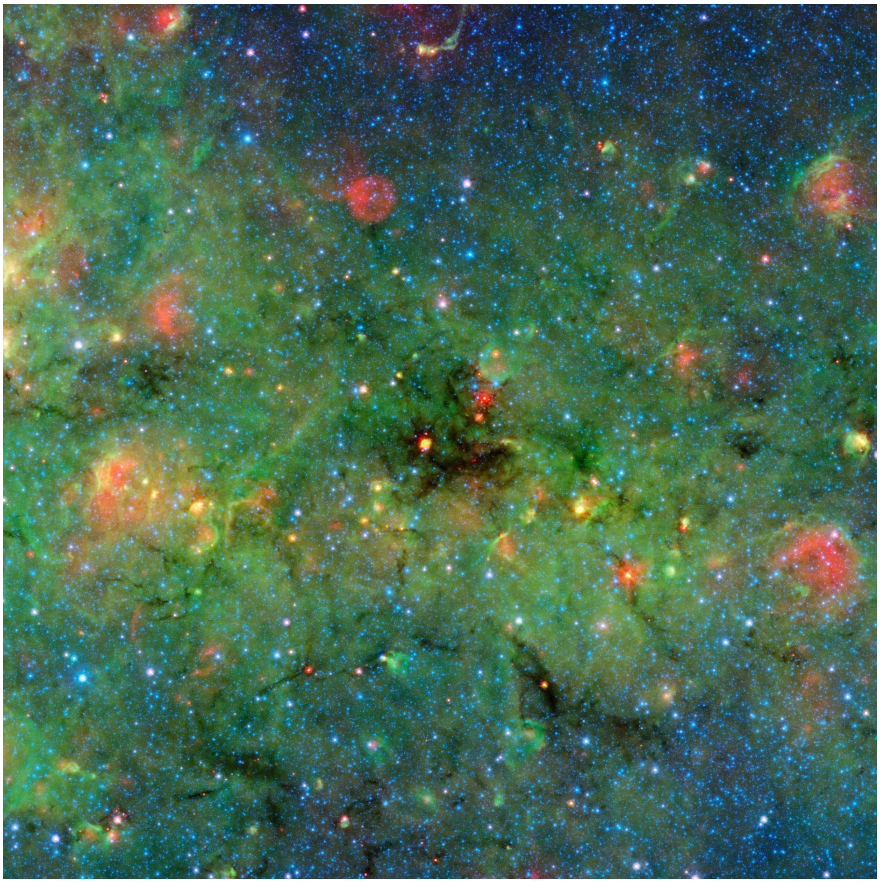


Spectroscopic Study of the Astrophysics and Astrochemistry of Protostars in a Clustered Region of Massive Star Formation

Master's thesis in Physics

ARVID ANDERSSON

DEPARTMENT OF SPACE, EARTH AND ENVIRONMENT

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2024-06-07
www.chalmers.se

Spectroscopic Study of the Astrophysics and Astrochemistry of Protostars
in a Clustered Region of Massive Star Formation
ARVID ANDERSSON

Supervisors: Dr. Mélisse Bonfand, University of Virginia, Department of Astronomy
Dr. Prasanta Gorai, University of Oslo, Institute of Theoretical Astrophysics
Examiner: Prof. Jonathan C. Tan, Chalmers University of Technology, Department of Space, Earth and Environment

Master's Thesis 2024
Department of Space, Earth and Environment
Division of Astronomy and Plasma Physics
Part of Chalmers Initiative on Cosmic Origins
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Spitzer Space Telescope composite image of infrared dark cloud G028.37+00.07, NASA, JPL-Caltech, University of Zurich.

Typeset in LaTeX
Gothenburg, Sweden 2024-06-07

Abstract

Massive stars play crucial roles in the energy budget and evolution of galaxies, and also show great chemical complexity in their protostellar phases. Still, the processes through which they form remain areas of active research. With the advent of state-of-the-art radio interferometers, the hot molecular core phase has proven to be a window into the physical and chemical properties of massive star formation. Spectroscopic observations of molecules released into gas phase from heated dust grains in the protostellar envelope can now be used to probe the star forming regions. In this study, data from observations using the Submillimeter Array (SMA) interferometer toward the infrared dark cloud and active high-mass star forming region G028.37+00.07 (G28.37), has been analyzed. G28.37 is one of several clustered massive protostellar regions whose emission from potential hot molecular cores has been targeted through a continuous 204-248 GHz systematic spectral line survey.

Seven sources have been identified based on their millimeter-continuum emission, and spectra extracted from beam sized apertures centered on each source. Then, methods have been developed to automate an objective line identification algorithm, in order to prepare for future analysis of these surveyed regions. Detected peaks in the spectra above a set signal to noise threshold are automatically cross-matched to the Cologne Database for Molecular Spectroscopy (CDMS), and their line profiles investigated through Gaussian fitting. The chemical inventory of the sources is presented and compared, revealing CH_3OH as the most prevalent species. Emission from $C^{34}S$, H_2CS, and $C^{17}O$, among other species, is also detected and presented. Column densities and abundances are estimated and compared to previous studies of hot molecular cores. Finally, extracted kinematic properties of the species are also discussed. Concluding, G28.37 is suggested to be an earlier stage, less evolved, massive protostellar region, motivated by its relatively low chemical complexity compared to other hot molecular cores.

Acknowledgements

I would like to express my gratitude toward Prof. Jonathan C. Tan for allowing me to carry out this Master's thesis work and research project within his group, and welcoming me into a great community. Thank you for inviting me into the world of astronomy and research through the weekly science discussions, the colloquia, and the opportunities to meet renowned researchers in the field. Furthermore, thank you to Dr. Mélisse Bonfand and Dr. Prasanta Gorai, for your support and guidance throughout the project. Thank you for constantly cheering me on, and for your passion in teaching me the ways of your areas of expertise.

Contents

1 Introduction

1.1 High-mass star formation

High-mass stars have masses from about $8\,M_\odot$ and upward, living short but impactful lives where they play crucial roles through for example supernovae events and energetic radiation (Motte et al. 2018). Furthermore, massive star forming regions have proven to contain organic molecules, of interest in studying the origin of prebiotic chemistry in space. Understanding the processes through which such giant stars form thus has relevance for understanding star formation in galaxies as a whole, as well as the overall evolution of galaxies and their energy budget, while also relating to the search for traces of life in space. However, contrary to those of lower-mass stars, the mechanisms governing high-mass star formation are still significantly less understood and therefore remains an active area of research.

Studying the formation process of high-mass stars comes with a number of difficulties. On astronomical timescales, the high-mass protostellar evolution is rather quick. As this rapid evolution is believed to occur in dense molecular clouds, the protostars are often still heavily obscured by dense envelopes during critical stages such as upon reaching the main sequence. This abundant dust in which the young star is embedded hinders observations in infrared or shorter wavelengths, complicating the study. Furthermore, massive stars are rare. As a consequence, regions of massive star formation are generally located far away, thus requiring extremely high angular resolution observations. The required angular resolution to in detail study such regions has not been possible until the recent advent of large radioastronomical interferometer arrays, such as the Submillimetre Array (SMA) which began operations in 2003 (Ho et al. 2004).

With the help of state-of-the-art radio interferometers, a promising frontier toward early stages of high-mass star formation has emerged. In particular, the hot molecular core stage, which is characterized by a hot and dense envelope of gas and dust surrounding the newly formed massive protostar. At temperatures $\gtrsim 100\,K$, ice mantles on dust grains in the dense envelope around the central protostar sublimate, releasing high number densities of gas phase molecules (Jørgensen et al. 2020). Previous studies of hot cores have detected several complex organic molecules, such as CH_3CHO, C_2H_5OH, CH_3COCH_3, among others (Gorai et al. 2024). Their rotational transitions can then be studied in the millimeter to sub-millimeter wavelength range with interferometers such as SMA, in order to gain understanding of the protostar evolving within the envelope. From intensity spectra toward these high-mass star forming regions, spectral line features can be used as powerful diagnostic tools to study both the chemical and physical conditions of massive protostars.

1.2 The G028.37+00.07 star-forming region

G028.37+00.07 (G28.37) is a massive, $> 10^4\,M_\odot$, infrared dark cloud (IRDC) located at a distance of $5\,kpc$, studied by e.g. Carey et al. (1998), Simon et al. (2006a, 2006b), Rathborne et al. (2006), and Butler & Tan (2009). The dense molecular cloud is classified as an IRDC because of its high opacity filamentary structure, which can be observed as a dark region against the diffuse infrared background emission from our galactic plane. The high densities of molecular gas and dust found in this cloud are strong indications of showing an early view of clustered and massive star formation. G28.37 has been studied for its magnetic fields through polarization observations, allowing comparisons between the roles of gravity and magnetic field strengths in massive star formation (Liu et al. 2020).

G28.37 has also been studied as one of > 50 high-mass star forming regions in the SOFIA Massive Star Formation (SOMA) survey. The SOMA survey, led by Primary Investigator J. Tan, used the Faint Object infrared CAmera of the SOFIA telescope to observe these star forming regions in the range ~ 10 to $40\,\mu m$ (De Buizer et al. 2017). By compiling spectral energy distributions, physical properties were derived through fitting of theoretical protostellar models. Further, imaging using the Spitzer Space Telescope indicated G28.37 to be a highly clustered star forming region with the presence of > 5 protostellar objects. The brightest emission object was then estimated to contain a best fit of $\approx 13\,M_\odot$ protostellar mass with a surrounding $\approx 30\,M_\odot$ envelope mass (Telkamp et al. In prep.).

2 Observational data

2.1 The SMA line survey of G28.37

SMA is a radio interferometer consisting of eight antenna dishes, each with a diameter of 6 m, located at an altitude of over 4000 m in Hawaii (Ho et al. 2004). As a follow-up to the previous observations with SOFIA and the Spitzer Space Telescope, G28.37 was one of several clustered SOMA targets selected for follow-up spectroscopic observations using SMA. The SMA achieves 44 GHz broadband observations with a frequency resolution of 0.14 MHz, enabling an unbiased, broad and continuous spectral line survey toward these protostellar regions.

This study utilized SMA data (Project: 2022A-S006, PI: J. Tan) obtained toward the high-mass star-forming region G28.37, over a continuous bandwidth from 204 to 248 GHz. The data has been calibrated and reduced prior to this project using the python interface Pyuvdata, and the Common Astronomy Software Applications Package (CASA) (The CASA Team et al. 2022), where the spectral data was smoothed by a factor of 4. The reduced data is in the form of 3D data cubes (position-position-frequency), which can then be either split into maps of individual channels, or integrated along the frequency axis to produce spectra at different positions in the field of view. Additionally, the data has been corrected to account for the primary beam attenuation which lowers the antenna response in the outer regions of the field of view.

The 44 GHz bandwidth consists of 24 spectral windows, each with an individual bandwidth of $\approx$ 2 GHz, with a slight overlap in order to ensure continuous spectral data. For short, the spectral windows are referred to as 'Spw' followed by a numbering, and detailed information for all spectral windows can be found in Table 1. As such, the data is split into 24 individual data cubes, from which the continuum emission level was subtracted using the *uvcontsub* task in CASA, to only keep the emission from spectral lines. Two pairs of spectral windows are degenerate, Spw6 & Spw16 and Spw7 & Spw17, in the sense that they encompass the same frequency range. For the purpose of this study, only Spw6 and Spw17 have been kept for further analysis in order to ensure an approximately homogeneous sensitivity throughout the spectral range.

Table 1: Frequency range and synthesized beam size for all the 24 individual spectral windows of the SMA observations.

Spectral window	Frequency range [MHz]	Synthesized beam [$''\times''$]	Spectral window	Frequency range [MHz]	Synthesized beam [$''\times''$]
Spw0	203381−206118	4.91×3.78	Spw12	215836−218124	4.61×3.88
Spw1	205843−208131	3.92×3.46	Spw13	217848−220136	4.56×3.86
Spw2	207832−210119	4.81×3.70	Spw14	219831−222118	4.53×3.80
Spw3	209844−212132	4.71×3.66	Spw15	221843−224131	4.45×3.76
Spw4	211833−214120	4.68×3.63	Spw16	223832−226119	4.40×3.72
Spw5	213845−216133	4.63×3.62	Spw17	225845−228132	3.76×3.17
Spw6	223836−226123	4.44×3.46	Spw18	235835−238123	4.18×3.56
Spw7	225849−228136	4.35×3.43	Spw19	237848−240135	4.15×3.53
Spw8	227837−230124	4.35×3.39	Spw20	239837−242124	4.10×3.52
Spw9	229850−232137	4.27×3.36	Spw21	241849−244136	3.67×2.96
Spw10	231838−234126	4.29×3.35	Spw22	243837−246125	4.09×3.50
Spw11	233851−236139	3.45×3.07	Spw23	245850−248138	3.59×2.87

A continuum image was prepared by integrating only the spectral line-free emission in Spw20, which has a rest frequency of 240 919 MHz. The SMA observations have an average beam size of $3.9''$, leading to resolved physical scales of $\sim$20 000 AU in the image. The synthesized beam for each spectral window can be found in Table 1. Further, the noise level in the image was extracted from a manually selected polygon region, avoiding the strongest emission regions. From the intensity data in this region, the root mean square (rms, $\sigma_{\rm rms}$) deviation from zero was calculated, yielding $\sigma_{\rm rms} = 0.014$ K. From the continuum image, the positions of seven cores with emission above the $3\,\sigma_{\rm rms}$ noise level have been manually extracted, labeled Core 1 through 7, as can be seen in Figure 1. Regions with emission stronger than $3\,\sigma_{\rm rms}$ at the edge of the primary beam field of view were excluded. The cores are labeled starting from the strongest continuum emission region Core 1, and

then increasing with distance from Core 1. A summary of properties of all seven cores is presented in Table 4.

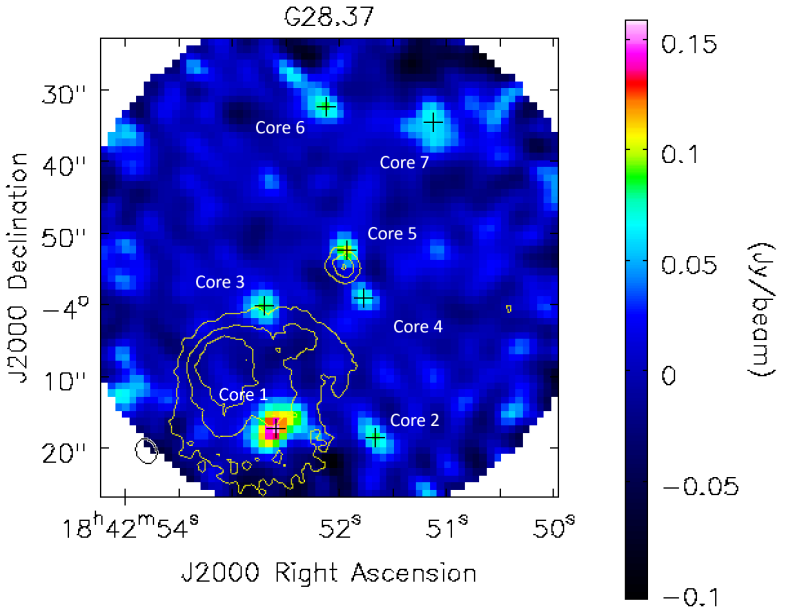


Figure 1: Map of the continuum emission observed with SMA toward G28.37 for spectral window 20, i.e. at a rest frequency of 240 919 MHz. The synthesized beam can be seen depicted as the small black ellipse in the lower left corner. Yellow contours are 5, 10, and 15 σ_{RMS} levels from 31 µm observations with SOFIA from the SOMA survey (Telkamp et al. In prep.). Seven regions of strong emission are labeled as Core 1 through 7, marked with black crosses.

2.2 Spectrum extraction

In order to analyze the spectral data from the seven identified cores, the spectra must be extracted from each core and from each separate spectral cube. To handle the spectral cubes, this project used CASA. For the spectrum extraction, a Python script was developed in order to automatically call CASA functions and perform the task for several data cubes and cores. The script takes the core coordinates as input, and is then run for all spectral cubes. The spectra are then extracted from beam-sized apertures centered on the central pixel in each core. To define these beam-sized apertures, the average synthesized beam $\theta_{\mathrm{maj}} \times \theta_{\mathrm{min}} = 4.29'' \times 3.50''$ was used for all spectral windows. By calling the CASA function *imhead()* with the argument *hdkey="cdelt1"*, the pixel size is directly read from each spectral cube and then the beam in arcseconds is converted into the corresponding region expressed in pixels from which the spectrum is extracted.

A few preparatory steps were taken to prepare the spectra for further analysis. First, in order to concatenate the 24 spectra extracted toward each core, Grenoble Image and Line Analysis Software (GILDAS) (iram.fr/IRAMFR/GILDAS) was used for its *stitch /weight equal* function, which correctly interprets the frequency channels and averages the overlapping channels together. Then, in the concatenated 44 GHz spectra for each core, the average noise level was measured by manually selecting line-free channels in the spectra. The root mean square deviation σ_{rms} per channel was calculated, and the resulting values for the spectra extracted toward each core are presented in Table 4.

2.3 Spectral line analysis

By matching observed spectral emission features to known rotational transition rest frequencies of molecules, the origin of spectral emission can be determined. For this, accurate spectroscopic information for a multitude of molecules is available from laboratory experiments and quantum chemical calculations through databases such as the Cologne Database for Molecular Spectroscopy (CDMS) (Müller et al. 2001, 2005) and Jet Propulsion Laboratory (JPL) (Pickett et al. 1998). Still, the

procedure comes with a number of intricacies. Although the observed spectral emission is caused by discrete energy transitions, the measured spectral line features will be spread out over a broader energy range. The line profile of a spectral feature is a result of the physical conditions of the gas emitting the signal.

A major contributor to this appearance of astrophysical spectra is the Doppler effect. The frequency of light is shifted as a result of the relative line-of-sight motion of the emitting material and the observer. The line profile therefore contains information about the local velocity structure of the emitting gas, where the velocity distribution will lead to a broadening of the line profile. In a thermodynamic system where the velocity distribution is entirely due to thermal motion, the line will appear as a Gaussian distribution around the center frequency, with a full width half maximum (FWHM) $(\Delta\nu)_{\rm FWHM}$ directly relating to the one dimensional velocity dispersion σ_v, as seen in Equation 1 (Draine 2011),

$$(\Delta\nu)_{\rm FWHM} = \sqrt{8\ln 2}\sigma_v. \quad (1)$$

In less ideal situations, the effects of opacity and absorption in the intermediate matter can affect the line profile. In high opacity situations, at frequencies where the observed medium has become optically thick, only signal emitted from the surface of the observed region will escape the gas and reach us. The spectral line signal will then not be representative of the full physical conditions and abundance of the emitting molecule, and might deviate noticeably from the expected Gaussian distribution. Similarly, the process of self-absorption, in which emitted light is reabsorbed before escaping the source, can also cause deviations from a Gaussian profile. A typical situation is that of a cooler, intermediate gas component absorbing the emission from warmer gas. Self-absorption is a resonance process, and is most probable at the line centre, leading to incorrect line intensities and absorption features in the underlying Gaussian profile (Gudimenko et al. 2012).

Additionally, any systematic motion of gas such as infall, outflow or rotation will lead to alterations of the line profile, and can serve as a tool for analysis of the observed kinematic structure. Species tracing outflowing material may exhibit a significantly different velocity shift or line width than those of other emission lines observed toward the same source. The conditions in an outflow may not allow as high upper state energies $E_{\rm up}$ for the species as can be present closer to the central protostar, but can still be an important effect for lower $E_{\rm up}$ transitions. CO is known to trace such outflows, but as studied by Prasanta et al. (2023), more complex molecules such as CH_3OH may also be common in outflows from hot molecular cores.

2.4 Vlsr correction

The peculiar motion of the individual cores and their environment will lead to a shift in all spectral emission through the Doppler effect. The relation between the rest frequency of the transition ν_0 and the observed frequency ν shifted by the line-of-sight velocity relative to the local standard of rest $v_{\rm LSR}$ is then $\nu = \nu_0(1 - v_{\rm LSR}/c)$. A correction of $v_{\rm LSR} = 76.89\,{\rm km\ s^{-1}}$ was initially applied to the spectra of all cores based on the results of recent studies of G28.37 by Light et al. (In prep.), which allowed the identification of some prominent spectral lines. The velocity correction was then refined by measuring the velocity shift of these initially identified emission lines. Four CH_3OH transitions were chosen, and are listed in Table 2, based on their occurrence in all cores. Several high $E_{\rm up}$ lines were intentionally included as they are expected to be least affected by eventual outflows.

Table 2: List of lines used in automatic $v_{\rm LSR}$ calculations for the seven cores.

Species	Transition $[J'_{K'_a,K'_b} - J''_{K''_a,K''_b}]$	Frequency [GHz]	$E_{\rm up}$ [K]	$A_{\rm ij}$ [s^{-1}]
E-CH_3OH	$8_{-1,0} - 7_{0,0}$	229.7588	81.20	4.19E-5
	$5_{0,0} - 4_{0,0}$	241.7002	40.04	6.03E-5
A-CH_3OH	$5_{3,0} - 4_{3,0}$	241.8327	84.62	3.86E-5
	$5_{2,-0} - 4_{2,-0}$	241.8423	72.53	5.11E-5

To automatically update the velocity estimate for each core, a script was developed which attempts to fit Gaussian line profiles to the data within 10 MHz windows centered on the rest frequencies of the selected transitions. For all N successful Gaussian fits, the velocity shift $v_{\rm i}$ of the Gaussian mean with its corresponding standard deviation error $\sigma_{v,{\rm i}}$ was used to calculate an inverse variance weighted average $v_{\rm avg}$ according to Equation 2, together with the standard uncertainty of the weighted mean $\sigma_{v,{\rm avg}}$ according to Equation 3. In both equations, the weights are defined as $w_{\rm i} = 1/\sigma^2_{v,{\rm i}}$. The velocity shifts were then used to define new $v_{\rm LSR}$ values with uncertainties for each core, which are presented in Table 4.

$$v_{\rm avg} = \frac{\sum_{\rm i} w_{\rm i} v_{\rm i}}{\sum_{\rm i} w_{\rm i}} \tag{2}$$

$$\sigma_{v,{\rm avg}} = \sqrt{\frac{\sum_{\rm i} w_{\rm i}(v_{\rm i} - v_{\rm avg})^2}{\sum_{\rm i} w_{\rm i}(N-1)}} \tag{3}$$

2.5 Cleaning artifacts

All seven spectra overlaid can be seen in the top panel of Figure 2, where a periodic nature of spike artifacts can be seen marked with red dashed lines, appearing every 12 GHz at the overlap between pairs of spectral windows.

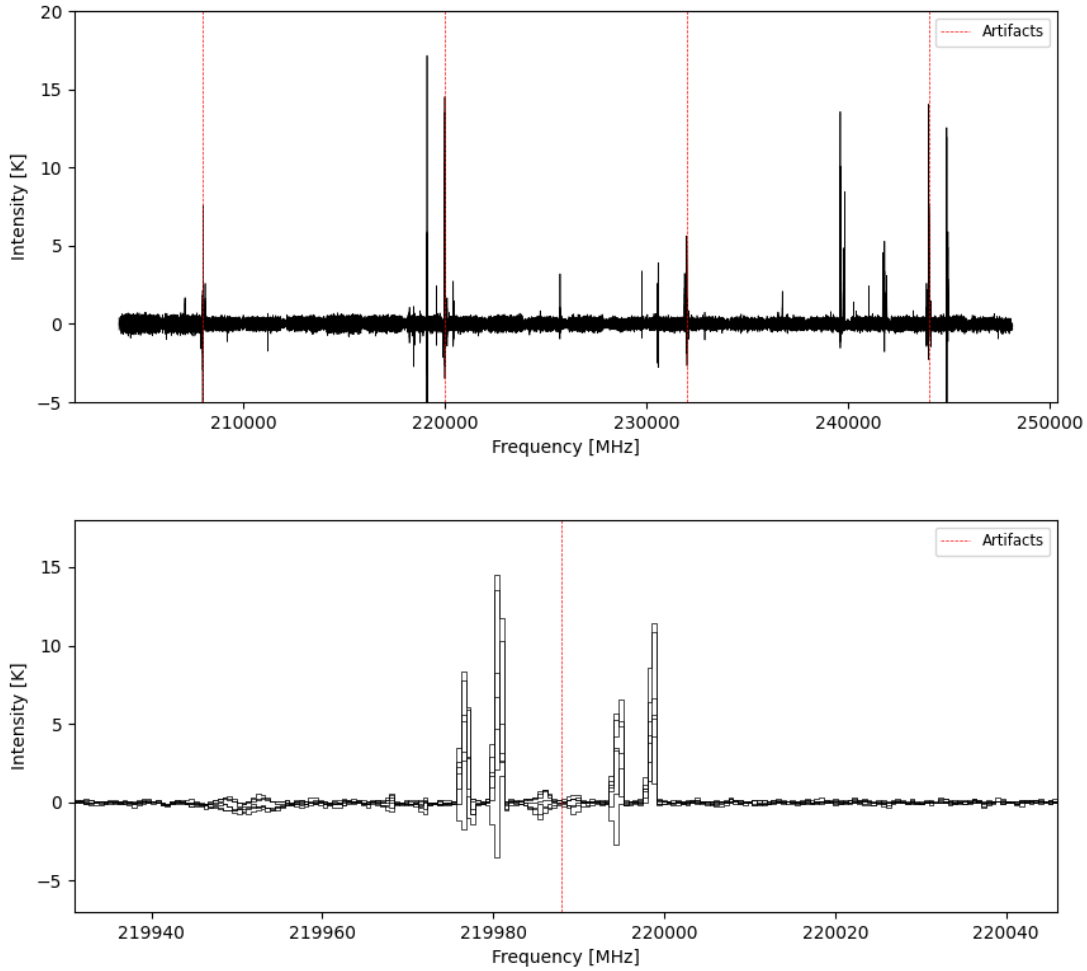


Figure 2: In the top panel, an overview of the periodic nature of spikes can be seen marked with red dashed lines, appearing in the overlapped spectra from all cores. In the bottom panel, a zoomed-in view is showing the profile of one of the spike features, which was similar for all four regions marked in the top panel.

These spikes appear near 208 GHz, 220 GHz, 232 GHz, and 244 GHz. A zoomed in view of the artifact at 220 GHz can be seen in the bottom panel of Figure 2, which shows the general appearance of

the artifacts. These periodic features are abnormally strong and narrow, often covering only $2-3$ channels, and overall do not resemble other Gaussian-like emission lines. As a consequence, any potential other spectral emission measured in these channels will not be recoverable. Based on this, all channels containing artifacts were manually masked and removed from further analysis of the spectra.

3 Methods of analysis

3.1 Automatic peak detection

As a first step in an automated line identification algorithm, a Python script was developed to extract the frequencies of all peaks in a spectrum. The script uses the in-spectrum noise values measured for each core from Table 4 in order to define a $5\,\sigma_{\rm rms}$ signal-to-noise threshold for the intensities of the found peaks. Further, to avoid multiple detections of a single emission line, the script excludes weaker peaks within $5\,\mathrm{km\ s^{-1}}$ of a stronger peak. As a first result, this script provides the spectral line density of the spectrum, which serves as a simple measure of chemical richness. A visualization of the detected peaks in the spectrum toward Core 1 can be seen in Figure 3.

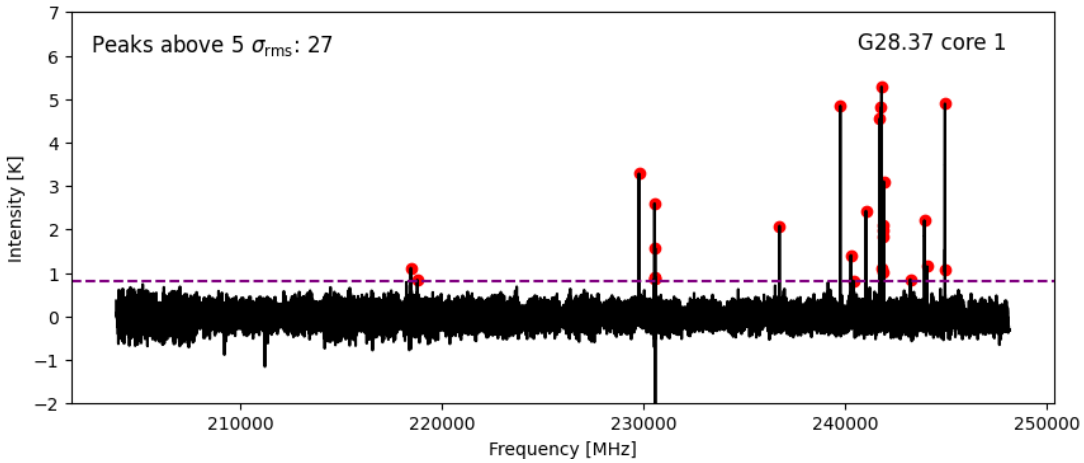


Figure 3: Automatically detected peaks in the spectrum extracted toward Core 1.

3.2 Automatic line identification

To continue the line identification process, a Python script was developed using the *linelists.cdms* module of the *Astroquery* Python package (Ginsburg et al. 2019). Through *linelists.cdms*, the CDMS database is queried for rotational transitions to be cross-matched with result of the peak finding script. The script queries the database for transitions from a list of species based on the ones previously detected toward hot molecular cores in the work of Prasanta et al. (2023), and also species detected by Light et al. (In prep.). Also, based on the same studies, an upper limit for the query of $E_{\rm up} < 300\,\mathrm{K}$ was used. The full list of species can be found in Table 3. Finally, all cross-matched transitions in the database whose rest frequency lies within a $2\,\mathrm{km\ s^{-1}}$ shift in velocity space from a found peak are saved.

For all lines matched to a peak, another Python script was developed which attempts to fit Gaussian profiles to the lines using the Python package *scipy.optimize*. Based on the widths of observed line profiles in the preliminary line identification, an initial 10 MHz window around each peak frequency is extracted for the Gaussian fitting. From the Gaussian fits, the fit amplitude $I_{\rm max}$, mean velocity shift $v_{\rm i}$ of the peak from the rest transition, and standard deviation σ of the Gaussian distribution are directly extracted. This standard deviation of the velocity distribution can then be used to calculate the FWHM of each line according to Equation 1. Further, the integrated intensity of each line is calculated according to Equation 4,

$$W = \frac{\sqrt{2\pi}}{\sqrt{8\ln 2}} \times \mathrm{FWHM} \times I_{\rm max}. \tag{4}$$

Then, for all successful fits, a χ^2 test for the goodness of the fit as in Riener et al. (2019) was calculated using Equation 5,

$$\chi^2 = \sum_{i=1}^{N} \frac{(y_i - Y_i)^2}{\sigma^2_{\rm rms}}. \tag{5}$$

Here, $y_i - Y_i$ is the residual of the Gaussian fit and the data, scaled by the rms squared $\sigma^2_{\rm rms}$, and summed over all N data points in the fitting window. The χ^2 value was used together with the uncertainties of the fit Gaussian parameters $I_{\rm max}$, $v_{\rm i}$, and σ, in order to detect poor Gaussian fits. For these poor fits, no Gaussian parameters or any of the metrics calculated using them will be presented as results.

Table 3: Species queried from the CDMS database in the automatic line identification script

Diatomic	Three- to five-atomic	COMs
CO	$H^{13}CO+$	CH_3OH
$C^{17}O$	$HN^{13}C$	$^{13}CH_3OH$
$C^{18}O$	SO_2	CH_3CN
SO	$^{34}SO_2$	CH_3OCH_3
^{34}SO	H_2S	C_2H_3CN
SiO	H_2CO	C_2H_5OH
CS	H_2CN	C_2H_5CN
$C^{34}S$	H_2CS	
^{13}CS	HNCO	
NS	HC_3N	
	OCS	

3.3 Column density calculations

When the gas density is much higher than the critical density, collisions dominate the excitation over radiative processes, and then the emitting gas is assumed to be in local thermal equilibrium (LTE). In LTE conditions, the number densities of molecules in the upper and lower energy states corresponding to a given rotational transition is assumed to be described by a single excitation temperature $T_{\rm ex}$. The excitation temperature is then also assumed to correspond to the kinetic temperature $T_{\rm K}$ of the gas. For Gaussian line profiles, the integrated intensity of the line together with information about the specific energy transition can be used to estimate the total column density of the emitting gas, $N_{\rm tot}$, as presented in Equation 6 (Bonfand 2019),

$$N_{\rm tot} = \frac{8\pi k_{\rm b}\nu^2}{b_{\rm ff} h c^3} \frac{Q_{\rm rot}(T_{\rm ex})}{A_{\rm ul} g_{\rm u}} e^{\frac{E_{\rm u}}{k_{\rm b} T_{\rm ex}}} W. \tag{6}$$

Here, $k_{\rm b}$ is the Boltzmann constant, h is the Planck constant, c is the speed of light, $Q_{\rm rot}$ is the rotational partition function at the excitation temperature $T_{\rm ex}$, and $b_{\rm ff}$ is the beam filling factor. For the sake of this work, $b_{\rm ff}$ is assumed to be equal to one, implying that the source is fully covered by the beam. $A_{\rm ul}$ is the Einstein coefficient of the transition, $g_{\rm u}$ is the degeneracy of the upper energy state, $E_{\rm u}$ is the upper energy level of the transition, and finally W is the integrated intensity of the Gaussian line profile. For each identified transition, $Q_{\rm rot}(T_{\rm ex})$, $E_{\rm u}$, $A_{\rm ul}$, and $g_{\rm u}$ are automatically saved from the CDMS query and used for the calculations. W is calculated from Gaussian fits to the identified lines, and finally $T_{\rm ex}$ which also in turn determines $Q_{\rm rot}(T_{\rm ex})$ remains as the only unknown parameter.

As the column density is dependent on the line-of-sight size of the observed cloud, the relative abundances as compared to the column density of Hydrogen nuclei $N_{\rm H}$ are often used as a metric to compare the chemical composition of molecular cores. Based on the assumption of optically thin thermal emission from dust, $N_{\rm H}$ for a hot molecular core can be estimated from the observed millimeter continuum flux F_ν through Equation 7 (Law et al. 2022, Gorai et al. 2024),

$$N_{\rm H} = 369 \frac{F_\nu \lambda^3}{\Omega \kappa_\nu \mu_{\rm H}} \times (e^{\frac{0.111}{T_{\rm d}\lambda}} - 1). \tag{7}$$

Here, the equation is normalized such that F_ν is the total integrated flux in Jy from within a solid angle Ω in arcseconds, at wavelength λ normalized to $\lambda/1.3\,\mathrm{mm}$=1. κ_ν is the dust absorption coefficient used by Law et al. (2022), $\kappa_\nu = 0.006\,38\,\mathrm{cm^2 g^{-1}}$. Finally, the dust temperature T_d is normalized to $T_\mathrm{d}/100\,\mathrm{K} = 1$.

4 Results

4.1 Parameters of the cores

Coordinates for the seven cores extracted from the continuum map can be found in Table 4, together with the measured integrated flux F_ν from within a beam-sized region of $4.29'' \times 3.50''$ centered on the respective cores. F_ν was then used to calculate an estimated column density of Hydrogen nuclei using Equation 7, where the dust temperature T_d was set to 100 K, similar to previous studies by Gorai et al. (2024). Furthermore, from the spectra extracted toward each core, the in-spectrum noise σ_rms and calculated velocity shift v_LSR are also presented. The observed variation in σ_rms could in part be a result of the primary beam correction from the preliminary data reduction, as the central cores show significantly lower noise levels. As for the velocity shifts, all cores have v_LSR values within a $1\,\mathrm{km\ s^{-1}}$ range, showing little signs of relative motion between the cores.

Core 1 is the brightest of all cores by a factor of ~ 2, which then translates into the same relation between the cores regarding the mass surface density Σ and the Hydrogen column density N_H. Compared to the hot molecular cores observed by Gorai et al. (2024), the G28.37 surface mass and column densities are lower by about an order of magnitude. This difference can partially be explained by the smaller beam size of the observational data analyzed by Gorai et al. (2024). At similar distances, the higher angular resolution probes the gas conditions closer to the central protostar, and is then expected to result in larger densities.

Table 4: Coordinates of the seven cores identified in G28.37, along with information extracted from the continuum image and from the spectra.

Name	R.A. J2000	Dec. J2000	F_ν [Jy]	Σ [g cm^{-2}]	N_H [cm^{-2}]	σ_rms [K]	v_LSR [km s^{-1}]
Core 1	18:42:52.598	-4.00.17.072	1.164	3.086	13.19×10^{23}	0.1611	76.42 ± 0.12
Core 2	18:42:51.649	-4.00.17.830	0.5736	1.521	6.50×10^{23}	0.1636	77.00 ± 0.17
Core 3	18:42:52.674	-4.00.00.220	0.6378	1.691	7.23×10^{23}	0.1003	76.06 ± 0.16
Core 4	18:42:51.763	-3.59.58.516	0.4713	1.250	5.34×10^{23}	0.0909	76.97 ± 0.31
Core 5	18:42:51.890	-3.59.52.457	0.6522	1.729	7.39×10^{23}	0.0906	76.40 ± 0.23
Core 6	18:42:52.143	-3.59.32.196	0.5985	1.587	6.78×10^{23}	0.1424	76.01 ± 0.20
Core 7	18:42:51.119	-3.59.34.244	0.5521	1.464	6.26×10^{23}	0.1584	76.65 ± 0.24

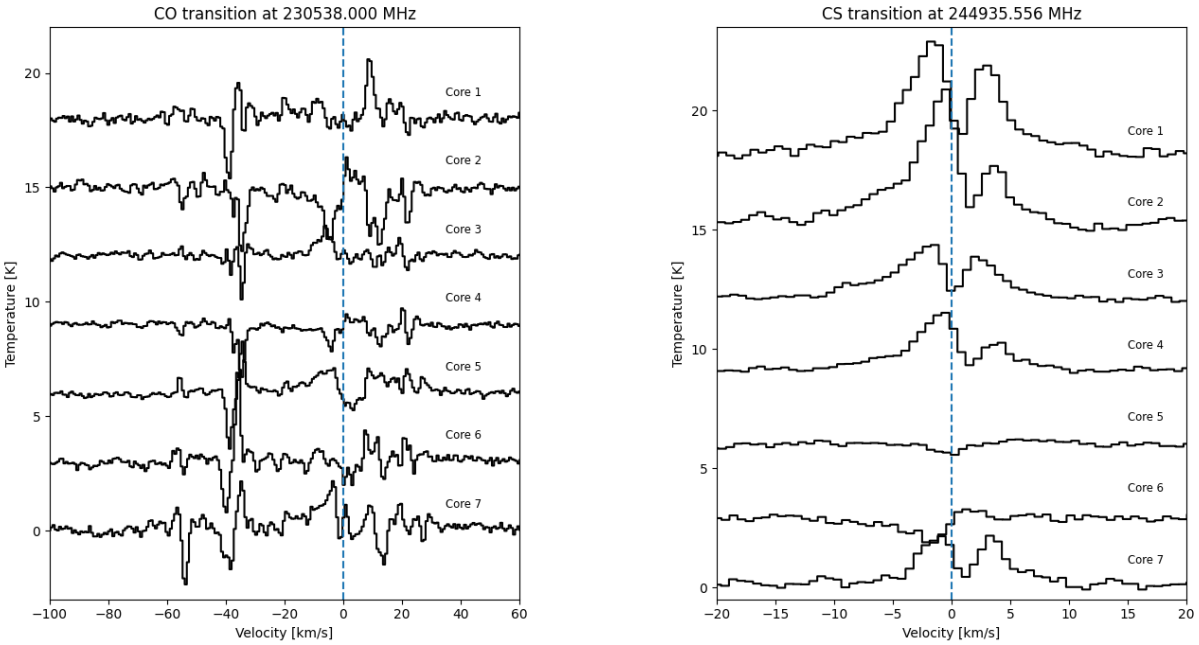
4.2 Species detected

A complete list of species tentatively detected using the automatic line identification script can be found in Table 5. Species written inside parentheses are thought to be misidentifications from overlapping transitions, in such cases, the simpler species are expected to be the true cause of the signal as they are most likely to be detected toward this type of source. Notably, emission from CH_3OH and $C^{34}S$ was detected toward all cores, which suggests some level of chemical similarity. Other species such as ^{13}CS, OCS, H_2CO, HNCO and HC_3N were only detected toward one or two cores each. Weak emission from H_2CO was observed only in Core 4, and a strong line with self-absorption in Core 5. Thus, the automatic detections point to a possibly varied chemical inventory of the cores, despite them forming close-by in a clustered region.

Table 5: List of transitions detected toward the seven cores. Uppercase letters mark automatic detections, with "G" having acceptable Gaussian fits, and "N" no acceptable Gaussian fits. Additionally, non-detections in other cores are labeled with lowercase letters as "a" for absorption lines, and "d" for missed detections caused by distorted line profiles.

Species	Transition	ν_0	$E_{\rm up}$	Core 1	Core 2	Core 3	Core 4	Core 5	Core 6	Core 7
	$(J'_{K'_a,K'_b} - J''_{K''_a,K''_b})$	[MHz]	[K]							
CO	$2-1$	230538.000	16.60	d	N	d	N	d	d	d
(C_2H_5CN)	$12_{8,5} - 13_{7,6}$	230524.102	104.6	d	d	N	N	d	d	N
(C_2H_5OH)	$13_{3,10,2} - 12_{12,3,9,2}$	230523.689	88.20	d	d	N	N	d	N	N
$C^{17}O$	$2_2 - 1_3$	224713.533	16.18		G	G				
	$2_3 - 1_3$	224714.187	16.18		N	N				
	$2_2 - 1_2$	224714.743	16.18		N	N				
	$2_3 - 1_2$	224715.310	16.18			N				
CS	$5_0 - 4_0$	244935.556	35.27	d	N	N	N	a	a	d
$C^{34}S$	$5_0 - 4_0$	241016.089	34.70	N	G	G	G	G	G	G
^{13}CS	$5_0 - 4_0$	231220.685	33.29					G		
(C_2H_5OH)	$10_{3,7,1} - 10_{2,9,0}$	231220.821	118.4					N		
OCS	$19-18$	231060.993	110.9					G		
	$20-19$	243218.036	122.6	N						
H_2CO	$3_{1,3} - 2_{1,2}$	211211.468	32.06	a	a	a	G	a	a	a
	$3_{1,2} - 2_{1,1}$	225697.775	33.45				a	N	d	
HNCO	$11_{0,11} - 10_{0,10}$	241774.032	69.62	N		N				
HC_3N	$26-25$	236512.788	153.2						G	
H_2CS	$7_{1,7} - 6_{1,6}$	236727.020	58.62	G	G	G		N		G
	$7_{0,7} - 6_{0,6}$	240266.872	46.14	G		G				
	$7_{3,5} - 6_{3,4}$	240393.037	164.6	G						
	$7_{3,4} - 6_{3,3}$	240393.761	164.6	N						
	$7_{1,6} - 6_{1,5}$	244048.504	60.03	G		G				
CH_3OH	$8_{1,8,2} - 7_{0,7,1}$	229758.756	89.10	G	a	G		G	G	G
	$5_{1,5,0} - 4_{1,4,0}$	239746.219	49.06	G	G	G		G	G	G
	$5_{0,5,1} - 4_{0,4,1}$	241700.159	47.93	G	G	G	G	G	G	G
	$5_{1,5,2} - 4_{1,4,2}$	241767.234	40.39	N	a	G	a	G	d	G
(C_2H_5OH)	$14_{2,13,2} - 13_{13,2,12,2}$	241768.799	92.56	N						N
	$5_{0,5,0} - 4_{0,4,0}$	241791.352	34.82	N	a	G	a	G	d	G
	$5_{3,3,0} - 4_{3,2,0}$	241832.718	84.62	G	G	G	G	G	G	N
	$5_{3,2,0} - 4_{3,1,0}$	241833.106	84.62	N	N	N	N	N	N	N
	$5_{2,4,0} - 4_{2,3,0}$	241842.284	72.53	G	G	G		G	G	G
	$5_{3,3,1} - 4_{3,2,1}$	241843.604	82.53			N		G	G	N
	$5_{1,4,1} - 4_{1,3,1}$	241879.025	55.87	G		G		G	G	G
	$5_{2,3,0} - 4_{2,2,0}$	241887.674	73.53	G						
	$5_{2,3,2} - 4_{2,2,2}$	241904.147	60.72	G		G		G	G	G
	$5_{2,4,1} - 4_{2,3,1}$	241904.643	57.07	N		N		N	N	N
	$5_{1,4,0} - 4_{1,3,0}$	243915.788	49.66	G	G	d	G			G

CO was only detected toward two cores, despite being the most abundant observable molecule in the interstellar medium. This is likely a result of widespread CO emission with varying velocity components, which is not correctly interpreted by the SMA in its compact configuration and leads to a distorted signal. Surrounding the frequency of the CO(2-1) transition, a mixture of emission and absorption was found, with no detected Gaussian line profile, as can be seen in the left panel of Figure 4. Further, CS was detected toward cores 2, 3, and 4, and strong signal was also seen toward cores 1 and 7. However, in the latter cases, the molecule was not detected because of a strong self-absorption feature in the line profile which can be seen in the right panel of Figure 4.



(a) (b)

Figure 4: Spectra from all cores in the vicinity of the CO(2-1) transition, left panel, and the CS(5-4) transition, right panel, marked with blue dashed lines. The spectra are shifted vertically to show the individual line profiles.

4.3 Gaussian fitting results

Following the line identification, Gaussian fitting was attempted for all lines. For the lines which were able to be automatically fitted with Gaussian profiles, an overview of the line profiles for each core can be seen in Figures 5, 6, 7, and 8. The lines profiles are all displayed in velocity space relative to their respective rest frequencies, such that individual velocity shifts of lines can be seen. Further, the panels all use a common velocity axis, such that differences in line widths throughout the cores can be seen.

Further details of the fitted lines and the respective transitions are provided in Tables 6 and 7. CO, CS, HNCO, and a number of lines from other detected species could not be fitted with acceptable Gaussian lines, and thus column densities and abundances were not estimated. Again, this was caused by a distorted signal for CO, and the self-absorption feature in the CS line. In the case of HNCO, the line was blended with the $E_{\rm up} = 40.39\,{\rm K}$ CH_3OH line, and as such, neither lines showed Gaussian profiles.

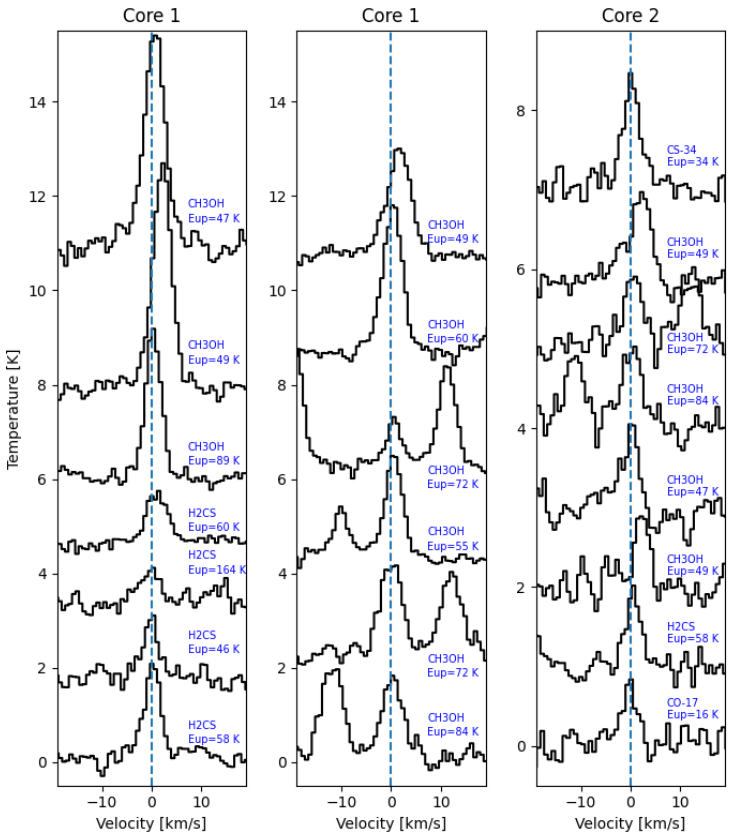


Figure 5: Overview of spectral line profiles observed toward Core 1 & Core 2.

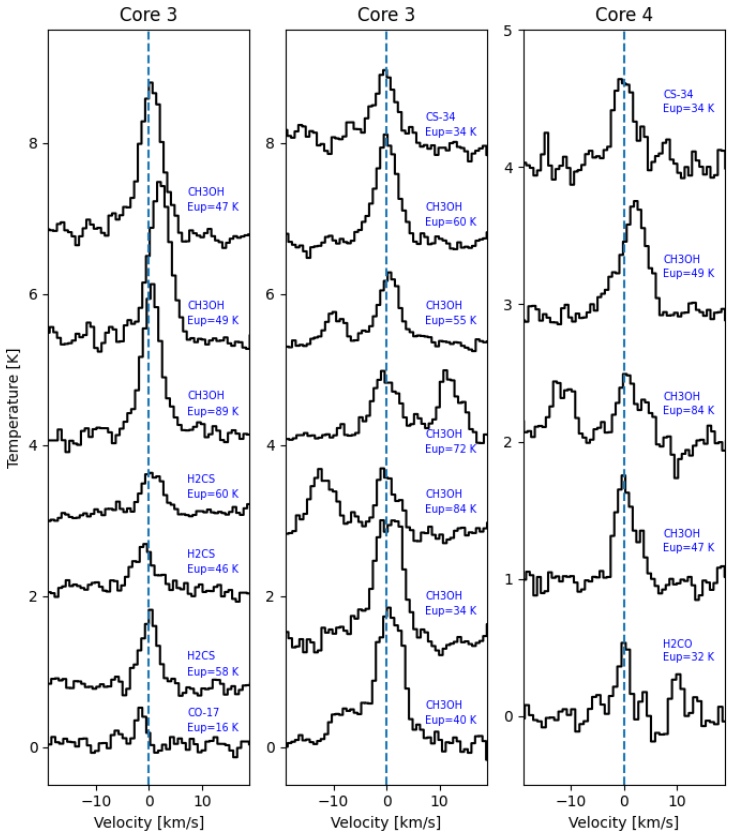


Figure 6: Overview of spectral line profiles observed toward Core 3 & Core 4.

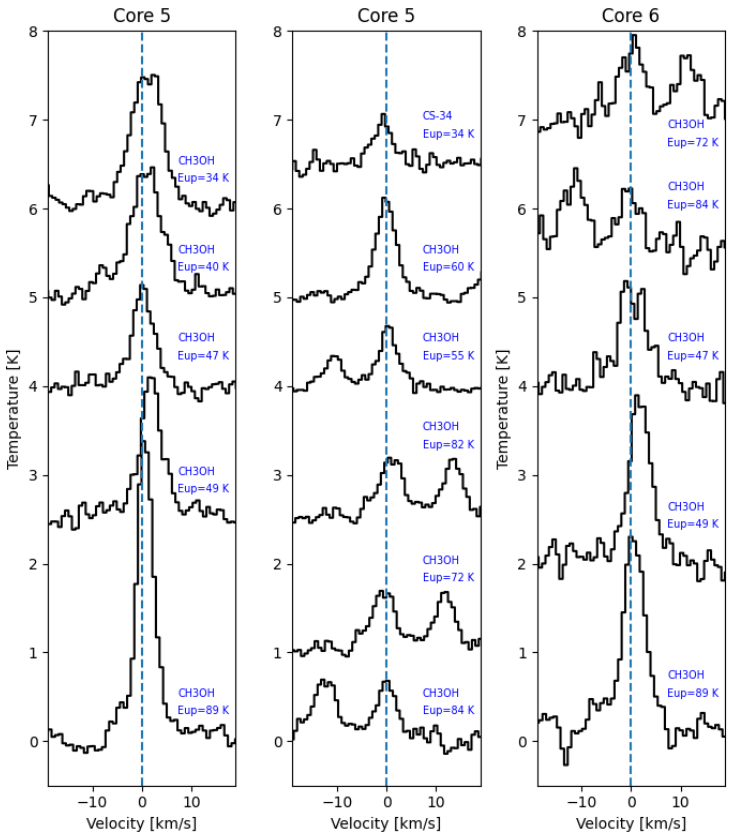


Figure 7: Overview of spectral line profiles observed toward Core 5 & Core 6.

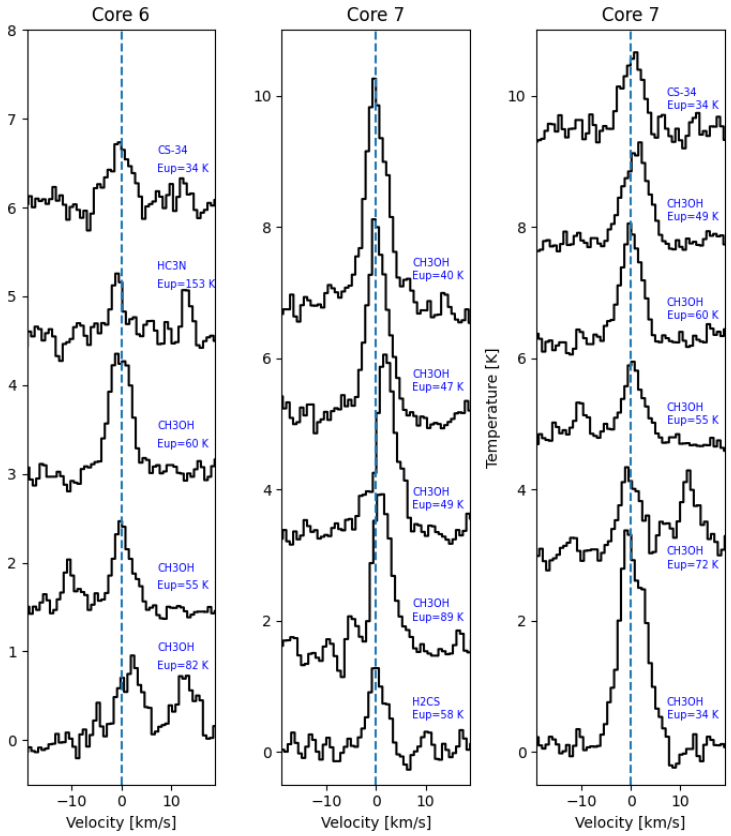


Figure 8: Overview of spectral line profiles observed toward Core 6 & Core 7.

Table 6: Results from Gaussian fitting of the automatically identified lines.

Species	Transition $(J'_{K'_a,K'_b} - J''_{K''_a,K''_b})$	ν_0 [MHz]	$E_{\rm up}$ [K]	FWHM [km s^{-1}]	W [K km s^{-1}]	$N_{\rm tot}$ [cm^{-2}]	Abundance $N_{\rm tot}/N_{\rm H}$
			Core 1				
H_2CS	$7_{1,7} - 6_{1,6}$	236727.020	58.62	5.991	12.8	2.64×10^{15}	2.49×10^{-9}
	$7_{0,7} - 6_{0,6}$	240266.872	46.14	6.340	8.27	4.19×10^{15}	4.12×10^{-9}
	$7_{3,4} - 6_{3,3}$	240393.037	164.6	8.294	6.35	6.37×10^{15}	4.22×10^{-9}
	$7_{1,6} - 6_{1,5}$	244048.504	60.03	7.554	8.75	1.79×10^{15}	1.68×10^{-9}
CH_3OH	$8_{1,8,2} - 7_{0,7,1}$	229758.756	89.10	5.600	19.1	6.58×10^{16}	5.99×10^{-8}
	$5_{1,5,0} - 4_{1,4,0}$	239746.219	49.06	6.473	32.7	8.24×10^{16}	8.57×10^{-8}
	$5_{0,5,1} - 4_{0,4,1}$	241700.159	47.93	7.229	35.0	8.28×10^{16}	8.65×10^{-8}
	$5_{3,3,0} - 4_{3,2,0}$	241832.718	84.62	7.926	14.5	8.74×10^{16}	8.07×10^{-8}
	$5_{2,4,0} - 4_{2,3,0}$	241842.284	72.53	7.770	16.7	6.50×10^{16}	6.25×10^{-8}
	$5_{1,4,1} - 4_{1,3,1}$	241879.025	55.87	5.719	12.7	3.40×10^{16}	3.46×10^{-8}
	$5_{2,3,0} - 4_{2,2,0}$	241887.674	72.53	5.348	5.26	2.04×10^{16}	1.97×10^{-8}
	$5_{2,3,2} - 4_{2,2,2}$	241904.147	60.72	6.657	21.5	7.19×10^{16}	7.19×10^{-8}
	$5_{1,4,0} - 4_{1,3,0}$	243915.788	49.66	7.842	18.2	4.54×10^{16}	4.71×10^{-8}
			Core 2				
$C^{17}O$	$2_2 - 1_3$	224713.533	16.18	5.333	4.23	1.13×10^{18}	2.17×10^{-6}
$C^{34}S$	$5_0 - 4_0$	241016.089	34.70	6.241	8.71	3.26×10^{14}	5.87×10^{-10}
H_2CS	$7_{1,7} - 6_{1,6}$	236727.020	58.62	5.011	4.23	1.02×10^{15}	1.95×10^{-9}
CH_3OH	$5_{1,5,0} - 4_{1,4,0}$	239746.219	49.06	4.449	4.59	1.16×10^{16}	2.44×10^{-8}
	$5_{0,5,1} - 4_{0,4,1}$	241700.159	47.93	5.501	5.59	1.32×10^{16}	2.81×10^{-8}
	$5_{3,3,0} - 4_{3,2,0}$	241832.718	84.62	7.543	7.71	4.66×10^{16}	8.73×10^{-8}
	$5_{2,4,0} - 4_{2,3,0}$	241842.284	72.53	6.712	6.44	2.50×10^{16}	4.88×10^{-8}
	$5_{1,4,0} - 4_{1,3,0}$	243915.788	49.66	5.891	6.02	1.50×10^{16}	3.17×10^{-8}
			Core 3				
$C^{17}O$	$2_2 - 1_3$	224713.533	16.18	3.123	1.72	4.59×10^{17}	7.93×10^{-7}
$C^{34}S$	$5_0 - 4_0$	241016.089	34.70	7.703	7.25	2.71×10^{14}	4.40×10^{-10}
H_2CS	$7_{1,7} - 6_{1,6}$	236727.020	58.62	5.380	5.32	1.10×10^{15}	1.89×10^{-9}
	$7_{0,7} - 6_{0,6}$	240266.872	46.14	6.532	3.66	1.86×10^{15}	3.33×10^{-9}
	$7_{1,6} - 6_{1,5}$	244048.504	60.03	6.481	3.75	7.67×10^{14}	1.31×10^{-9}
CH_3OH	$8_{1,8,2} - 7_{0,7,1}$	229758.756	89.10	6.521	13.1	4.54×10^{16}	7.54×10^{-8}
	$5_{1,5,0} - 4_{1,4,0}$	239746.219	49.06	6.778	14.8	3.74×10^{16}	7.11×10^{-8}
	$5_{0,5,1} - 4_{0,4,1}$	241700.159	47.93	7.216	14.9	3.52×10^{16}	6.71×10^{-8}
	$5_{1,5,2} - 4_{1,4,2}$	241767.234	40.39	8.996	17.3	3.86×10^{16}	7.54×10^{-8}
	$5_{0,5,0} - 4_{0,4,0}$	241791.352	34.82	9.995	18.5	3.69×10^{16}	7.33×10^{-8}
	$5_{3,3,0} - 4_{3,2,0}$	241832.718	84.62	7.680	6.36	3.84×10^{16}	6.48×10^{-8}
	$5_{2,4,0} - 4_{2,3,0}$	241842.284	72.53	8.570	7.35	2.86×10^{16}	5.01×10^{-8}
	$5_{1,4,1} - 4_{1,3,1}$	241879.025	55.87	6.278	5.72	1.53×10^{16}	2.83×10^{-8}
	$5_{2,3,2} - 4_{2,2,2}$	241904.147	60.72	7.050	10.0	3.34×10^{16}	6.10×10^{-8}
			Core 4				
$C^{34}S$	$5_0 - 4_0$	241016.089	34.70	7.040	4.78	1.79×10^{14}	3.92×10^{-10}
H_2CO	$3_{1,3} - 2_{1,2}$	211211.468	32.06	3.826	2.16	2.18×10^{14}	5.52×10^{-10}
CH_3OH	$5_{0,5,1} - 4_{0,4,1}$	241700.159	47.93	6.114	4.50	1.07×10^{16}	2.75×10^{-8}
	$5_{3,3,0} - 4_{3,2,0}$	241832.718	84.62	9.795	4.31	2.60×10^{16}	5.94×10^{-8}
	$5_{1,4,0} - 4_{1,3,0}$	243915.788	49.66	7.322	5.61	1.40×10^{16}	3.59×10^{-8}

Continued

Table 7: Results from Gaussian fitting of the automatically identified lines *continued*.

Species	Transition $(J'_{K'_a,K'_b} - J''_{K''_a,K''_b})$	ν_0 [MHz]	$E_{\rm up}$ [K]	FWHM [kms^{-1}]	W [K kms^{-1}]	$N_{\rm tot}$ [cm^{-2}]	Abundance
			Core 5				
$C^{34}S$	$5_0 - 4_0$	241016.089	34.70	6.875	3.45	1.29×10^{14}	2.04×10^{-10}
^{13}CS	$5_0 - 4_0$	231220.685	33.29	7.347	3.38	1.35×10^{14}	2.15×10^{-10}
OCS	$19 - 18$	231060.993	110.9	5.364	3.39	2.84×10^{15}	3.51×10^{-9}
CH_3OH	$8_{1,8,2} - 7_{0,7,1}$	229758.756	89.10	5.918	20.9	7.23×10^{16}	1.17×10^{-7}
	$5_{1,5,0} - 4_{1,4,0}$	239746.219	49.06	7.765	12.6	3.18×10^{16}	5.89×10^{-8}
	$5_{0,5,1} - 4_{0,4,1}$	241700.159	47.93	7.506	8.49	2.01×10^{16}	3.75×10^{-8}
	$5_{1,5,2} - 4_{1,4,2}$	241767.234	40.39	10.66	16.1	3.60×10^{16}	6.87×10^{-8}
	$5_{0,5,0} - 4_{0,4,0}$	241791.352	34.82	10.60	17.6	3.50×10^{16}	6.81×10^{-8}
	$5_{3,3,0} - 4_{3,2,0}$	241832.718	84.62	7.752	5.26	3.18×10^{16}	5.24×10^{-8}
	$5_{2,4,0} - 4_{2,3,0}$	241842.284	72.53	9.367	6.59	2.56×10^{16}	4.39×10^{-8}
	$5_{3,3,1} - 4_{3,2,1}$	241843.604	82.53	9.367	6.59	3.86×10^{16}	6.40×10^{-8}
	$5_{1,4,1} - 4_{1,3,1}$	241879.025	55.87	6.665	4.46	1.19×10^{16}	2.16×10^{-8}
	$5_{2,3,2} - 4_{2,2,2}$	241904.147	60.72	7.222	8.19	2.73×10^{16}	4.88×10^{-8}
			Core 6				
$C^{34}S$	$5_0 - 4_0$	241016.089	34.70	7.813	5.63	2.11×10^{14}	3.64×10^{-10}
HC_3N	$26 - 25$	236512.788	153.2	5.812	3.94	2.04×10^{14}	2.38×10^{-10}
CH_3OH	$8_{1,8,2} - 7_{0,7,1}$	229758.756	89.10	8.298	19.7	6.81×10^{16}	1.20×10^{-7}
	$5_{1,5,0} - 4_{1,4,0}$	239746.219	49.06	7.556	14.9	3.76×10^{16}	7.60×10^{-8}
	$5_{0,5,1} - 4_{0,4,1}$	241700.159	47.93	9.664	11.2	2.65×10^{16}	5.37×10^{-8}
	$5_{3,3,0} - 4_{3,2,0}$	241832.718	84.62	8.114	5.87	3.54×10^{16}	6.37×10^{-8}
	$5_{2,4,0} - 4_{2,3,0}$	241842.284	72.53	8.571	7.48	2.91×10^{16}	5.44×10^{-8}
	$5_{3,3,1} - 4_{3,2,1}$	241843.604	82.53	8.571	7.48	4.38×10^{16}	7.92×10^{-8}
	$5_{1,4,1} - 4_{1,3,1}$	241879.025	55.87	6.236	5.95	1.59×10^{16}	3.14×10^{-8}
	$5_{2,3,2} - 4_{2,2,2}$	241904.147	60.72	7.230	10.7	3.57×10^{16}	6.95×10^{-8}
			Core 7				
$C^{34}S$	$5_0 - 4_0$	241016.089	34.70	6.611	7.86	2.94×10^{14}	5.51×10^{-10}
H_2CS	$7_{1,7} - 6_{1,6}$	236727.020	58.62	6.027	7.81	1.61×10^{15}	3.20×10^{-9}
CH_3OH	$8_{1,8,2} - 7_{0,7,1}$	229758.756	89.10	6.570	16.9	5.83×10^{16}	1.12×10^{-7}
	$5_{1,5,0} - 4_{1,4,0}$	239746.219	49.06	7.174	19.7	4.97×10^{16}	1.09×10^{-7}
	$5_{0,5,1} - 4_{0,4,1}$	241700.159	47.93	7.122	21.3	5.04×10^{16}	1.11×10^{-7}
	$5_{1,5,2} - 4_{1,4,2}$	241767.234	40.39	7.772	26.9	5.99×10^{16}	1.35×10^{-7}
	$5_{0,5,0} - 4_{0,4,0}$	241791.352	34.82	9.100	31.0	6.15×10^{16}	1.41×10^{-7}
	$5_{2,4,0} - 4_{2,3,0}$	241842.284	72.53	7.425	9.61	3.73×10^{16}	7.57×10^{-8}
	$5_{1,4,1} - 4_{1,3,1}$	241879.025	55.87	6.780	7.64	2.04×10^{16}	4.37×10^{-8}
	$5_{2,3,2} - 4_{2,2,2}$	241904.147	60.72	6.880	12.0	4.01×10^{16}	8.46×10^{-8}
	$5_{1,4,0} - 4_{1,3,0}$	243915.788	49.66	8.135	11.9	2.96×10^{16}	6.48×10^{-8}

4.4 Abundances

A summary of the abundances relative to Hydrogen that are calculated based on the results of the Gaussian fitting can be seen in Figure 9. The central points are the values calculated at dust and excitation temperatures of 100 K, while the upper and lower limit presents an estimated range by calculating the abundances at 37.5 and 150 K. Further, based on previous studies of $C^{17}O$ and its low upper state energy $E_{\rm up} = 16.18$ K, the species is expected to trace different, cooler gas components as compared to the other species presented in the figure. Thus, the abundances shown are calculated at 20 K, with an estimated range from calculations at 10 and 30 K. Then, upper limits of $3\,\sigma_{\rm rms}$ are also included in cases of non-detections for $C^{17}O$. In Figure 9, there is a trend throughout all cores

in which $C^{34}S$ is the least abundant, followed by H_2CS, and then CH_3OH and $C^{17}O$ at similar levels. Despite being the most complex detected species, CH_3OH appears to also be the most abundant in all cores except Core 2.

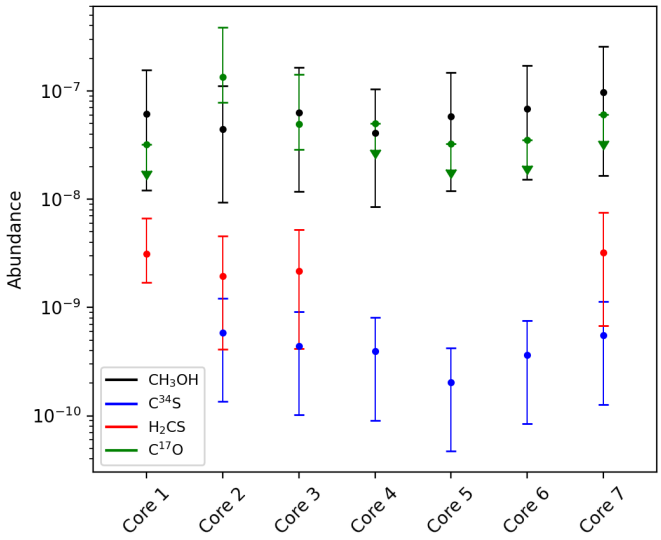


Figure 9: A diagram showing the abundances relative to Hydrogen nuclei of automatically detected species, for all cores. The values are presented as a range, based on calculations at assumed dust and excitation temperatures of 37.5 and 150 K.

Further investigating the CH_3OH abundance in the region, the values are compared to those from previous hot cores discussed by Gorai et al. (2024), seen in Figure 10. G28.2-HMC3 is located further from the central continuum emission of G28.20-0.05, estimated to have the lowest temperature of the three hot cores in the region, and shows a CH_3OH abundance in the lower range of the G28.37 intervals. Meanwhile, the remaining hot cores display increased abundances by approximately an order of magnitude, while also showing emission from several other complex organic molecules. This comparison suggests that G28.37 could contain earlier stage massive protostars, with lower abundances of complex molecules yet released into gas phase by the heating of the central protostars.

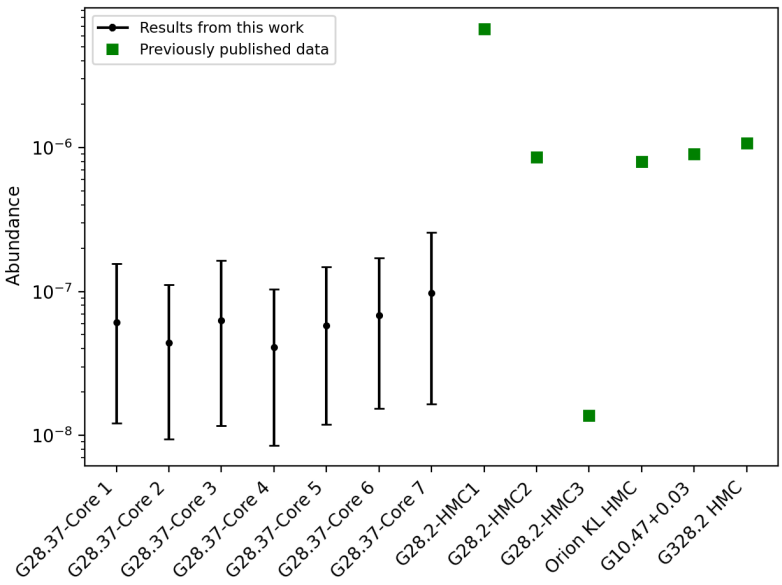


Figure 10: Abundances of CH_3OH relative to Hydrogen, for the seven cores observed in G28.37 and also compared to values presented by Gorai et al. (2024).

As a final comparison to the chemical properties of hot molecular cores presented by Gorai et al. (2024), Figure 11 is adapted to include G28.37 in the correlation between column densities of CH_3OH

and H_2CO. The species are astrochemically linked, as H_2CO is believed to be a stage in the chemical pathway in forming CH_3OH through cool dust-grain surface chemistry. Now, CH_3OH is widespread in G28.37, while H_2CO emission is only detected toward two cores, and again only one core whose Gaussian parameters could be extracted. Thus, Core 4 is the only data point included in the figure. Comparing G28.37 to the previous trend, Core 4 is found to contain a larger relative amount of CH_3OH to H_2CO. Overall, G28.37 shows abundant and widespread CH_3OH, with low or even no emission from H_2CO. This can be used to constrain the astrochemical evolution of these massive protostars, and could be related to the mechanisms through which molecules leave dust grains and enter the gas phase, or the chemical pathways for molecule formation.

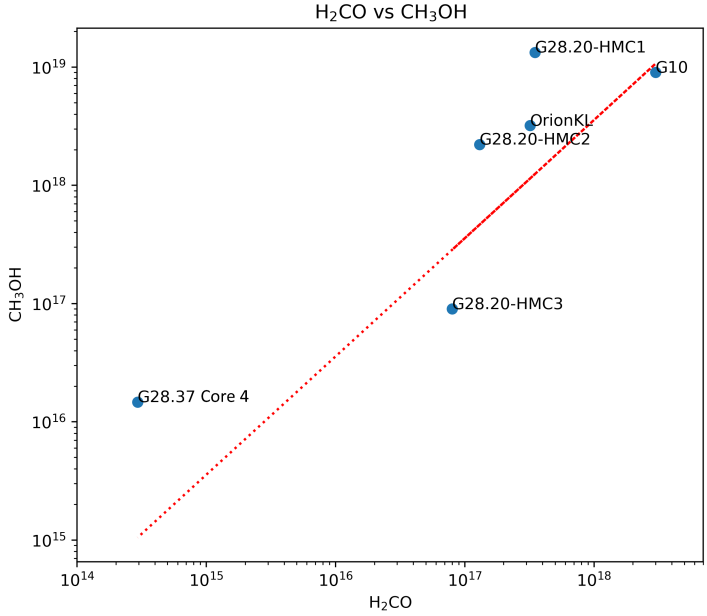


Figure 11: Correlation between column densities of CH_3OH and H_2CO, adapted from Gorai et al. (2024). Core 4 is the only G28.37 core with a column density estimate for H_2CO, and is found to show a high relative amount of CH_3OH, while still similar to some of the previously studied hot cores.

4.5 Kinematics

The kinematic properties extracted from the emission lines of species in all cores can be seen in figures 12 and 13, giving a summary of the previously shown line profiles from figures 5 through 8. From the velocity shifts, some general trends can be seen. In all cores where the simpler species $C^{34}S$ and $C^{17}O$ are detected, except Core 7, they are distinctly shifted from the CH_3OH emission. H_2CS is generally less shifted, and closer to the v_{LSR} of CH_3OH. These trends give indications that the species may be tracing different gas components, with slightly different relative motions, which could be caused by outflows or different spatial distances from the central protostar.

Furthermore, the large uncertainties in the CH_3OH emission could be a result of a mixture of velocity components for the higher and lower E_{up} transitions. Some of these can be observed in the overview of the line profiles, such as the $E_{up} = 34\,K$ line in Core 3 from Figure 6 and Core 5 from Figure 7, showing wide profiles and uncertain peak positions. These could perhaps indicate a separate CH_3OH population present in outflowing gas. Such possibilities could be the focus of further studies of the region, focusing on the spatial distribution and velocity gradients of the detected species from this work.

The second kinematic property extracted from the lines is their FWHM, which for a Gaussian profile is directly related to the velocity dispersion of the species. Figure 13 shows an overview of the FWHM as a function of E_{up} of each transition, for all cores. Generally, the linewidths appear to show a flat scaling with E_{up}. The general population can be seen to have widths scattered around a

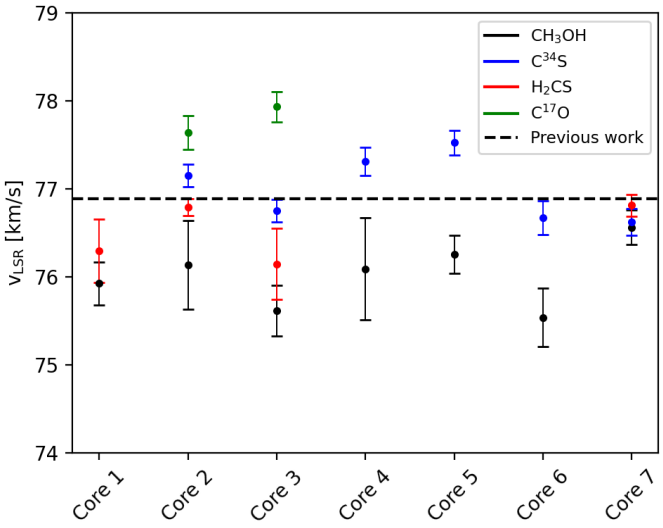


Figure 12: A diagram showing the v_{LSR} and one standard deviation uncertainty of automatically detected species, for all cores. In the case of species with several detected lines, the values are the inverse variance weighted averages with respective standard uncertainties. The v_{LSR} estimated by Light et al. (In prep.) is also included as the dashed black line.

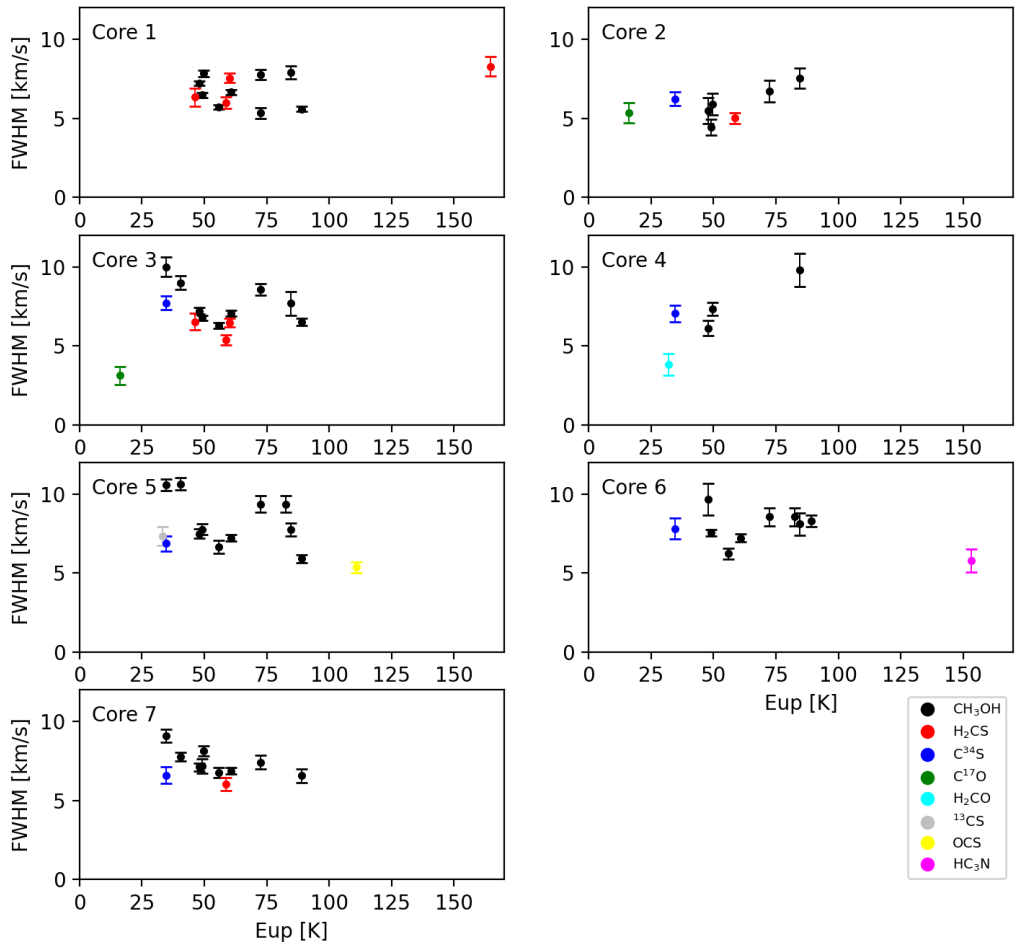


Figure 13: Overview of FWHM and one standard deviation uncertainties plotted as a function of E_{up} for all lines from the Gaussian fitting, for all seven cores.

typical value of 7 km s^{-1}, ranging from 5 to 10 km s^{-1}. Core 2 is a slight exception, with most lines at the lower end of the range at 5 km s^{-1}, which could indicate a lower temperature core.

At the higher end of the range, in cores 3, 5, 6, and 7, the lowest $E_{\rm up}$ CH_3OH transitions show widths of $\gtrsim 10$ km s^{-1}. Lower $E_{\rm up}$ transitions can be excited at lower excitation temperatures, which would then suggest a generally smaller velocity dispersion. Instead, these amplified linewidths could be the result of outflows, leading to separate velocity components blending and broadening the lines. On a similar note, $C^{17}O$ in the two detections shows a FWHM of $3-5$ km s^{-1}, which especially in Core 3 is seen to deviate from the remaining population. This further suggests that $C^{17}O$ is tracing another gas component, likely with cooler gas indicated by the lower velocity dispersion.

4.6 Limitations of the method

The line analysis algorithm proved to be a useful tool in quickly extracting the molecular inventory, as well as abundance and kinematic results. The method could prove especially useful in more line-rich regions, which would otherwise require vast amounts of manual careful work. However, there are still a number of areas of improvement. To completely assess the chemical inventory of hot molecular cores, the script should be expanded to also search for and distinguish absorption lines. The H_2CO $E_{\rm up} = 32.06$ K transition was present as an absorption line in all cores except the emission line in Core 4, as can be seen in Table 5. Such absorption lines could be included in the current algorithm without much alteration by also searching for peaks in the inverted spectrum. Further, self-absorption features such as the ones observed for CS and H_2CO could be sought after by investigating the profile of the residual. Comparing the line profile with its Gaussian fit, self-absorption could be detected by identifying an absorption in an otherwise Gaussian profile. Lastly, blended lines could become a crucial part of future studies of more line-rich regions, and could be handled through fitting of multiple simultaneous Gaussian components. The unsuccessful Gaussian fit of HNCO serves as a prime example of this limitation.

5 Conclusions

Observations carried out with the Submillimeter Array in a continuous 44 GHz spectral range toward the high-mass star forming region G28.37 have been analyzed. From the millimeter dust continuum emission seven cores were extracted, confirming the clustered nature of the star forming region. The two strongest emission cores, Core 1 and Core 5, could be related to the regions of strongest emission as seen in 31 µm SOFIA observations, however, several new cores are also seen. For each, the surface mass and hydrogen column densities are estimated, yielding $N_{\rm H}$ values in the range $5.34 \times 10^{23} - 1.32 \times 10^{24}$ cm^{-2}.

From each core, the 44 GHz spectrum was extracted and studied using an automatic line analysis algorithm. The spectral data was systematically analyzed by detecting peaks and cross-matching to species previously detected toward hot molecular cores, using the CDMS database. Several molecules, including H_2CO, H_2CS, HNCO, and CH_3OH, were detected in assessing the chemical inventories of the cores. In general, the cores showed a lower level of chemical richness than typical hot molecular cores, indicating a less chemically evolved region of massive star formation. Then, Gaussian fitting was performed for each line while also automatically excluding poor fits, and parameters were extracted to study the abundances and kinematics. The automatic line analysis algorithm proved to be a useful tool, and could serve as a key step in handling the large amount of data expected in the continued SMA line survey. There are however still limitations to the method, which could be expanded upon in continued work.

The abundances showed a similar trend in all cores, where CH_3OH despite being the most complex detected molecule, also was the most abundant. C^{17} was observed in similar abundances, although only detected above the signal-to-noise threshold in two of the cores. A few species with strong signals were not included in the abundance calculations, such as CS in several cores and H_2CO in one core, because of self-absorption features in the line profile. CH_3OH, being the most widespread species, was compared to previous studies of hot molecular cores. The cores presented by Gorai et al. (2024) in general showed CH_3OH abundances higher by an order of magnitude. Finally, the abundant detections of CH_3OH in all cores without simultaneous emission of the chemically linked species H_2CO is also discussed.

From the kinematic properties, $v_{\rm LSR}$ and FWHM, CH_3OH was both suggested to trace gas closer to the hot cores together with the emission from H_2CS, but lower $E_{\rm up}$ transitions were found to show velocity shifts and suggestions of line broadening through multiple velocity components. As such, CH_3OH could also show signs of potential outflows in the region. Further, $C^{17}O$ deviated from the remaining population both in terms of a velocity shift, but also through a generally smaller linewidth. As such, the $C^{17}O$ emission is also indicating a separate, cooler gas component. Overall, the FWHM was scattered between $\sim$5 and 10 km s^{-1} for all cores, showing a flat scaling with $E_{\rm up}$.

References

Bonfand, M. 2019, *Complex organic chemistry in high-mass star forming regions.* Ph. D. Thesis. University of Bonn. Available at: https://hdl.handle.net/20.500.11811/8080

Butler, M. J. & Tan, J. C. 2009, The Astrophysical Journal, 696, 484, doi: 10.1088/0004-637X/696/1/484

Carey, S. J., Clark, F. O., Egan, M. P., et al. 1998, The Astrophysical Journal, 508, 721, doi: 10.1086/306438

De Buizer, J. M., Liu, M., Tan, J. C., et al. 2017, The Astrophysical Journal, 843, 33, doi: 10.3847/1538-4357/aa74c8

Draine, B. 2011, *Physics of the Interstellar and Intergalactic Medium.* Princeton: Princeton University Press. doi: 10.1515/9781400839087

Ginsburg, A., Sipőcs, B. M., Brasseur, C. E., et al. 2019, The Astronomical Journal, 157, 98, doi: 10.3847/1538-3881/aafc33

Gorai, P., Law, C-Y., Tan, J. C., et al. 2024, The Astrophysical Journal, 960, 126, doi: 10.3847/1538-4357/ad09bb

Gudimenko, E., Milosavljevic, V., & Daniels, S. 2012, Optics Express, 20, 12699, doi: 10.1364/OE.20.012699

Ho, P. T. P., Moran, J. M., & Lo, K. Y. 2004, The Astrophysical Journal, 616, L1, doi: 10.1086/423245

Jørgensen, J. K., Belloche, A., Garrod, R. T. 2020, Annual Review of Astronomy and Astrophysics, 58, 727, doi: 10.1146/annurev-astro-032620-021927

Liu, J., Zhang, Q., Qiu, K., et al. 2020, The Astrophysical Journal, 895, 142, doi: 10.3847/1538-4357/ab9087

Law, C.-Y., Tan., J. C., Gorai, P., et al. 2022, The Astrophysical Journal, 939, 120, doi: 10.3847/1538-4357/ac90c7

Liu, M., Qin, S.-L., Liu, T., et al. 2023, The Astrophysical Journal, 958, 174, doi: 10.3847/1538-4357/ad00aa

Motte, F., Bontemps, S., Louvet, F. 2018, Annual Review of Astronomy and Astrophysics, 56, 41, doi: 10.1146/annurev-astro-091916-055235

Müller, H. S. P., Schlöder, F., Stutzki, J., & Winnewisser, G. 2005, Journal of Molecular Structure, 742, 215, doi: 10.1016/j.molstruc.2005.01.027

Müller, H. S. P., Thorwirth, S., Roth, D. A., & Winnewisser, G. 2001, Astronomy & Astrophysics, 370, L49, doi: 10.1051/0004-6361:20010367

Pickett, H. M., Poynter, R. L., Cohen, E. A., et al. 1998, Journal of Quantitative Spectroscopy and Radiative Transfer, 60, 883, doi: 10.1016/S0022-4073(98)00091-0

Rathborne, J. M., Jackson, J. M. & Simon, R. 2006, The Astrophysical Journal, 641, 389, doi: 10.1086/500423

Riener, M., Kainulainen, J., Henshaw, J. D., et al. 2019, Astronomy & Astrophysics, 628, A78, doi: 10.1051/0004-6361/201935519

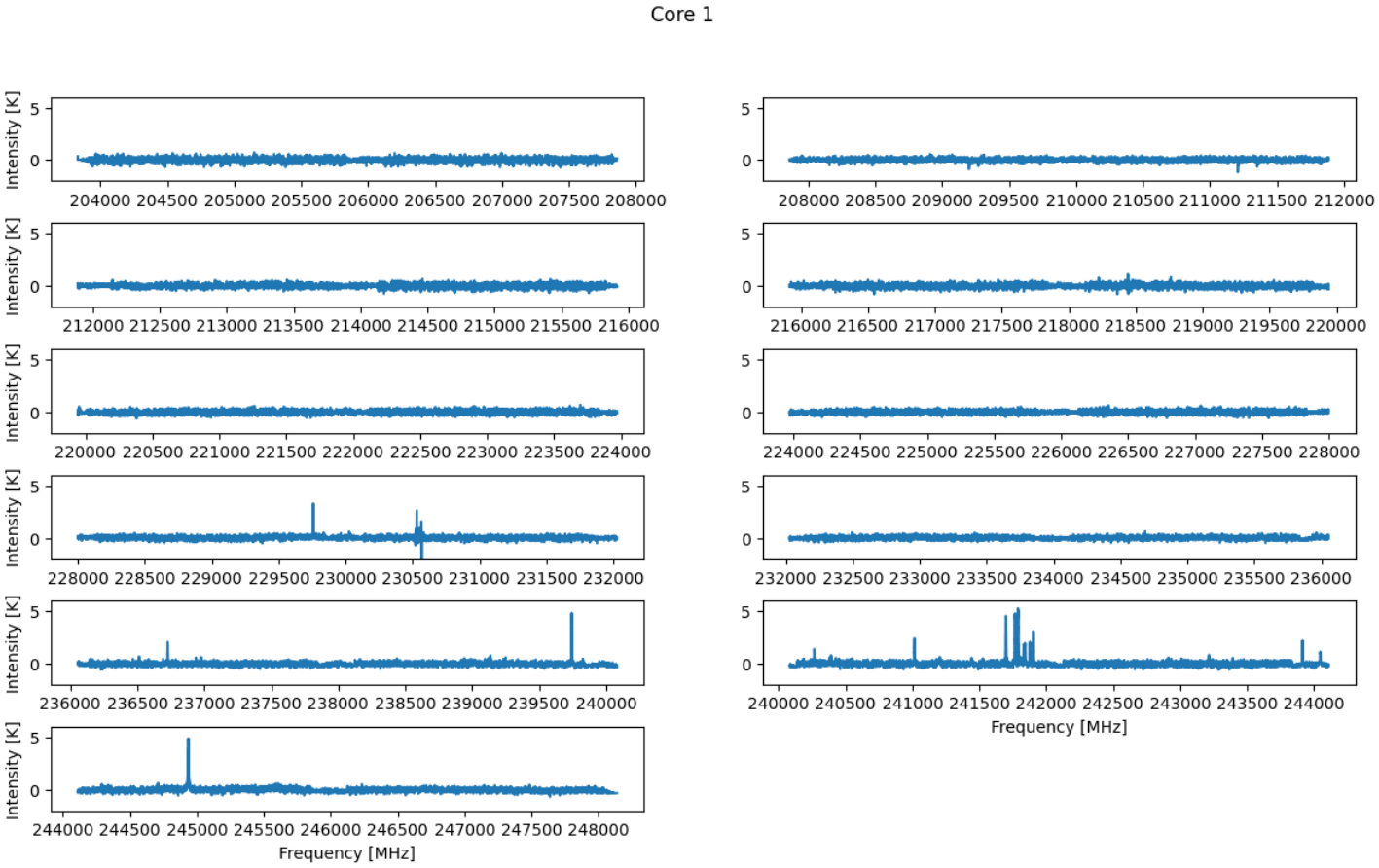
Simon, R., Jackson, J. M., Rathborne, J. M. & Chambers, E. T. 2006a, The Astrophysical Journal, 639, 227, doi: 10.1086/499342

Simon, R., Rathborne, J. M., Shah, R. Y., et al. 2006b, The Astrophysical Journal, 653, 1325, doi: 10.1086/508915

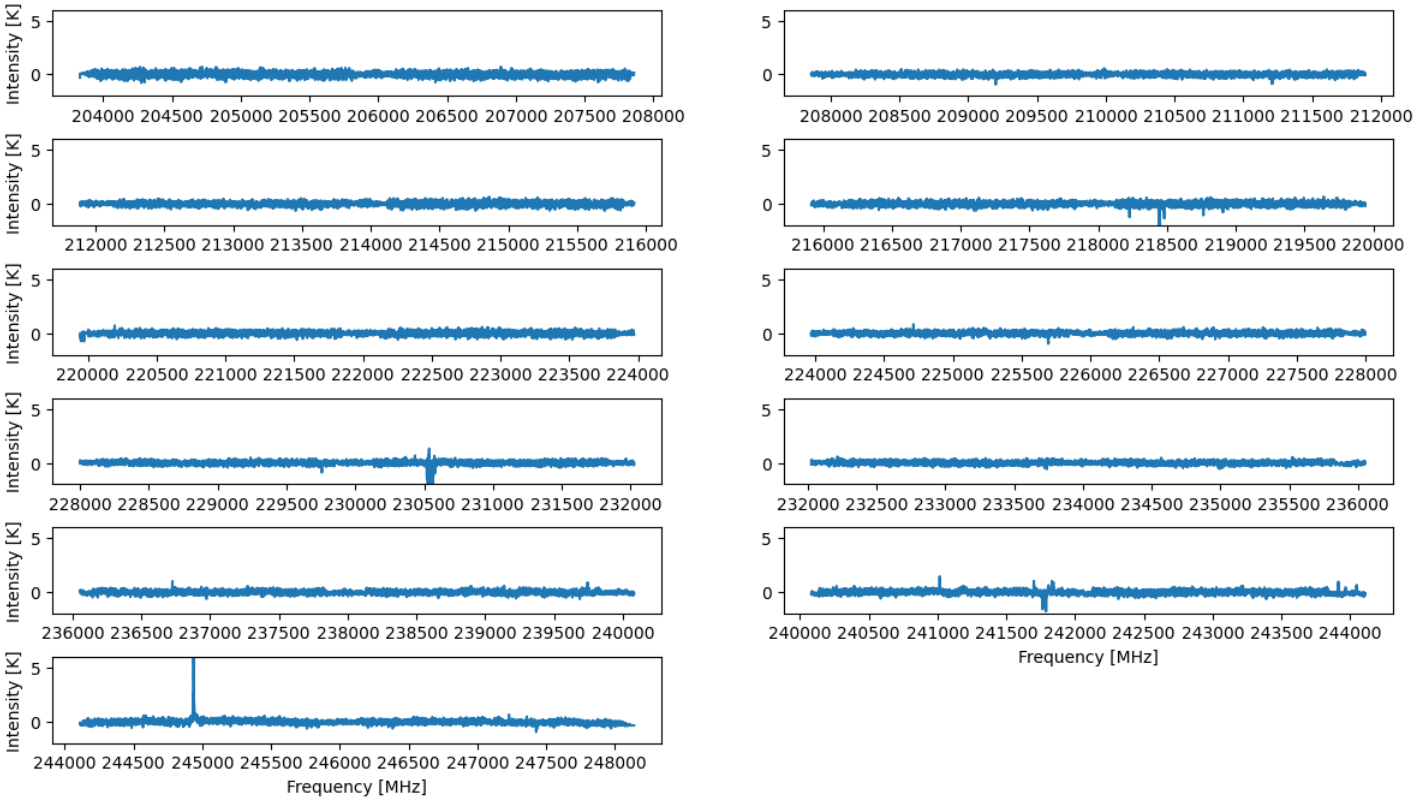
The CASA Team et al. 2022, Publications of the Astronomical Society of the Pacific, 134, 114501, doi: 10.1088/1538-3873/ac9642

A. Appendix

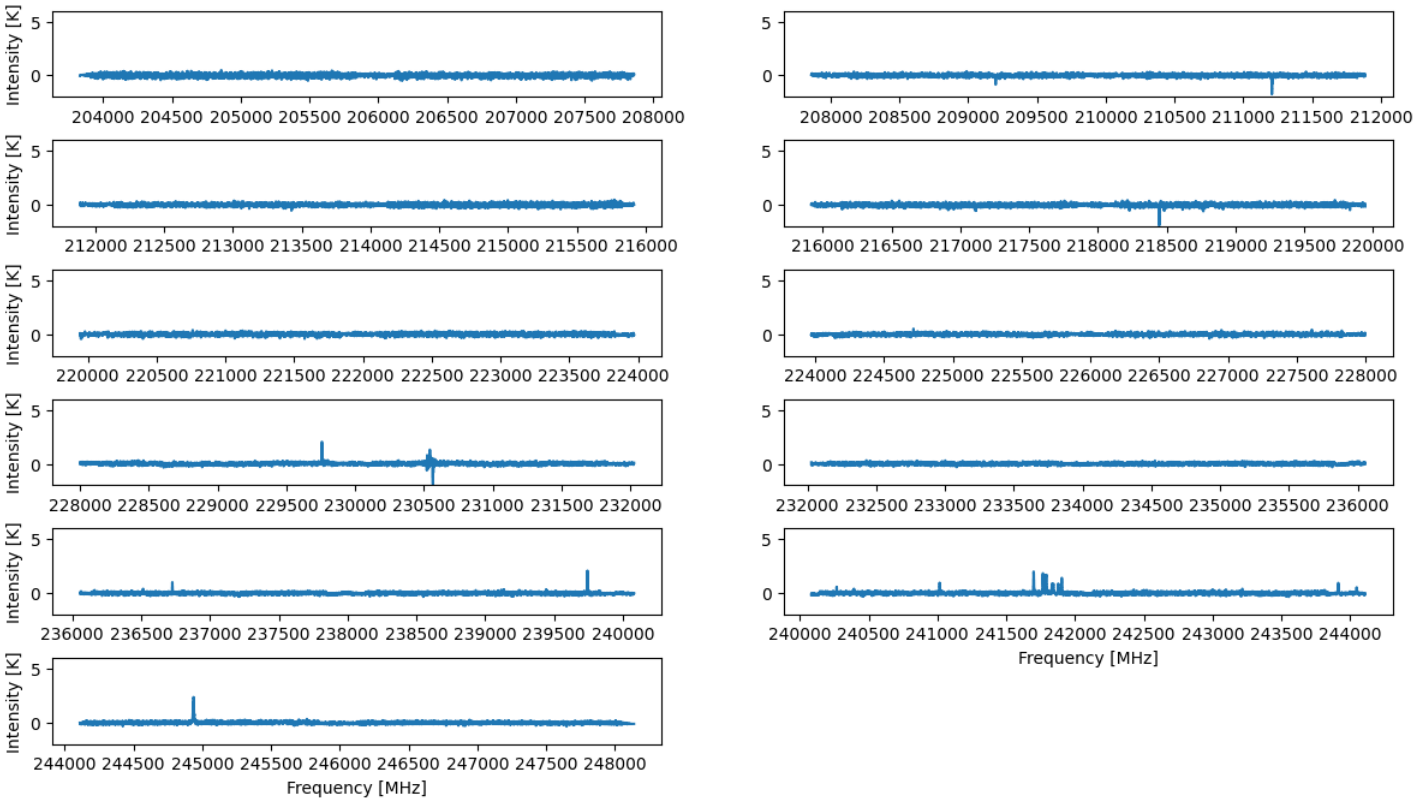
On the following pages, an overview of the entire spectral dataset is provided.



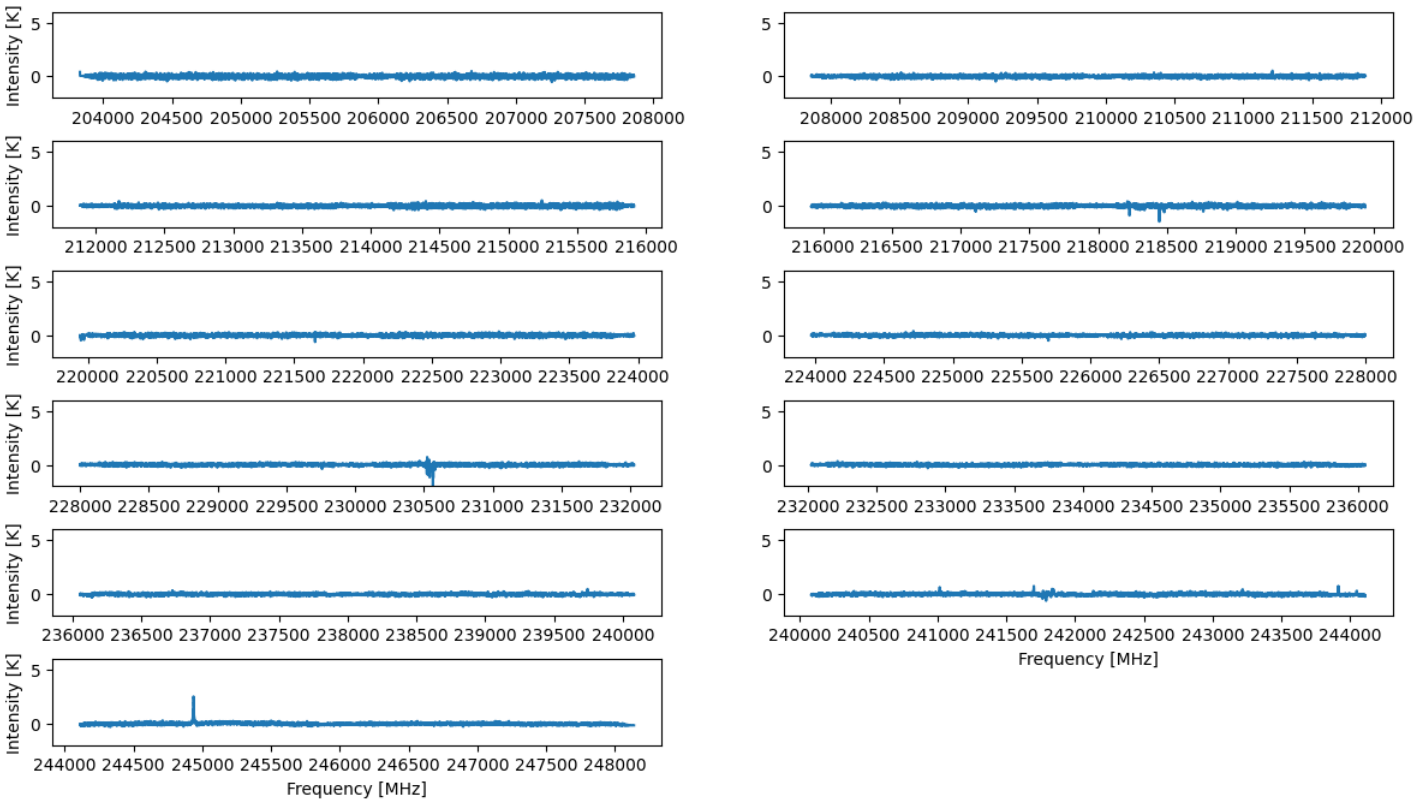
Core 2


Intensity [K]
5
0
204000 204500 205000 205500 206000 206500 207000 207500 208000
208000 208500 209000 209500 210000 210500 211000 211500 212000
212000 212500 213000 213500 214000 214500 215000 215500 216000
216000 216500 217000 217500 218000 218500 219000 219500 220000
220000 220500 221000 221500 222000 222500 223000 223500 224000
224000 224500 225000 225500 226000 226500 227000 227500 228000
228000 228500 229000 229500 230000 230500 231000 231500 232000
232000 232500 233000 233500 234000 234500 235000 235500 236000
236000 236500 237000 237500 238000 238500 239000 239500 240000
240000 240500 241000 241500 242000 242500 243000 243500 244000
244000 244500 245000 245500 246000 246500 247000 247500 248000
Frequency [MHz]

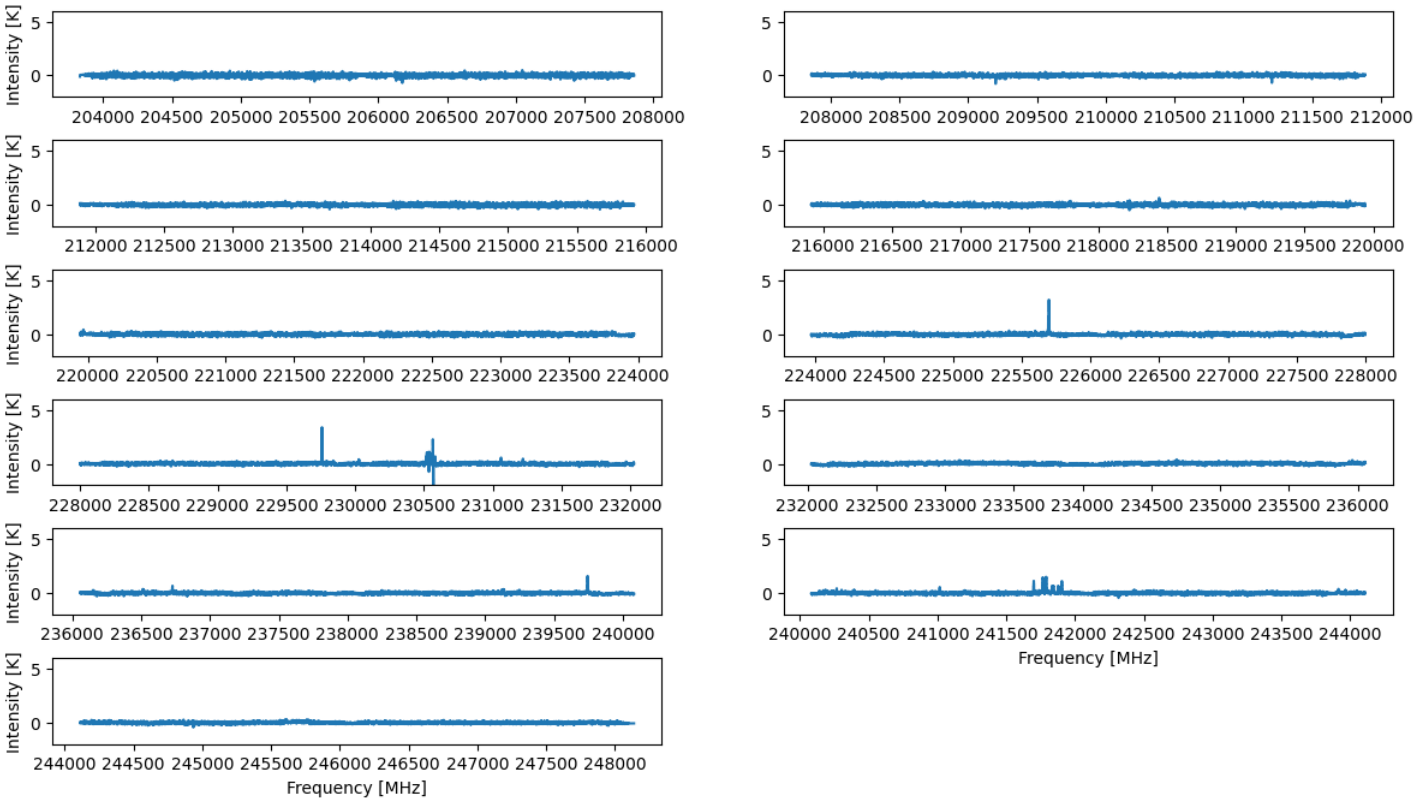
Core 3


Intensity [K]
5
0
204000 204500 205000 205500 206000 206500 207000 207500 208000
208000 208500 209000 209500 210000 210500 211000 211500 212000
212000 212500 213000 213500 214000 214500 215000 215500 216000
216000 216500 217000 217500 218000 218500 219000 219500 220000
220000 220500 221000 221500 222000 222500 223000 223500 224000
224000 224500 225000 225500 226000 226500 227000 227500 228000
228000 228500 229000 229500 230000 230500 231000 231500 232000
232000 232500 233000 233500 234000 234500 235000 235500 236000
236000 236500 237000 237500 238000 238500 239000 239500 240000
240000 240500 241000 241500 242000 242500 243000 243500 244000
244000 244500 245000 245500 246000 246500 247000 247500 248000
Frequency [MHz]

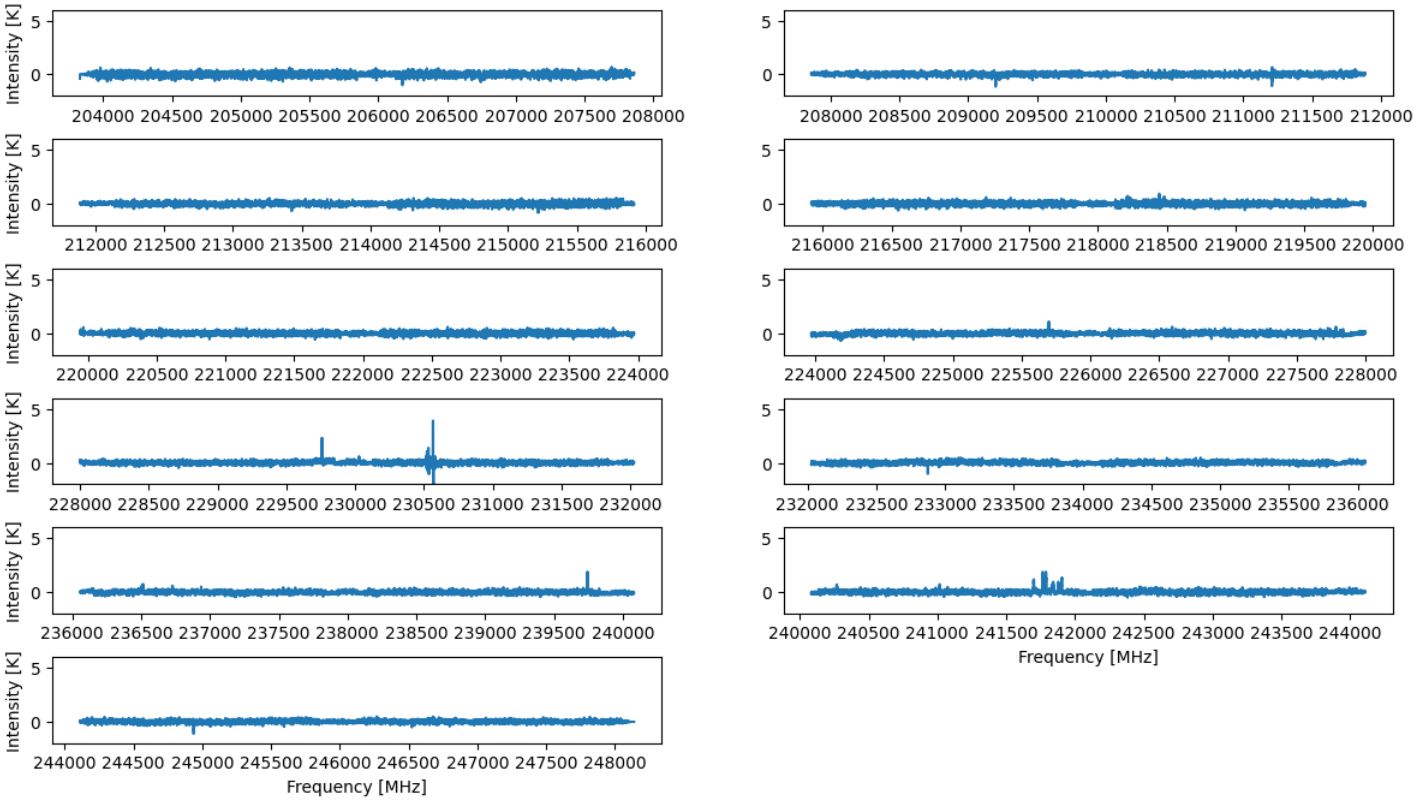
Core 4


Intensity [K]
5
0
204000 204500 205000 205500 206000 206500 207000 207500 208000
208000 208500 209000 209500 210000 210500 211000 211500 212000
212000 212500 213000 213500 214000 214500 215000 215500 216000
216000 216500 217000 217500 218000 218500 219000 219500 220000
220000 220500 221000 221500 222000 222500 223000 223500 224000
224000 224500 225000 225500 226000 226500 227000 227500 228000
228000 228500 229000 229500 230000 230500 231000 231500 232000
232000 232500 233000 233500 234000 234500 235000 235500 236000
236000 236500 237000 237500 238000 238500 239000 239500 240000
240000 240500 241000 241500 242000 242500 243000 243500 244000
244000 244500 245000 245500 246000 246500 247000 247500 248000
Frequency [MHz]

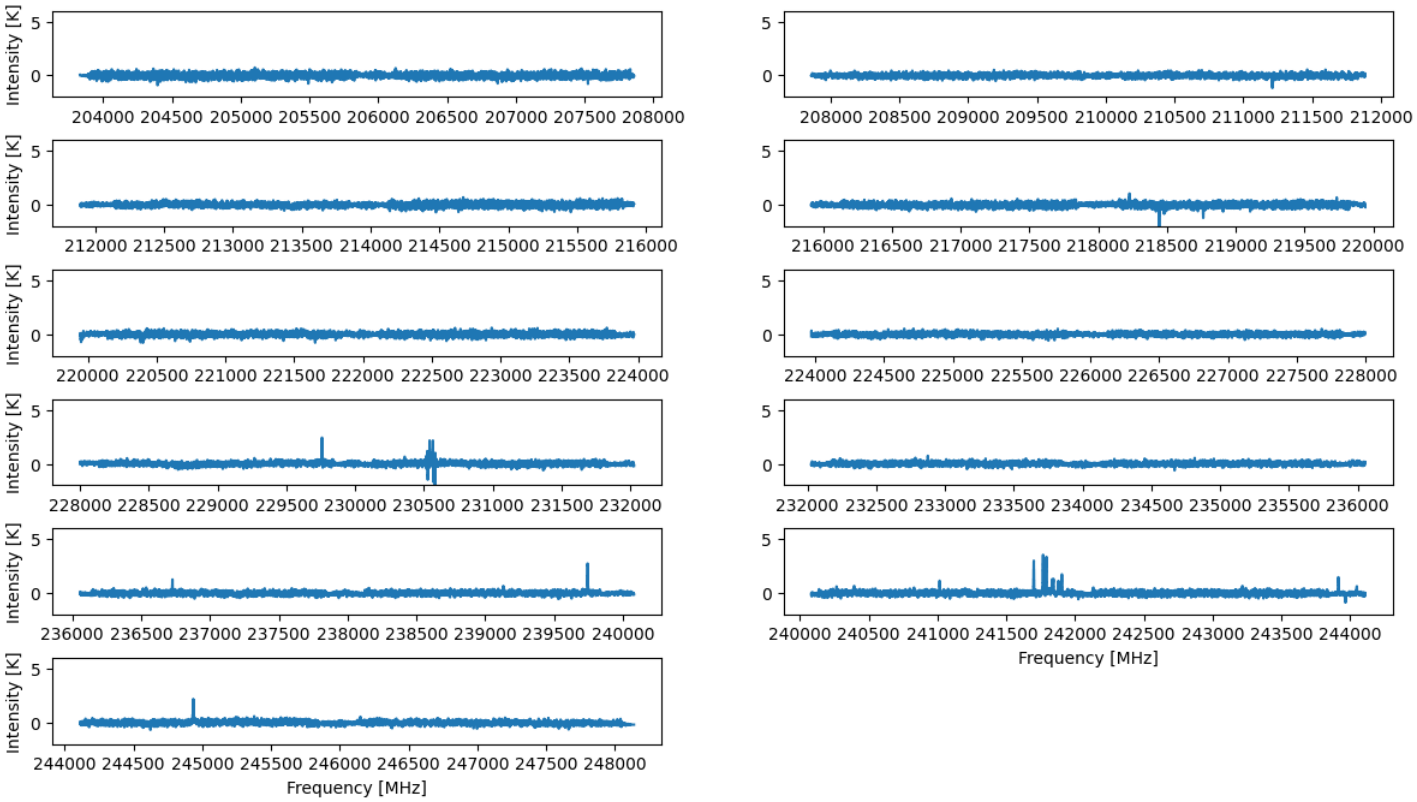
Core 5


Intensity [K]
5
0
204000 204500 205000 205500 206000 206500 207000 207500 208000
208000 208500 209000 209500 210000 210500 211000 211500 212000
212000 212500 213000 213500 214000 214500 215000 215500 216000
216000 216500 217000 217500 218000 218500 219000 219500 220000
220000 220500 221000 221500 222000 222500 223000 223500 224000
224000 224500 225000 225500 226000 226500 227000 227500 228000
228000 228500 229000 229500 230000 230500 231000 231500 232000
232000 232500 233000 233500 234000 234500 235000 235500 236000
236000 236500 237000 237500 238000 238500 239000 239500 240000
240000 240500 241000 241500 242000 242500 243000 243500 244000
244000 244500 245000 245500 246000 246500 247000 247500 248000
Frequency [MHz]

Core 6


Intensity [K]
5
0
204000 204500 205000 205500 206000 206500 207000 207500 208000
208000 208500 209000 209500 210000 210500 211000 211500 212000
212000 212500 213000 213500 214000 214500 215000 215500 216000
216000 216500 217000 217500 218000 218500 219000 219500 220000
220000 220500 221000 221500 222000 222500 223000 223500 224000
224000 224500 225000 225500 226000 226500 227000 227500 228000
228000 228500 229000 229500 230000 230500 231000 231500 232000
232000 232500 233000 233500 234000 234500 235000 235500 236000
236000 236500 237000 237500 238000 238500 239000 239500 240000
240000 240500 241000 241500 242000 242500 243000 243500 244000
244000 244500 245000 245500 246000 246500 247000 247500 248000
Frequency [MHz]

Core 7


Intensity [K]
5
0
204000 204500 205000 205500 206000 206500 207000 207500 208000
208000 208500 209000 209500 210000 210500 211000 211500 212000
212000 212500 213000 213500 214000 214500 215000 215500 216000
216000 216500 217000 217500 218000 218500 219000 219500 220000
220000 220500 221000 221500 222000 222500 223000 223500 224000
224000 224500 225000 225500 226000 226500 227000 227500 228000
228000 228500 229000 229500 230000 230500 231000 231500 232000
232000 232500 233000 233500 234000 234500 235000 235500 236000
236000 236500 237000 237500 238000 238500 239000 239500 240000
240000 240500 241000 241500 242000 242500 243000 243500 244000
244000 244500 245000 245500 246000 246500 247000 247500 248000
Frequency [MHz]