. J., eds, ASP Conf. Ser. Vol. 376, *Astronomical Data Analysis Software and Systems XVI*. Astron. Soc. Pac., San Francisco, p. 127
- Martini P., Leroy A. K., Mangum J. G., Bolatto A., Keating K. M., Sandstrom K., Walter F., 2018, *ApJ*, 856, 61
- Morganti R., Oosterloo T., Tsvetanov Z., 1998, *AJ*, 115, 915
- Saito T. et al., 2022a, *ApJ*, 927, L32
- Saito T. et al., 2022b, *ApJ*, 935, 155
- Schneider E. E., Robertson B. E., 2017, *ApJ*, 834, 144
- Schneider E. E., Robertson B. E., Thompson T. A., 2018, *ApJ*, 862, 56
- Schneider E. E., Ostriker E. C., Robertson B. E., Thompson T. A., 2020, *ApJ*, 895, 43
- Schruba A., Bialy S., Sternberg A., 2018, *ApJ*, 862, 110
- Seaquist E. R., Clark J., 2001, *ApJ*, 552, 133
- Simonson S. C., III, 1973, *A&A*, 23, 19
- Snowden S. L. et al., 1997, *ApJ*, 485, 125
- Sofue Y., Handa T., 1984, *Nature*, 310, 568
- Stanimirović S., Murray C. E., Lee M.-Y., Heiles C., Miller J., 2014, *ApJ*, 793, 132
- Su M., Slatyer T. R., Finkbeiner D. P., 2010, *ApJ*, 724, 1044
- Thompson T. A., Quataert E., Zhang D., Weinberg D. H., 2016, *MNRAS*, 455, 1830
- Veilleux S., Maiolino R., Bolatto A. D., Aalto S., 2020, *A&AR*, 28, 2
- Wang G. C. P. et al., 2021, *MNRAS*, 508, 3754
- Wannier P. G., Lichten S. M., Morris M., 1983, *ApJ*, 268, 727
- Werk J. K. et al., 2014, *ApJ*, 792, 8
- Wong T. et al., 2013, *ApJ*, 777, L4
- Yang H. Y. K., Ruszkowski M., Zweibel E. G., 2022, *Nat. Astron.*, 6, 584
- Yuan Y., Krumholz M. R., Martin C. L., 2022, *MNRAS*, 518, 4084

## APPENDIX A: COMPARISON OF MEERKAT DATA WITH GREEN BANK TELESCOPE

Here, we compare the MeerKAT data for C1 and C2 to those obtained with the GBT as presented in Di Teodoro et al. (2018). To conduct the comparison, we first rebin all data sets to the same spatial and spectral bins and only consider data within the same velocity width ( $143\text{--}176\text{ km s}^{-1}$  for C1 and  $242\text{--}319\text{ km s}^{-1}$  for C2) and spatial area  $[(l, b) = (-1.76^\circ \text{ to } -1.96^\circ, -3.73^\circ \text{ to } -3.93^\circ)$  for C1 and  $(l, b) = (-2.36^\circ \text{ to } -2.46^\circ, 5.49^\circ \text{ to } 5.59^\circ)$  for C2. We then smooth the MeerKAT resolution to that of the GBT (569 arcsec) and compare the mass contained within the above-described areas. Our MeerKAT mass results are within 1 per cent of the GBT mass results for both clouds. The differences in HI mass reported in Table 1 compared to those from Di Teodoro et al. (2020) are due to the undefined nature of the clouds, with neither possessing distinct edges. As such, mass estimates depend on the spectral and spatial area over which the clouds’ masses are calculated. The fact that the masses are consistent when these calculations are performed over the same PPV volume indicates that, despite the lack of short spacings, there is no significant missing flux in the MeerKAT data.

## APPENDIX B: DISTANCE ESTIMATES

In this section, we consider how distance measurements affect mass estimates. To calculate the mass of a cloud, we require a distance measurement as described in equations (3) and (7). The results presented in Table 1 use a distance of 8.2 kpc (the distance from the Sun to the GC; GRAVITY Collaboration 2019) for all clouds. However, Di Teodoro et al. (2020) determine distances from the GC for each cloud using a kinematic model of the MW GC wind. Using these distances, we calculate the masses of each cloud as shown in Table B1.

**Table B1.** Derived masses of observed clouds. Distance is the cloud's distance from the Sun measured in kpc as determined by a kinematic wind model as described in Di Teodoro et al. (2020). Atomic mass,  $M_{\text{at}}$ , is the derived mass as traced by the H I gas in solar masses from equation (3) using the distances in the distance column. Molecular mass,  $M_{\text{mol}}$ , is the derived mass as traced by the CO in solar masses from equation (7) using the distances in the distance column.

| Cloud ID | Distance<br>(kpc) | $M_{\text{at}}$<br>( $M_{\odot}$ ) | $M_{\text{mol}}$<br>( $M_{\odot}$ ) |
|----------|-------------------|------------------------------------|-------------------------------------|
| C1       | 8.7               | 385                                | 443                                 |
| C2       | 9.7               | 1009                               | 799                                 |
| C3       | 8.3               | 3426                               | –                                   |

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