

Turning JWST/MIRI backgrounds into a survey of diffuse molecular hydrogen

E. Nigou^{1,2,*}, B. Godard^{1,2}, P. Guillard³, G. Pineau des Forêts^{2,4},
M. A. Miville-Deschênes^{1,2}, and P. Lesaffre^{1,2}

¹ Laboratoire de Physique de l'École Normale Supérieure, ENS, Université PSL, CNRS, Sorbonne Université, 75005 Paris, France

² LUX, Observatoire de Paris, Université PSL, Sorbonne Université, CNRS, 75014 Paris, France

³ Institut d'Astrophysique de Paris, Sorbonne Université, CNRS, 98 bis bd Arago, 75014 Paris, France

⁴ Institut d'Astrophysique Spatiale, Université Paris-Saclay, CNRS, 91405 Orsay, France

Received 22 January 2026 / Accepted 13 March 2026

ABSTRACT

Context. A statistically significant sampling of H₂ rotational excitation in the diffuse interstellar medium (ISM) is essential to identifying its excitation mechanisms and assessing the importance of H₂ in the cooling of the gas and the regulation of thermal pressure.

Aims. To complement the statistics provided by ancillary telescopes, we conducted a search for pure rotational H₂ emission lines in all publicly available background observations obtained with the Medium Resolution Spectrometer (MRS) aboard the JWST.

Methods. The sample consists of 276 background observations acquired over the past three years. Departing from the standard pipeline, each uncalibrated MRS background file was reprocessed, enabling the analysis of H₂ pure rotational emission. Lines of sight likely associated with star-forming complexes were excluded to focus on emission from the diffuse ISM. The results were compared with FUSE absorption data and were analyzed in relation to the column densities of H and H₂ and to dust emission derived from HI4PI, *Planck*, and WISE data.

Results. This analysis reveals widespread H₂ emission throughout the Galaxy. We report the first detections of the pure rotational S(4), S(5), and S(7) lines in the diffuse ISM. The S(1) line is detected along 84 lines of sight, corresponding to a detection rate of 41%. Its integrated intensity decreases steeply with Galactic latitude, spanning nearly two orders of magnitude, in remarkable agreement with absorption measurements. The T_{34} and T_{35} excitation temperatures vary between 200 and ~1000 K, are correlated with each other, and are anticorrelated with the column density of H₂, as expected from ancillary data. All lines of sight in the sample have undergone the H–H₂ transition, at $N_{\text{H}} \gtrsim 10^{20} \text{ cm}^{-2}$, and are partly molecular, with $f_{\text{H}_2} \gtrsim 0.1$. Under these conditions, the cooling rate associated with the S(1) line, expressed per hydrogen atom, is found to be remarkably constant, with a characteristic value of $\sim 4 \times 10^{-27} \text{ erg s}^{-1} \text{ H}^{-1}$.

Conclusions. This study demonstrates that the high sensitivity of the JWST enables measurements that both strengthen and complement those from absorption studies. Observations collected over just a fraction of JWST's lifetime have already yielded detections along dozens of lines of sight, significantly expanding the statistical sample of H₂ rotational excitation in the diffuse ISM.

Key words. methods: data analysis – methods: observational – techniques: spectroscopic – ISM: lines and bands – ISM: molecules

1. Introduction

Understanding the formation, survival, and excitation of molecular hydrogen is a long-standing objective in studies of the interstellar medium (ISM). First, H₂ surface density correlates with the star formation rate (e.g., Lin et al. 2019) and provides a valuable tracer for understanding the Kennicutt–Schmidt relation and its spatial variations in galaxies (e.g., Leroy et al. 2013; Pessa et al. 2022). Second, as the most abundant molecule in the Universe, H₂ initiates interstellar chemistry and the formation of complex species in space. Last, while the cooling of the two stable phases of the diffuse ISM, the warm neutral medium (WNM) and the cold neutral medium (CNM), is presumably dominated by atomic species (Wolfire et al. 2003), H₂ plays a subtle but essential role in cooling the gas. By preventing the CNM from evaporating into the WNM, H₂ cooling suppresses thermal instability and regulates the thermal pressure of the diffuse gas and, hence, the relative distributions of the WNM and CNM (Godard et al. 2023).

The first detection of molecular hydrogen in the local diffuse ISM was obtained with UV rocket observations (Carruthers 1970) and was later confirmed and extensively characterized by the *Copernicus* satellite via resonant electronic absorption lines of H₂ toward nearby OB stars (e.g., Spitzer & Jenkins 1975; Savage et al. 1977; Shull & Beckwith 1982). These early studies were later expanded upon using measurements by the Far Ultraviolet Spectroscopic Explorer (FUSE), which targeted more distant and fainter sources both in the Galactic disk (e.g., Rachford et al. 2002; Shull et al. 2021) and at higher Galactic latitudes toward extragalactic background sources (e.g., Gillmon et al. 2006; Gillmon & Shull 2006; Wakker 2006). The first detection of H₂ pure rotational emission from diffuse Galactic regions came with the Infrared Space Observatory (ISO), which measured the S(0) to S(3) lines along a line of sight across the Galactic midplane (Falgarone et al. 2005). This discovery was later extended with *Spitzer* observations of the first pure rotational lines of H₂ in emission toward two translucent clouds at high Galactic latitudes associated with infrared cirrus emission (Ingalls et al. 2011).

On the one hand, observations performed in absorption have provided access to the so-called H–H₂ transition and, more

* Corresponding author: emma.nigou@phys.ens.fr

importantly, to the statistical distribution of the molecular fraction along random lines of sight. In the local diffuse ISM, this transition occurs at total hydrogen column densities $N_{\text{H}} \sim 10^{20} - 10^{21} \text{ cm}^{-2}$, with a mean transition column density that depends on the Galactic latitude (Gillmon et al. 2006). Across this range, the molecular fraction increases by several orders of magnitude and approaches a mean value of $f_{\text{H}_2} \gtrsim 0.1$, with decreasing dispersion at higher column densities, reflecting line-of-sight averaging over long distances. Initially interpreted as a signature of H-to-H₂ conversion within individual clouds, this transition is now understood to reflect the underlying multiphase structure of the gas (Bellomi et al. 2020). The distribution of H₂ column densities traces the probability of encountering CNM structures along a given line of sight, which itself results from thermal instability and is ultimately regulated by Galactic-scale properties such as the mean density of the diffuse ISM, the intensity of the ambient UV radiation field, and the typical separation between OB associations.

On the other hand, observations of H₂ in both emission and absorption have provided access to its excitation in its pure rotational levels. While the lowest two rotational states ($J = 0, 1$) appear to be thermally coupled with the gas at a kinetic temperature close to that of the CNM ($\sim 80 \text{ K}$), higher rotational states display excitation temperatures of a few hundred Kelvin that anticorrelate with the column density of molecular hydrogen (Spitzer et al. 1974; Savage et al. 1977; Wakker 2006; Shull et al. 2021). Two competing processes are invoked to account for the excitation of these rotational states: collisions in warm environments and far-UV pumping of the electronic transitions followed by fluorescent decay. Multiple studies have shown that the populations of several rotational levels cannot be explained by far-UV pumping alone (Gry et al. 2002; Falgarone et al. 2005; Nehmé et al. 2008; Ingalls et al. 2011) and instead require collisional excitation in gas at temperatures of a few hundred Kelvin along the line of sight. This warm H₂ component, which is connected to the presence of CH⁺ in the diffuse ISM (e.g., Frisch & Jura 1980; Lambert & Danks 1986), may arise from the dissipation of mechanical energy (e.g., Godard et al. 2009, 2014; Myers et al. 2015; Lesaffre et al. 2020; Moseley et al. 2021) or from the evaporation of CNM structures (Valdivia et al. 2016; Godard et al. 2023; Godard 2025). In this context, assembling a large observational sample is essential to enable statistical analyses of H₂ excitation and to constrain models of multiphase, turbulent gas, as such statistics provide insight far beyond what can be inferred from individual lines of sight.

The present study aims to expand the statistical sampling of H₂ rotational excitation in the diffuse ISM by searching for pure rotational line emission in all available background observations collected by the Medium Resolution Spectrometer (MRS) aboard the *James Webb* Space Telescope (JWST). This approach potentially enables the detection of H₂ from the S(1) to S(8) transitions.

2. JWST MIRI-MRS data

2.1. MRS backgrounds and data reduction

The recommended observation strategy for JWST/MRS programs is to measure the background level using dedicated, co-temporal observations of a nearby region in the sky selected to be free of target emission. These background observations are primarily used to remove the sky and telescope contributions (including zodiacal light and thermal emission¹), and to

¹ See the latest JWST technical report for details.

mitigate detector artifacts such as temporal variations in the hot and warm pixels and coherent striping produced by drifts in the effective detector dark current that are currently not captured by the standard calibration pipeline². Such detector systematics are best calibrated using dedicated backgrounds with the same readout ramp length as the science data.

Dithering offsets the target on the detector between exposures, improving spatial sampling and helping reject bad pixels and cosmic-ray residuals when the exposures are combined. A large fraction of MRS backgrounds used in this paper are taken with a two-point extended source dither pattern with detector readout mode and parameters (including exposure time set by the N_{groups} parameter) matched to that used for the science exposures. This is, in most cases, sufficient to mitigate the effects of bad pixels and cosmic rays (Law et al. 2025). In other cases, the data consist of either a single background position (i.e., no dithering) or a four-point extended dither pattern, with a total exposure depth comparable to that of the science observation, as typically used for very faint targets. The MRS covers a total wavelength range from 4.9 to 27.9 μm , separated into four integral field units (IFUs) referred to as channels, each divided into three bands. The channels cover slightly different fields of view (FoVs), from 3.2'' \times 3.7'' (channel 1) up to 6.6'' \times 7.7'' (channel 4), and have different spectral resolutions (from ~ 3700 to ~ 1500 , Labiano et al. 2021).

Dedicated backgrounds (i.e., indicated as such in the Astronomer Proposal Tool) are generally not used directly for scientific purposes, and they are automatically subtracted from science observations in the pipeline. Therefore, for our purpose, we re-reduced the uncalibrated background MRS files up to level3 calibrated spectral cubes, the standard analysis-ready data product for IFU observations. The data reduction was done with the JWST Science Calibration Pipeline (version 1.17.1, Bushouse et al. 2025), with the context 1350 for the Calibration References Data System following the standard procedures (see, e.g., Labiano et al. 2016 and Álvarez-Márquez et al. 2023 for detailed examples of MRS data reduction and calibration). Careful examination of the data showed that the stage 1 corrections (Morrison et al. 2023) could be run with default parameters, and did not reveal significant residual impacting line emission. Naturally, we switched off the master background correction and sky matching steps in stage 3 of the pipeline, before producing the final fully reconstructed science cubes (Law et al. 2023). We did not realign the MRS astrometry thanks to the simultaneous imaging data. Finally, the science cubes were shifted to the usual orientation with north up and east to the left, increasing the total FoV in channel 1 to 6.1'' \times 5.6'' and up to 11.5'' \times 11.5'' in channel 4.

2.2. Lines of sight selection

The data reduction described above was applied to all publicly available Mid-InfraRed Instrument-MRS (MIRI-MRS) observations retrieved from the Mikulski Archive for Space Telescopes (MAST). As of November 2025, this resulted in 297 individual lines of sight, each observed in at least one of the four spectral channels of the MRS instrument. Among these, 21 lines of sight were excluded from the sample. Programs targeting bodies in the Solar System were rejected due to contamination of the background spectra by diffuse light scattered from the primary target (e.g., Bockelée-Morvan et al. 2024), while lines of sight toward the Galactic center were excluded because their mid-infrared

² While the pipeline corrects for dark current offsets in the dark count rate it cannot compensate for changes in the coherent pattern noise.

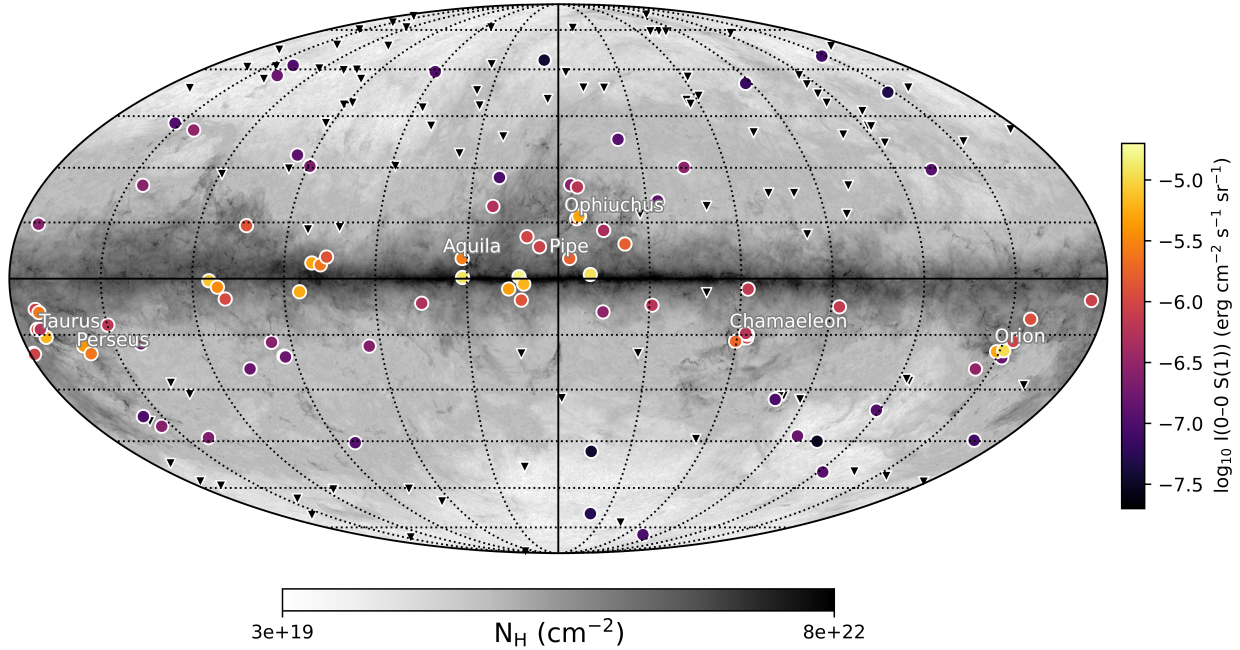


Fig. 1. Aitoff projection of the MRS background lines of sight analyzed in this paper. Filled circles indicate detections of the 0–0 S(1) line of H₂, color-coded by integrated intensity, while triangles show non-detections. All lines of sight are overlaid on the total hydrogen column density map derived from the dust opacity at 353 GHz measured by *Planck* (Planck Collaboration XI 2014).

emission is dominated by dense environments rather than by the diffuse ISM.

The spectral properties of the final sample, which comprises 276 lines of sight, and the procedures used to extract and identify H₂ rotational emission are described in Appendix A. The corresponding measurements are summarized in Tables A.1 and A.2, which report, respectively, the integrated intensities measured for all detected H₂ rotational lines in each spectrum and the detection statistics for each transition. To support the interpretation of the data, we also compiled estimates of the atomic and total hydrogen column densities, as well as the dust temperature and its 12 μm emission, along each line of sight using all-sky surveys. These complementary data are discussed in Appendix B and listed in Table B.1. Although these estimates are derived over solid angles at least 25 times larger than the MRS FoV, they provide valuable contextual information.

A subset of the final sample is displayed in Fig. 1, which shows the distribution of lines of sight in Galactic coordinates, as well as the integrated intensity of the H₂ 0–0 S(1) line, used throughout this paper as a reference for the study of warm and diffuse molecular hydrogen. Several background lines of sight exhibit rotational H₂ emission significantly higher than the bulk of the sample. Upon further inspection, a fraction of these highly emissive lines of sight are associated with the Orion, Ophiuchus, Perseus, and Chamaeleon cloud complexes, all located near regions of active star formation. To exclude such environments and focus on the diffuse ISM, we devised selection criteria based on the total hydrogen column density and the dust temperature, as described in Appendix B.3. Applying these criteria led to the flagging of 43 observations as non-diffuse, leaving a final sample of 233 lines of sight intersecting the diffuse ISM.

3. Diffuse H₂ emission and excitation

We report widespread detections of H₂ pure rotational line emission originating from Galactic material. After excluding the

flagged lines of sight, the S(1) line is detected along 84 lines of sight, up to Galactic latitudes of 78°. Most detections correspond to molecular hydrogen column densities between 10¹⁹ and 10²² cm⁻², with line intensities and excitation properties consistent with absorption measurements in the solar neighborhood. This consistency indicates that the emission predominantly arises from diffuse, largely local interstellar gas³. In this context, we also report the first detections of the pure rotational S(4), S(5), and S(7) lines in the local diffuse ISM. As the detection rate generally decreases with increasing rotational quantum number (see Table A.2), the S(6) and S(8) lines are not detected. The spontaneous Einstein coefficient of the $J \rightarrow J - 2$ pure rotational transition approximately scales as J^6 between $J = 2$ and $J = 9$ (Shull & Hollenbach 1978). This decrease in detection rate therefore reflects the sharp drop in the level populations as J increases.

3.1. Variations with galactic latitude

The integrated intensities of the H₂ 0–0 S(1) line are shown in Fig. 2 as functions of the absolute Galactic latitude of the lines of sight, $|b|$. The measured intensities span more than two orders of magnitude and display a steep decline with increasing latitude, in excellent agreement with values inferred from absorption studies. For exposure times of a few minutes – which apply to most MRS background pointings – the sensitivity limit is $\sim 10^{-7}$ erg cm⁻² s⁻¹ sr⁻¹. This limit corresponds to a column density of the $J = 3$ level of 2×10^{16} cm⁻² (see Eq. (A.1)), approximately two orders of magnitude higher than the detection threshold reached in absorption.

³ We note that two lines of sight have H₂ column densities above 10²² cm⁻². Those lie at Galactic latitudes $|b|$ below $\sim 0.5^\circ$ and intersect the full column of Galactic disk material, similarly to the line of sight analyzed by Falgarone et al. (2005).

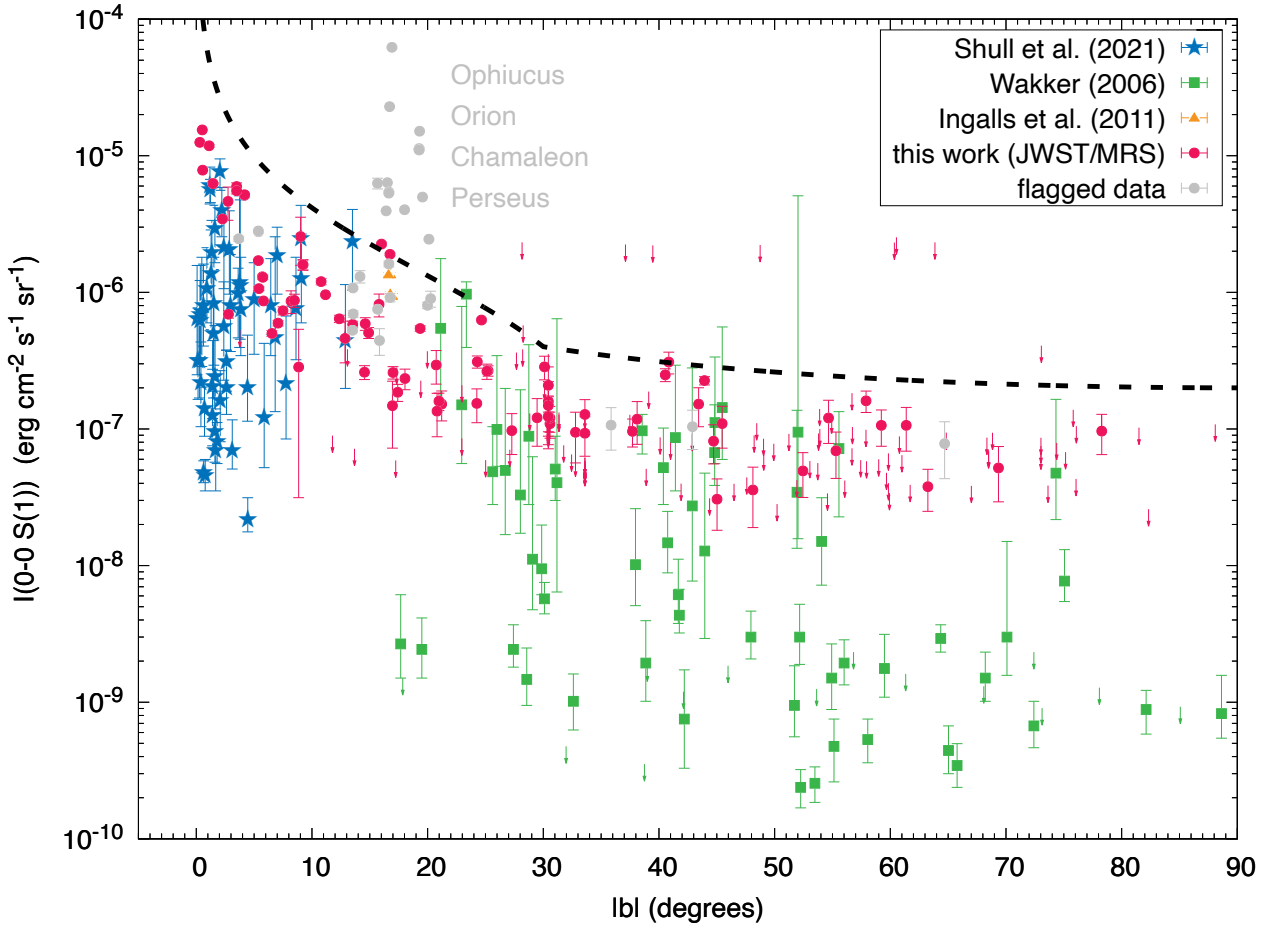


Fig. 2. Integrated intensities of the 0–0 S(1) line of H₂ as a function of the absolute Galactic latitude, $|b|$. Green squares and blue stars correspond to values inferred from absorption measurements toward extragalactic targets and nearby OB stars observed with FUSE (Wakker 2006; Shull et al. 2021), using Eq. (A.1) and the column densities of the $J = 3$ rotational level. Yellow triangles show measurements obtained with *Spitzer* by Ingalls et al. (2011). Gray and red points represent intensities measured along flagged and un-flagged MRS background lines of sight, respectively, while red arrows indicate upper limits set by the sensitivity of the MRS over the wide range of exposure times in the sample. The dashed line shows a geometrically motivated analytical estimate of the maximum expected emission (Eq. (1)).

Interestingly, the upper envelope of the emission follows a well-defined trend that reflects the vertical distribution of molecular gas in the solar neighborhood. Assuming that the local diffuse ISM can be described by two partially molecular components, a spherical component associated with the Local Bubble and a slab representing the more extended Galactic ISM, the maximum expected emission can be written as

$$I_{\max}(|b|) = \frac{C_1}{\max(\sin |b|, r)} + C_2 \left(\frac{1}{\sin |b|} - \frac{1}{\max(\sin |b|, r)} \right). \quad (1)$$

$C_1 = 2 \times 10^{-7}$ and $C_2 = 10^{-6}$ are the maximum emissions (in $\text{erg cm}^{-2} \text{s}^{-1} \text{sr}^{-1}$) of the S(1) line perpendicular to the Galactic plane associated with the Local Bubble component and the extended Galactic ISM, respectively, each adjusted to match the data, while $r = 0.5$ represents the fixed ratio between the characteristic height of the molecular gas layer and the radius of the Local Bubble (e.g., Dame et al. 2001; Lallement et al. 2014). The resulting curve, shown as a dashed line in Fig. 2, closely traces the upper envelope of the observed intensities, providing further indication that the emission arises predominantly from diffuse molecular gas.

This envelope and the distribution of observed S(1) intensities are found to be symmetric with respect to $b = 0$. There

is, however, a statistically significant difference in the detection rate of the S(1) line between northern and southern lines of sight. Although the sample contains twice as many northern lines of sight as southern ones, the detection rate is a factor of two higher in the south, independently of Galactic latitude. This difference cannot be attributed to a bias in integration time, as the northern and southern subsamples have identical integration time distributions. The asymmetry therefore reflects a higher probability of intercepting molecular regions along southern lines of sight, while the intrinsic emission properties of the gas remain unchanged, consistent with the distribution of molecular structures at high Galactic latitude revealed by Röhser et al. (2016).

3.2. Excitation properties

The distribution of H₂ among its rotationally excited states is shown in Fig. 3, which presents excitation diagrams constructed from the column densities of individual rotational levels, normalized by their statistical weights and by the estimated H₂ column density (Appendix B), as functions of the level energies. These excitation diagrams exhibit a striking consistency with those derived from absorption studies, in terms of both population ratios and the range of column densities probed (see

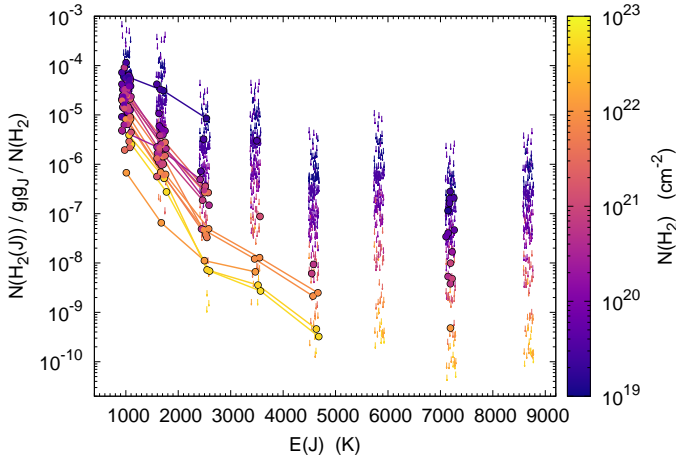


Fig. 3. Excitation diagrams of H₂ derived along the MRS background lines of sight, excluding the flagged data. The column densities of each rotational level are derived from the corresponding line intensities using Eq. (A.1) and are normalized and color-coded by the H₂ column density estimated in Appendix B. For visual clarity, data points are slightly offset along the energy axis by small random values. Detections, along a given line of sight, of all successive rotational transitions up to at least the S(3) line are connected by segments.

Fig. 1.16 of Godard 2025). As found in previous work (e.g., Gry et al. 2002), the rotational populations depart from Local Thermal Equilibrium and display excitation temperatures that increase with the rotational quantum number J . In addition, and as reported by Wakker (2006), the excitation of the rotational levels is linked to the column density of molecular hydrogen. As $N(\text{H}_2)$ increases, the fraction of H₂ in its excited rotational states generally decreases, and the excitation diagrams presumably converge toward a mean distribution resulting from the statistical averaging along increasingly complex lines of sight.

To further quantify the excitation of H₂ in the diffuse ISM, we display in Fig. 4 the excitation temperatures T_{34} and T_{35} , inferred from the populations of the $J = 3, 4$, and 5 rotational levels. Because it involves both para- and ortho-rotational levels, T_{34} serves as a proxy for the overall excitation of H₂ that encodes the combined effects of ortho–para conversion and nonreactive excitation processes. Keeping this important aspect in mind, T_{34} and T_{35} are found to lie between 200 and ~ 1000 K, in agreement with the values derived from absorption studies. As shown in Fig. 4, the two excitation temperatures, derived from our sample and the lines of sight observed in absorption, follow a simple power-law relation,

$$T_{34} = (106 \pm 10) \text{ K} \left(\frac{T_{35}}{150 \text{ K}} \right)^{1.39 \pm 0.12}, \quad (2)$$

where the errors on each parameter correspond to the 3σ marginalized uncertainties derived from the posterior distribution. Taken together, the results presented in Figs. 3 and 4 provide paramount empirical constraints on the excitation conditions of molecular hydrogen in the diffuse ISM. In particular, they encode statistical information on the variations of the ortho-to-para ratio and on the distribution of densities, temperatures, and UV irradiation along a random set of lines of sight.

It has been proposed that the rotational excitation of H₂ is enhanced by UV pumping due to the interaction with the background source along lines of sight targeting hot OB-type stars (e.g., Shull et al. 2021). In contrast, the JWST observations analyzed here probe H₂ emission along lines of sight that do not

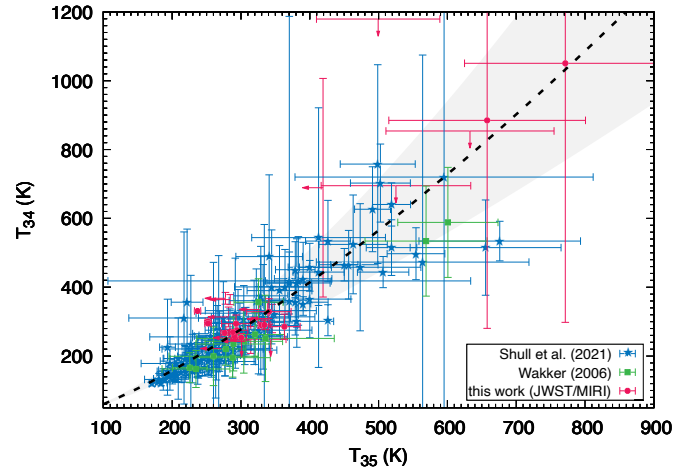


Fig. 4. Excitation temperature T_{34} between the $J = 3$ and $J = 4$ levels as a function of the excitation temperature T_{35} between the $J = 3$ and $J = 5$ levels. Green squares and blue stars correspond to values inferred from absorption measurements toward extragalactic targets and nearby OB stars observed with FUSE (Wakker 2006; Shull et al. 2021). Red points and arrows show measurements and upper limits, respectively, obtained along MRS background lines of sight after excluding flagged data. The dashed black line shows a power-law relation derived between the two temperatures (Eq. (2)), and the shaded region is an envelope obtained by varying each fit parameter independently within its marginalized $\pm 3\sigma$ uncertainty, which encloses 78% of the data points.

target specific background sources and are, a priori, not tied to the presence of nearby massive stars. The fact that the excitation properties, in terms of both excitation diagrams and T_{34} – T_{35} distributions, are similar suggests either that the contribution of local stellar environments is difficult to disentangle from the cumulative excitation along the line of sight, or that their net impact on the averaged excitation properties is limited. A natural interpretation is that the observed excitation arises from a combination of collisional excitation and radiative pumping by the ambient interstellar radiation field integrated along extended path lengths, as found in recent numerical simulations of the diffuse multiphase ISM (see Fig. 1.20 of Godard 2025).

3.3. Cooling rate

The contribution of pure rotational H₂ emission to the energy budget of the diffuse ISM is investigated in Fig. 5, which shows the fluxes of the 0–0 S(1) line as a function of the total hydrogen column density, N_{H} . Combining the MRS background measurements with the values inferred from absorption studies toward high-Galactic-latitude extragalactic sources reveals a sharp increase in S(1) emission at column densities between 10^{20} and 10^{21} cm^{-2} . This behavior coincides with the onset of the H–H₂ transition in diffuse interstellar gas, marking the transition from a predominantly atomic to a partially molecular medium (e.g., Bellomi et al. 2020). Because emission measurements are less sensitive than absorption, the entire MRS sample probes lines of sight with substantial H₂ column densities that have already undergone the H–H₂ transition.

As established from absorption studies, the cooling associated with H₂ emission per unit column density of gas, as well as its partition among the various rotational lines, is expected to vary from one line of sight to another. In particular, the 0–0 S(1) line typically contributes between 5 and 95% of the cooling carried by the rotational ladder. Nevertheless, Fig. 5

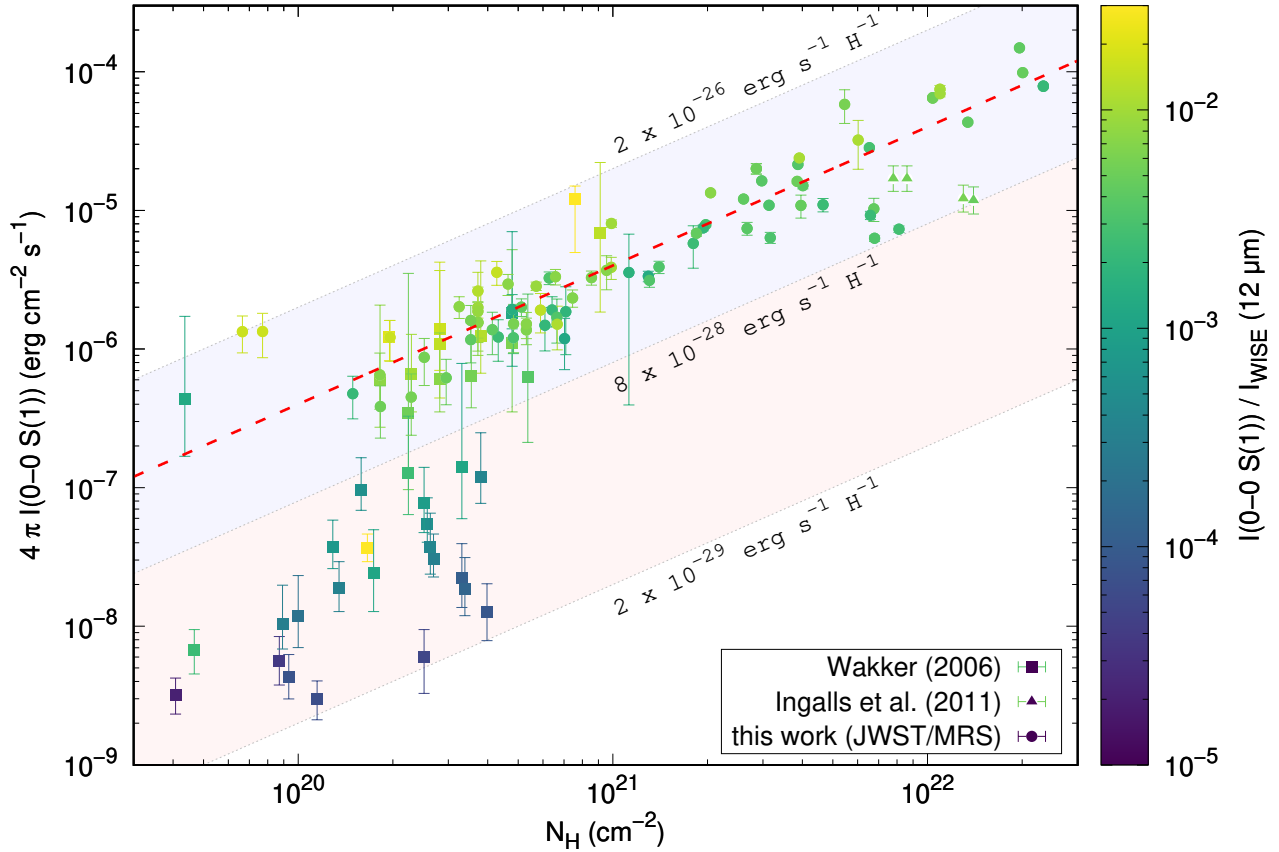


Fig. 5. Flux of the 0–0 S(1) line of H₂ as a function of the total hydrogen column density, N_{H} (Table B.1). Filled squares correspond to values inferred from absorption measurements toward extragalactic targets observed with FUSE (Wakker 2006). Filled circles show intensities measured along MRS background lines of sight after excluding flagged data. Points are color-coded by the ratio of the S(1) integrated intensity to the PAH surface brightness at 12 μm measured by WISE (Table B.1). The three dotted gray lines indicate constant values of the S(1) cooling rate per hydrogen atom. The dashed red line shows a fit to the S(1) cooling rate derived from the MRS sample only (Eq. (3)).

shows that once the H–H₂ transition is reached, the cooling rate of the S(1) line per hydrogen atom,

$$\Lambda_{S(1)} = \frac{4\pi I(0-0 S(1))}{N_{\text{H}}} = 10^{-26.4 \pm 0.3} \text{ erg s}^{-1} \text{ H}^{-1} \quad (3)$$

remains nearly constant, with a dispersion of less than a factor of two over nearly three orders of magnitude in total hydrogen column density. This cooling rate corresponds to about 20% of the cooling rate of the [C II] 158 μm line measured in diffuse gas at high Galactic latitudes (Bock et al. 1993; Lehner et al. 2004). Over the same range of column density, the S(1) line intensity amounts to 0.6% of the polycyclic aromatic hydrocarbon (PAH) surface brightness at 12 μm measured by the Wide-field Infrared Survey Explorer (WISE) (Meisner & Finkbeiner 2014), again with a dispersion smaller than a factor of two. These results are consistent with Hensley et al. (2016), who showed that the PAH fraction increases in diffuse gas and reaches a plateau at $N_{\text{H}} \gtrsim 8 \times 10^{20} \text{ cm}^{-2}$.

The fact that $\Lambda_{S(1)}$ and its ratio to the 12 μm PAH emission remain approximately constant beyond the H–H₂ transition constitutes a new and important result. It provides a link between the physics of WNM–CNM phase transitions, CNM irradiation, H₂ formation, ortho-to-para conversion processes, and H₂ excitation in diffuse gas. In particular, the value inferred from the MRS data (Eq. (3)) is larger than that predicted by models of static photodissociation regions as recently found by Goldsmith et al. (2025).

4. Conclusions and perspectives

This paper introduces a novel observational strategy for probing the diffuse ISM that requires no dedicated telescope time. It demonstrates that diffuse H₂ emission can be readily detected by systematically analyzing the backgrounds of existing JWST observations. This method significantly enhances the scientific yield of the JWST by leveraging data that would otherwise remain unused. The main results can be summarized as follows.

- Of the 297 background observations acquired over the past three years with the MRS instrument, 233 are identified as probing the diffuse, predominantly local ISM. We report the detection of the pure rotational S(1) line of H₂ along 41% of these lines of sight, together with the first detections of the S(4), S(5), and S(7) pure rotational lines in the diffuse ISM.
- Because of sensitivity limits, most detections correspond to lines of sight that have undergone the H–H₂ transition, with molecular hydrogen column densities between 10^{19} and 10^{22} cm^{-2} .
- The S(1) line is detected over a wide range of Galactic latitudes, up to $|b| = 78^\circ$. Its intensity spans more than two orders of magnitude and exhibits a steep decline with increasing Galactic latitude, reflecting both the finite vertical extent of molecular gas above the Galactic disk and the radial extent of the Local Bubble.
- The intrinsic emission properties of H₂ are symmetric with respect to $b = 0$. However, the detection rate differs between

northern and southern Galactic lines of sight, with the S(1) line detected twice as often in the south.

- The spatial distribution and excitation properties of H₂ emission are consistent with those inferred from absorption studies conducted with FUSE. In particular, the excitation temperatures T_{34} and T_{35} , which trace ortho–para conversion and nonreactive excitation processes, lie between 200 and ~1000 K. These temperatures are well correlated with each other and anticorrelated with the H₂ column density.
- Beyond expanding the statistical sample of Galactic H₂ emission, this study reveals a new cooling law for the diffuse ISM. When contextualized with *Planck* and WISE data, lines of sight that have undergone the H–H₂ transition show a remarkably constant S(1) cooling rate per hydrogen atom, with a mean value of $\sim 4 \times 10^{-27}$ erg s⁻¹ H⁻¹ and a dispersion of less than a factor of two over nearly three orders of magnitude in total column density. The S(1) line intensity amounts to 20% of the cooling rate of the [C II] 158 μm line derived at high Galactic latitudes, and 0.6% of the PAH surface brightness at 12 μm, again with a dispersion smaller than a factor of two.

The widespread emission from rotationally excited H₂ across the Galaxy offers a powerful complement to traditional absorption studies. While absorption data provide precise measurements of total H₂ column densities and the ortho-to-para ratio with unmatched sensitivity, they are limited to rare, bright UV background sources. In contrast, current and future JWST observations enable a dramatic increase in statistical coverage across the sky and access to higher rotational levels that were previously undetectable in the diffuse ISM. In this study we focused on MRS data to demonstrate that this emission arises from diffuse gas. Ongoing work focuses on NIRSpec data, opening the door to the first systematic detection of diffuse rovibrational H₂ emission.

Data availability

The full versions of Tables A.1 and B.1 are available at the CDS via <https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/709/A152>.

Acknowledgements. We thank the referee, J. M. Shull, for a thorough reading of the manuscript and for comments that greatly improved the paper. This work is based on observations made with the NASA/ESA/CSA James Webb Space Telescope. The data were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for JWST. The research leading to these results has received financial support by the DIM ORIGINES program from Région Île de France. We would also like to acknowledge the support from the Thematic Actions “Cosmologie et Galaxies” (CG) and “Physique et Chimie du Milieu Interstellaire” (PCMI) of INSU Programme National “Astro”, with contributions from CNRS Physique and CNRS Chimie, INP and IN2P3, CEA, and CNES. The data gathering and processing was made possible by using the CANDIDE cluster at the Institut d’Astrophysique de Paris. The cluster was funded through grants from the PNCG, CNES, DIM-ACAV, the Euclid Consortium, and the Danish National Research Foundation Cosmic Dawn Center (DNRF140). It is maintained by Stephane Roubertol.

References

- Álvarez-Márquez, J., Labiano, A., Guillard, P., et al. 2023, *A&A*, 672, A108
 Argyriou, I., Glasse, A., Law, D. R., et al. 2023, *A&A*, 675, A111
 Bellomi, E., Godard, B., Hennebelle, P., et al. 2020, *A&A*, 643, A36
 Bock, J. J., Hristov, V. V., Kawada, M., et al. 1993, *ApJ*, 410, L115
 Bockelée-Morvan, D., Lellouch, E., Poch, O., et al. 2024, *A&A*, 681, A27
 Bushouse, H., Eisenhamer, J., Dencheva, N., et al. 2025, <https://doi.org/10.5281/zenodo.6984365>
 Carruthers, G. R. 1970, *ApJ*, 161, L81
 Castro Neto, A. H., Guinea, F., Peres, N. M. R., Novoselov, K. S., & Geim, A. K. 2009, *Rev. Mod. Phys.*, 81, 109
 Compiègne, M., Verstraete, L., Jones, A., et al. 2011, *A&A*, 525, A103
 Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, *ApJ*, 547, 792
 Draine, B. T., & Li, A. 2007, *ApJ*, 657, 810
 Falgarone, E., Verstraete, L., Pineau des Forêts, G., & Hily-Blant, P. 2005, *A&A*, 433, 997
 Frisch, P. C., & Jura, M. 1980, *ApJ*, 242, 560
 Gillmon, K., & Shull, J. M. 2006, *ApJ*, 636, 908
 Gillmon, K., Shull, J. M., Tumlinson, J., & Danforth, C. 2006, *ApJ*, 636, 891
 Godard, B. 2025, Habilitation à diriger des recherches, Observatoire de Paris, France
 Godard, B., Falgarone, E., & Pineau des Forêts, G. 2009, *A&A*, 495, 847
 Godard, B., Falgarone, E., & Pineau des Forêts, G. 2014, *A&A*, 570, A27
 Godard, B., Pineau des Forêts, G., Hennebelle, P., Bellomi, E., & Valdivia, V. 2023, *A&A*, 669, A74
 Goldsmith, P. F., Wang, S., Wang, X., et al. 2025, *ApJ*, 985, L4
 Gry, C., Boulanger, F., Nehmé, C., et al. 2002, *A&A*, 391, 675
 Hensley, B. S., Draine, B. T., & Meisner, A. M. 2016, *ApJ*, 827, 45
 HI4PI Collaboration (Ben Bekhti, N., et al.) 2016, *A&A*, 594, A116
 Ingalls, J. G., Bania, T. M., Boulanger, F., et al. 2011, *ApJ*, 743, 174
 Labiano, A., Azzollini, R., Bailey, J., et al. 2016, *SPIE*, 9910, 947
 Labiano, A., Argyriou, I., Álvarez-Márquez, J., et al. 2021, *A&A*, 656, A57
 Lallement, R., Vergely, J.-L., Valette, B., et al. 2014, *A&A*, 561, A91
 Lambert, D. L., & Danks, A. C. 1986, *ApJ*, 303, 401
 Law, D. R., E. Morrison, J., Argyriou, I., et al. 2023, *AJ*, 166, 45
 Law, D. R., Argyriou, I., Gordon, K. D., et al. 2025, *AJ*, 169, 67
 Lehner, N., Wakker, B. P., & Savage, B. D. 2004, *ApJ*, 615, 767
 Leroy, A. K., Walter, F., Sandstrom, K., et al. 2013, *AJ*, 146, 19
 Lesaffre, P., Todorov, P., Levrier, F., et al. 2020, *MNRAS*, 495, 816
 Lin, L., Pan, H.-A., Ellison, S. L., et al. 2019, *ApJ*, 884, L33
 Meisner, A. M., & Finkbeiner, D. P. 2014, *ApJ*, 781, 5
 Miville-Deschênes, M.-A., Murray, N., & Lee, E. J. 2017, *ApJ*, 834, 57
 Morrison, J. E., Dicken, D., Argyriou, I., et al. 2023, *PASP*, 135, 075004
 Moseley, E. R., Draine, B. T., Tomida, K., & Stone, J. M. 2021, *MNRAS*, 500, 3290
 Myers, A. T., McKee, C. F., & Li, P. S. 2015, *MNRAS*, 453, 2747
 Nehmé, C., Le Bourlot, J., Boulanger, F., Pineau des Forêts, G., & Gry, C. 2008, *A&A*, 483, 485
 Pessa, I., Schinnerer, E., Leroy, A. K., et al. 2022, *A&A*, 663, A61
 Planck Collaboration XI. 2014, *A&A*, 571, A11
 Rachford, B., Snow, T., Tumlinson, J., et al. 2002, *ApJ*, 577, 221
 Rigby, J. R., Lightsey, P. A., García Marín, M., et al. 2023, *PASP*, 135, 048002
 Röhsler, T., Kerp, J., Lenz, D., & Winkel, B. 2016, *A&A*, 596, A94
 Roueff, E., Abgrall, H., Czachorowski, P., et al. 2019, *A&A*, 630, A58
 Savage, B. D., Bohlin, R. C., Drake, J. F., & Budich, W. 1977, *ApJ*, 216, 291
 Shull, J. M., & Beckwith, S. 1982, *ARA&A*, 20, 163
 Shull, J. M., & Hollenbach, D. J. 1978, *ApJ*, 220, 525
 Shull, J. M., Danforth, C. W., & Anderson, K. L. 2021, *ApJ*, 911, 55
 Skalidis, R., Goldsmith, P. F., Hopkins, P. F., & Ponnada, S. B. 2024, *A&A*, 682, A161
 Spitzer, Jr., L., & Jenkins, E. B. 1975, *ARA&A*, 13, 133
 Spitzer, Jr., L., Cochran, W. D., & Hirshfeld, A. 1974, *ApJS*, 28, 373
 Valdivia, V., Hennebelle, P., Gérin, M., & Lesaffre, P. 2016, *A&A*, 587, A76
 Wakker, B. P. 2006, *ApJS*, 163, 282
 Wolfire, M. G., McKee, C. F., Hollenbach, D., & Tielens, A. 2003, *ApJ*, 587, 278

Appendix A: Spectra and integrated intensities

Appendix A.1: Spectral features

As shown by Rigby et al. (2023), MRS background spectra include several structured emission components: the thermal emission from the telescope, which dominates above $15 \mu\text{m}$; the zodiacal light, which dominates below $15 \mu\text{m}$; and the stray light, which becomes significant below $5 \mu\text{m}$. We searched for H_2 emission on top of this structured continuum and instrumental features, and above the instrument noise, which increases significantly near the edges of the spectral bands (below 7 and above $27 \mu\text{m}$).

As an illustration, we display in Fig. A.1 a representative background spectrum observed in all four MRS IFU channels, along with zoomed-in panels centered on each H_2 pure rotational line. The continuum emission is subtracted by applying a sliding median filter with sigma clipping, using a kernel width of 20 spectral resolution elements. Galactic H_2 lines are searched for by performing Gaussian fits over a velocity range of 600 km s^{-1} around the corresponding rest frequencies of each transition. The S(1) and S(2) lines are considered detected when their integrated intensities and maximum amplitude exceed the 3σ level, where the uncertainties σ are estimated from the standard deviation of the continuum-subtracted spectrum. Higher- J rotational lines are considered detected if their integrated intensities and maximum amplitude exceed 3σ and if at least one of the S(1) or S(2) lines is detected. Due to the strong instrumental noise beyond $27 \mu\text{m}$ (Argyriou et al. 2023), the 0-0 S(0) line of H_2 is not included in our analysis.

Appendix A.2: Line intensities

A subset of the lines of sight retrieved from the MAST archive, together with the integrated intensities of all detected H_2 lines and their associated uncertainties, is presented in Table A.1. The full version of the table is available online at the Centre de Données astronomiques de Strasbourg (CDS). The sample covers a broad region of the sky, spanning all Galactic longitudes and latitudes (see Fig. 1). Because the background observations originate from programs targeting sources with widely varying luminosities, the exposure times span a large range, from a few tens of seconds up to 46 minutes.

All pure rotational lines from S(1) to S(7) are detected at least once across the full observational sample. As described in Sect. 2.2 and detailed in Appendix B.3, selection criteria based on the total hydrogen column density and the average dust temperature are applied to identify lines of sight intersecting dense and strongly irradiated material possibly associated with star-forming complexes. When these flagged observations are excluded, the S(6) line is no longer detected.

Appendix A.3: Detection statistics

Not all lines of sight were observed in every MRS IFU channel. The dataset presented in Table A.1 is therefore incomplete. The distribution of channel coverage across the sample is summarized in Table A.2, which also provides detection statistics for each H_2 pure rotational line. Restricting the analysis to diffuse lines of sight, the 0-0 S(1) line of H_2 is detected along 84 lines of sight, corresponding to a detection rate of 41%. The detection rate decreases steadily for higher-energy rotational transitions, with a few detections reported for the S(4), S(5), and S(7) lines. This trend reflects both the decreasing intrinsic brightness of

higher- J transitions and the limited sensitivity of the background observations.

Appendix A.4: Conversion to column densities

Interpreting the observed H_2 emission and performing comparisons with ancillary data derived from absorption studies requires converting integrated line intensities into column densities, and vice versa. Assuming optically thin emission, the integrated intensity of a transition, I_{ul} , can be converted into the column density of the upper level, N_u , according to

$$N_u = \frac{4\pi I_{ul}}{A_{ul} h \nu_{ul}}, \quad (\text{A.1})$$

where A_{ul} and ν_{ul} are the Einstein coefficient and frequency of the transition, respectively, and h is the Planck constant. This approximation is valid provided that self-absorption by H_2 and absorption by PAHs can be neglected.

Because most of the H_2 column density resides in the $J = 0$ and $J = 1$ levels in the diffuse ISM, only the S(1) transition can potentially be affected by self-absorption among the pure rotational lines considered here. Using a lower-bound estimate for the line width of 1 km s^{-1} , the peak opacity of the S(1) line remains below 0.1 as long as $N(\text{H}_2) \lesssim 10^{23} \text{ cm}^{-2}$.

Absorption by PAHs can be estimated using conservative assumptions. Approximating PAHs as planar graphene flakes with a representative size of 6.4 \AA (Compiègne et al. 2011) yields ~ 50 carbon atoms per PAH, based on the graphene lattice geometry (Castro Neto et al. 2009). For a PAH mass fraction relative to dust of 4.6% in the diffuse ISM (Draine & Li 2007) and a dust-to-gas mass ratio of 1%, the PAH abundance relative to hydrogen is $x_{\text{PAH}} \sim 10^{-6}$. Using an upper-bound estimate of the absorption cross section per PAH over the MIRI wavelength range of 10^{-18} cm^2 (Draine & Li 2007), PAH absorption therefore remains negligible, with an optical depth below 0.1, as long as the total hydrogen column density $N_{\text{H}} \lesssim 10^{23} \text{ cm}^{-2}$.

As shown in Appendix B (Table B.1), both conditions, for negligible H_2 self-absorption and negligible PAH absorption, are satisfied throughout our sample, except for the flagged lines of sight (see Sect. 2.2 and Appendix B.3). Throughout this paper, we therefore use intensity and upper-level column density interchangeably, enabling direct comparisons with absorption measurements and the construction of excitation diagrams (see Sect. 3).

Appendix B: Estimates of gas and dust properties

Appendix B.1: All-sky surveys

The intensities of the pure rotational lines of H_2 measured along the lines of sight listed in Table A.1 arise from Galactic material, with systemic velocities and associated uncertainties consistent with the range expected from Galactic rotation. To interpret the origin and excitation properties of this emission, it is therefore essential to estimate the associated gas column densities and dust emission along the same lines of sight. These ancillary quantities provide important contextual information.

To this end, we compiled estimates of the atomic hydrogen column density, $N(\text{H})$, the total hydrogen column density, N_{H} , the PAH specific intensity at $12 \mu\text{m}$, $B_{\nu}(12 \mu\text{m})$, and the average dust temperature, T_{dust} , using publicly available all-sky surveys from HI4PI, Planck, and WISE (HI4PI Collaboration 2016; Planck Collaboration XI 2014; Meisner & Finkbeiner 2014). Previous combined analyses of HI, dust, and CO emission across

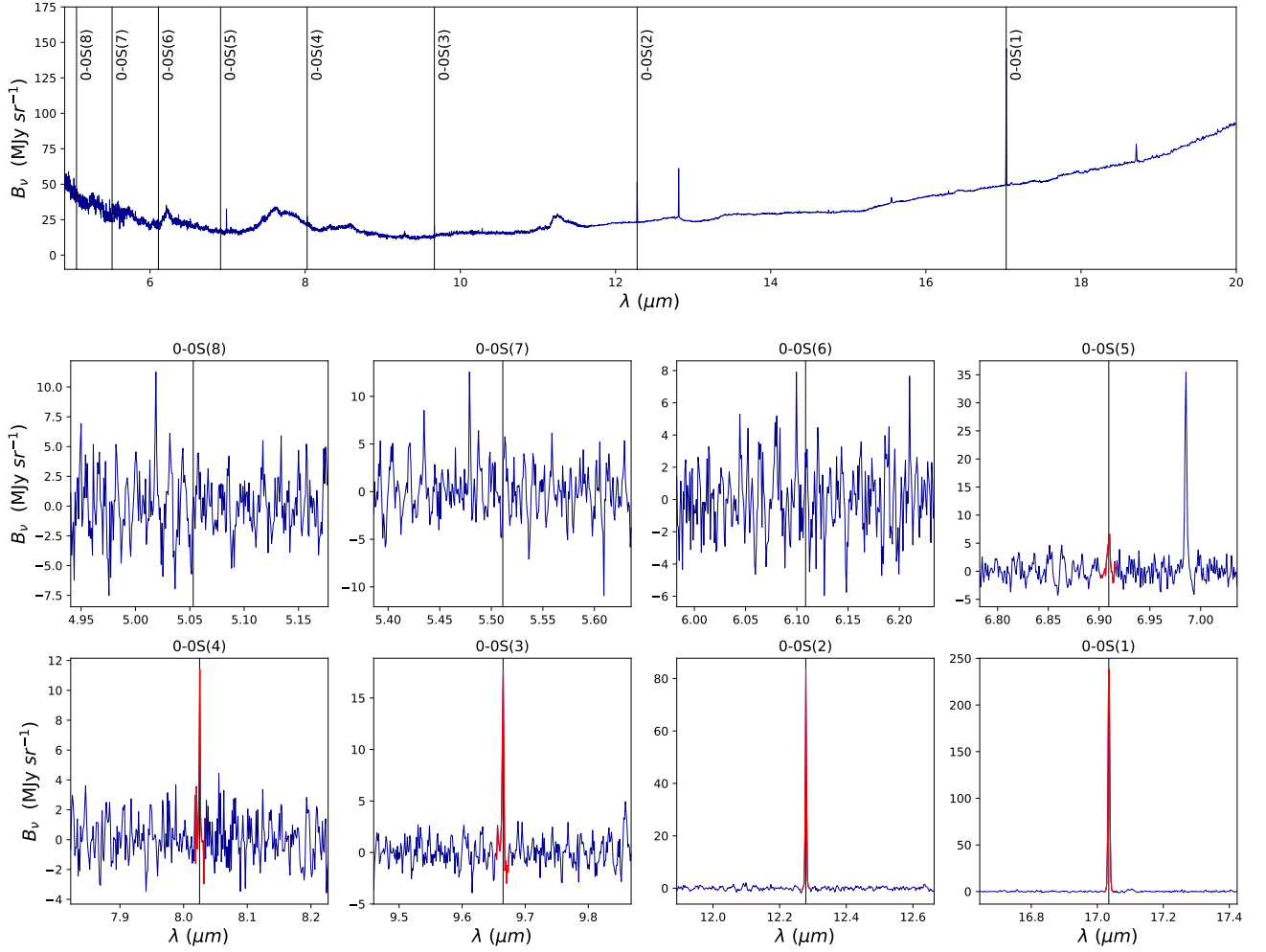


Fig. A.1: Example of a MRS background spectrum observed at Galactic coordinates $l, b = 12.881^\circ, 0.518^\circ$ (Program ID: 01290, Observation ID: 078). *Top panel*: Observed specific intensity as a function of wavelength, after stitching the channel-specific spectra. *Bottom panels*: Continuum-subtracted intensity in zoomed-in windows centered on each H_2 rotational line. Detected lines are highlighted in red.

the Galaxy have shown that N_H and the dust opacity at 353 GHz, τ_{353} , can be approximated by two distinct linear relations below and above $N_H \simeq 6 \times 10^{20} \text{ cm}^{-2}$ (Planck Collaboration XI 2014). Following these studies, the total hydrogen column density is computed as

$$N_H = \begin{cases} \frac{\tau_{353}}{6.3 \times 10^{-27}} \text{ cm}^{-2} & \text{if } \frac{\tau_{353}}{8.26 \times 10^{-27}} < 6 \times 10^{20} \text{ cm}^{-2} \\ \frac{\tau_{353}}{1.2 \times 10^{-26}} \text{ cm}^{-2} & \text{otherwise.} \end{cases} \quad (\text{B.1})$$

The resulting values of $N(H)$, N_H , τ_{353} , $B_v(12 \mu\text{m})$, and T_{dust} for each line of sight are listed in Table B.1.

A key limitation of this approach lies in the spatial resolution mismatch between the ancillary all-sky surveys and the JWST/MRS observations. The angular resolution of the HI4PI data product used in this paper is $16.2''$, while that of the dust and infrared maps from *Planck* and WISE is $\sim 5''$. In contrast, the MRS instrument aboard the JWST provides effective FoVs between $6''$ and $12''$ depending on the IFU channel (see Sect. 2.1). As a result, the derived column densities and infrared specific intensities represent large-scale averages that may not capture small-scale structure or fluctuations within the MRS FoV. These

quantities should therefore be interpreted as coarse contextual estimates rather than precise measurements of the conditions at the exact location of the H_2 emission.

Appendix B.2: Molecular fraction

Bearing this caveat in mind, the molecular hydrogen column density along each line of sight is estimated by subtracting the atomic hydrogen contribution from the total hydrogen column density:

$$N(H_2) = \frac{N_H - N(H)}{2}. \quad (\text{B.2})$$

This approach relies on the assumption that fully ionized phases do not contribute significantly to dust extinction, so that the column density derived from dust primarily traces neutral atomic and molecular gas. Because $N(H)$ and N_H are often comparable along the lines of sight studied in this paper—particularly at high Galactic latitude—this procedure frequently yields negative values of $N(H_2)$, reflecting the fact that the molecular component lies within the combined uncertainties of the two quantities. This behavior has been reported in previous studies (e.g., Skalidis et al. 2024). Given the uncertainties on both $N(H)$ and, more

Table A.1: Subset of the full catalog of background observations retrieved from the **MAST** archive and associated measured integrated intensities of H₂ pure rotational lines.

PID	OID	flag	long (°)	lat (°)	S(1)	err	S(2) (erg cm ⁻² s ⁻¹ sr ⁻¹)	err	S(3)	err	...
01192	002	0	104.190	14.155	1.3 (-6)	1.4 (-7)	9.4 (-7)	2.0 (-7)	<1.2 (-6)		...
01192	011	0	206.846	-16.758	9.2 (-7)	6.4 (-8)	4.9 (-7)	2.0 (-7)	<5.3 (-7)		...
01257	002	0	160.482	-18.006	4.0 (-6)	8.7 (-8)	1.8 (-6)	1.4 (-7)	8.3 (-7)	3.6 (-7)	...
01267	004	0	36.646	53.026	<7.5 (-8)		<1.2 (-7)		<2.8 (-7)		...
...
01023	038	1	271.870	-32.823			<9.3 (-6)				...
01039	006	1	96.451	29.752			<1.5 (-7)				...
01049	004	1	280.831	-32.783	<6.0 (-8)		<1.3 (-7)		<2.2 (-7)		...
01050	008	1	89.262	30.576	1.1 (-7)	2.4 (-8)	<1.1 (-7)		<1.8 (-7)		...
01050	010	1	89.262	30.576	1.2 (-7)	2.6 (-8)	1.6 (-7)	6.2 (-8)	<2.0 (-7)		...
01204	018	1	326.055	20.828			<1.1 (-7)				...
01204	020	1	22.219	19.352			1.5 (-7)	6.0 (-8)			...
01204	021	1	106.999	-24.266			<1.1 (-7)				...
01253	013	1	194.390	-28.832	<1.3 (-7)		<2.7 (-7)		<6.0 (-7)		...
01261	002	1	271.739	-32.774			<1.4 (-6)				...
01265	002	1	20.751	27.278	9.7 (-8)	3.2 (-8)	<1.8 (-7)		<2.7 (-7)		...
01269	003	1	309.626	19.409	<2.3 (-7)		<4.6 (-7)		<1.2 (-6)		...
01282	019	1	278.683	22.963	<2.1 (-7)		<4.9 (-7)		<9.6 (-7)		...
...

Notes. Program ID (PID) and Observation ID (OID) correspond to JWST identifiers. The flag column indicates the classification of each line of sight as diffuse (1) or non-diffuse (0) based on the selection criteria presented in Appendix B.3. Numbers in parentheses are powers of ten. The full version of this table is available online at the CDS.

 Table A.2: Summary of the 276 lines of sight analyzed in this paper (see Sect. 2.2) and corresponding numbers of detection of the pure rotational lines of H₂.

S(J)	λ_0 (μm)	A (s ⁻¹)	total sample		detections	
			all	diffuse	all	diffuse
S(0)	28.2188	2.94 (-11)	–	–	–	–
S(1)	17.0348	4.76 (-10)	248	206	116	84
S(2)	12.2786	2.76 (-09)	237	197	66	40
S(3)	9.6649	9.84 (-09)	217	178	42	23
S(4)	8.0250	2.64 (-08)	237	197	20	7
S(5)	6.9095	5.88 (-08)	248	206	24	9
S(6)	6.1086	1.14 (-07)	217	178	2	0
S(7)	5.5112	2.00 (-07)	237	197	23	17
S(8)	5.0531	3.24 (-07)	237	197	0	0

Notes. Diffuse and non-diffuse lines of sight (flagged 1 and 0 in Table A.1, respectively) are identified based on the selection criteria presented in Appendix B.3. The rest wavelengths, λ_0 , and spontaneous-emission Einstein coefficients, A, of the pure rotational transitions of H₂ are taken from Roueff et al. (2019). Numbers in parenthesis are power of ten.

critically, on the calibration used to derive N_{H} , we estimate that $N(\text{H}_2)$ cannot be robustly constrained below about 10% of the total hydrogen column density. However, because the sensitivity of the MRS limits detections to relatively high molecular hydrogen column densities, the majority of lines of sight with detected H₂ emission lie above this threshold.

The derived column densities of atomic and molecular hydrogen, the corresponding averaged molecular fraction

$$f_{\text{H}_2} = \frac{2N(\text{H}_2)}{N_{\text{H}}}, \quad (\text{B.3})$$

and the PAH and dust specific intensities, span more than four orders of magnitude across the full observational sample. With the exception of a few flagged lines of sight (see Appendix B.3 and Table B.1), both N_{H} and $N(\text{H}_2)$ remain below 10^{23} cm^{-2} , thereby validating the optically thin approximation and the omission of dust extinction adopted in Eq. A.1 to convert intensities into upper-level column densities.

Appendix B.3: Selection of diffuse lines of sight

Not all MRS background observations retrieved from the **MAST** archive probe purely diffuse interstellar gas. A subset of lines of sight shows enhanced H₂ rotational emission or elevated dust temperatures, indicative of contributions from irradiated environments such as star-forming complexes or from denser, possibly self-gravitating structures. To isolate sightlines representative of the diffuse ISM, additional selection criteria are required.

Figure B.1 shows the distributions of the logarithm of the total hydrogen column density projected perpendicular to the Galactic plane,

$$N_{\text{H}}^{\perp} = N_{\text{H}} \sin |b|, \quad (\text{B.4})$$

and of the line-of-sight-averaged dust temperature, T_{dust} . The bulk of the sample is distributed in approximately symmetric cores centered around $N_{\text{H}}^{\perp} \approx 3 \times 10^{20} \text{ cm}^{-2}$ and $T_{\text{dust}} \approx 20 \text{ K}$, characteristic of diffuse Galactic gas (Miville-Deschênes et al. 2017; Planck Collaboration XI 2014). A smaller fraction of the lines of sight extends toward higher values of both quantities, forming asymmetric wings indicative of denser or more strongly irradiated environments. We therefore classify as non-diffuse all lines of sight lying more than three standard deviations from the core of either distribution, corresponding to $N_{\text{H}}^{\perp} > 2 \times 10^{21} \text{ cm}^{-2}$ or $T_{\text{dust}} > 25 \text{ K}$.

Table B.1: Estimates of gas column densities and PAH and dust emission properties along the lines of sight listed in Table A.1.

PID	OID	flag	long (°)	lat (°)	T_{dust} (K)	$N(\text{H})$ (cm^{-2})	N_{H} (cm^{-2})	τ_{353}	$B_{\nu}(12\ \mu\text{m})$ (MJy sr^{-1})
01192	002	0	104.190	14.155	19.36	2.2 (+21)	1.1 (+22)	1.3 (-4)	2.219
01192	011	0	206.846	-16.758	27.43	2.0 (+21)	6.6 (+21)	7.9 (-5)	10.948
01257	002	0	160.482	-18.006	17.16	1.2 (+21)	5.5 (+22)	6.6 (-4)	4.988
01267	004	0	36.646	53.026	27.92	4.0 (+20)	5.0 (+20)	6.0 (-6)	0.080
...
01023	038	1	271.870	-32.823	19.34	4.1 (+20)	5.2 (+20)	3.3 (-6)	0.086
01039	006	1	96.451	29.752	21.63	3.7 (+20)	4.3 (+20)	2.7 (-6)	0.106
01049	004	1	280.831	-32.783	24.66	1.3 (+21)	7.0 (+20)	8.4 (-6)	0.314
01050	008	1	89.262	30.576	19.73	3.2 (+20)	5.3 (+20)	3.3 (-6)	0.067
01050	010	1	89.262	30.576	19.73	3.2 (+20)	5.3 (+20)	3.3 (-6)	0.067
01204	018	1	326.055	20.828	21.92	5.4 (+20)	6.6 (+20)	4.2 (-6)	0.158
01204	020	1	22.219	19.352	19.73	1.0 (+21)	1.8 (+21)	2.2 (-5)	0.488
01204	021	1	106.999	-24.266	21.32	9.0 (+20)	4.5 (+20)	5.5 (-6)	0.364
01253	013	1	194.390	-28.832	18.63	7.2 (+20)	5.8 (+20)	7.0 (-6)	0.151
01261	002	1	271.739	-32.774	18.86	3.9 (+20)	6.2 (+20)	3.9 (-6)	0.092
01265	002	1	20.751	27.278	22.65	5.5 (+20)	4.3 (+20)	5.2 (-6)	0.163
01269	003	1	309.626	19.409	23.89	2.7 (+20)	7.1 (+20)	8.6 (-6)	0.245
01282	019	1	278.683	22.963	19.69	4.6 (+20)	4.8 (+20)	5.7 (-6)	0.140
...

Notes. Atomic hydrogen column densities, $N(\text{H})$, are derived from the HI4PI survey at an angular resolution of $16.2'$ (HI4PI Collaboration 2016). The total hydrogen column densities, N_{H} , are inferred from *Planck* data (Planck Collaboration XI 2014), following equation B.1. The PAH specific intensity at $12\ \mu\text{m}$, $B_{\nu}(12\ \mu\text{m})$, the average dust temperature, T_{dust} , and the dust opacity at 353 GHz, τ_{353} , are obtained from WISE (Meisner & Finkbeiner 2014), and *Planck* all-sky maps (Planck Collaboration XI 2014), both smoothed to $5'$ resolution. Numbers in parentheses are powers of ten. The full version of this table is available online at the CDS.

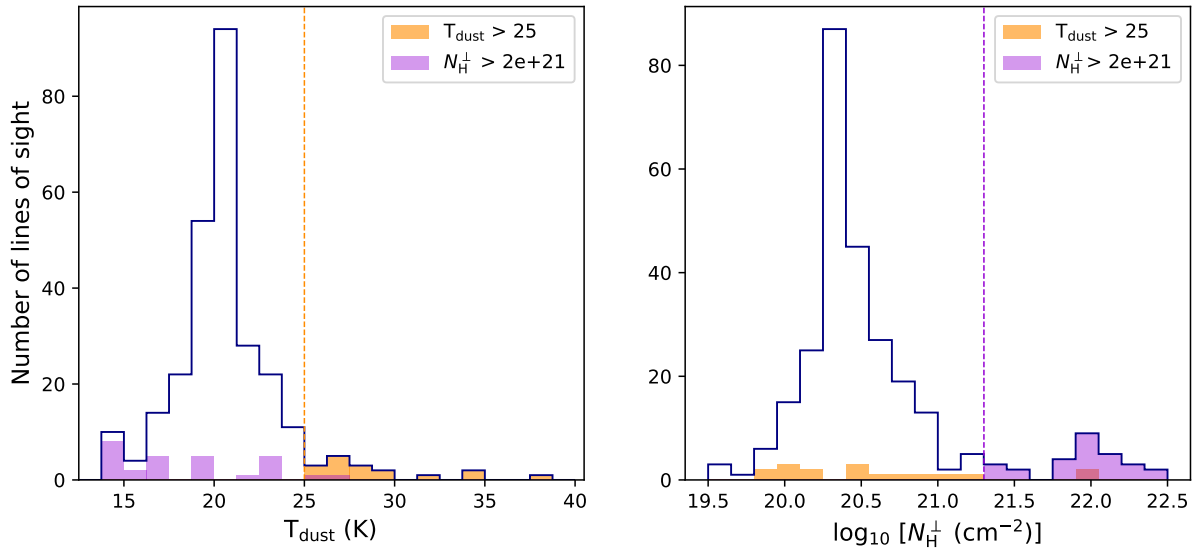


Fig. B.1: Integrated properties of the observational sample. *Left panel:* Distribution of the line-of-sight-averaged dust temperature, T_{dust} . *Right panel:* Distribution of the logarithm of the total hydrogen column density projected perpendicular to the Galactic plane, N_{H}^{\perp} . Lines of sight classified as non-diffuse ($N_{\text{H}}^{\perp} > 2 \times 10^{21}\ \text{cm}^{-2}$ or $T_{\text{dust}} > 25\ \text{K}$) are highlighted in color.

After excluding these flagged data, the H_2 column densities associated with detected rotational lines are typically above $10^{19}\ \text{cm}^{-2}$, corresponding to molecular fractions between 0.1 and 1. This indicates that most retained lines of sight have undergone the $\text{H}-\text{H}_2$ transition. Compared to absorption studies in the solar neighborhood, our sample occupies the higher end of the column density and molecular fraction distributions, more representative of sightlines in or near the Galactic plane (Shull et al. 2021). This reflects the sensitivity threshold of the MRS instrument,

which favors the detection of diffuse regions with relatively high molecular content.