
Chemical Segregation in the Early Stages of Star Formation

Katharina Giers

München 2025

Chemical Segregation in the Early Stages of Star Formation

Katharina Giers

Dissertation
an der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Katharina Giers
aus Esslingen am Neckar, Deutschland

München, den 13. August 2025

Erstgutachter: Prof. Dr. Paola Caselli

Zweitgutachter: Prof. Dr. Barbara Ercolano

Tag der mündlichen Prüfung: 1. Oktober 2025

Contents

Zusammenfassung	xv
Summary	xvii
1 Introduction	1
1.1 The interstellar medium and low-mass star formation	1
1.2 Astrochemistry	7
1.2.1 Fractionation processes	11
1.3 Observational studies	15
1.3.1 Radiative transfer	15
1.3.2 Spectral line emission	17
1.3.3 Radioastronomy	21
1.4 Machine learning in astrophysics	24
1.5 This thesis	26
2 Similar levels of deuteration in the prestellar core L1544 and the proto-stellar core HH211	29
2.1 Introduction	30
2.2 Observations	32
2.3 Results	32
2.3.1 L1544	35
2.3.2 HH211	38
2.4 Analysis	39
2.4.1 Excitation temperature and column density	39
2.4.2 Deuterium fraction	42
2.5 Radiative transfer modelling	44
2.5.1 Physical structure of L1544	44
2.5.2 Physical structure of HH211	46
2.5.3 Molecular abundances	46
2.5.4 Results	47
2.6 Discussion	57
2.6.1 Dilution of ^{13}C in CCH	57
2.6.2 Deuterium fractionation	57

2.6.3	Radiative transfer modelling	59
2.7	Conclusions	61
3	Chemical segregation analysed with unsupervised clustering	63
3.1	Introduction	64
3.2	Observations and data reduction	66
3.2.1	Data	66
3.2.2	Sources	66
3.3	Methodology	67
3.3.1	Preprocessing	67
3.3.2	Clustering	69
3.3.3	Case studies	71
3.4	Results	72
3.4.1	H ₂ column density gradient	72
3.4.2	Comparison of dataset features	73
3.4.3	Clustering	75
3.4.4	CH ₃ CCH abundances	85
3.5	Discussion	87
3.5.1	Density-based clustering	87
3.5.2	Evolution traced by CH ₃ CCH	89
3.6	Conclusions	91
4	Deuteration of HC₃N and CH₃CCH in the prestellar core L1544	93
4.1	Introduction	94
4.2	Data	96
4.3	Analysis	97
4.3.1	Spectral line emission	97
4.3.2	Non-LTE modelling at the dust peak	100
4.3.3	Column density and deuteration maps	105
4.4	Discussion	109
4.4.1	HC ₃ N	109
4.4.2	CH ₃ CCH	110
4.5	Conclusions	112
5	Conclusions	113
5.1	Summary of this thesis	113
5.2	Outlook and future prospects	114
A	Supplementary material for Chapter 2	119
A.1	Observed lines	119
A.2	Parameters of the additional velocity component towards HH211	125
A.3	Collisional rate coefficients	126
A.4	Derivation of the physical structure of HH211	127

A.5	Molecular abundance profiles of the best-fit modelling results	131
B	Supplementary material for Chapter 3	135
B.1	Integrated intensity maps	135
B.2	H ₂ column density gradient maps	139
B.3	Detailed clustering results	140
B.4	Spectra at dust peak	153
C	Supplementary material for Chapter 4	155
C.1	Deuteration of CH ₃ CCH	155
C.2	Spectra at molecular peaks	155
	Danksagung	171

List of Figures

1.1	The different stages of low-mass star formation.	6
1.2	Spectral energy distributions of protostar classes.	7
1.3	Scheme of the three basic steps of deuterium fractionation.	13
1.4	Sketch of an antenna beam pattern.	22
2.1	Selection of the spectra observed towards L1544 and HH211.	33
2.2	Central hyperfine of the 1-0 transitions of HCN and HC ₃ N, observed towards L1544.	36
2.3	Multi-component Gaussian fit of the four effective hyperfine components of HN ¹³ C observed towards L1544.	37
2.4	Profiles of the gas temperature, H ₂ number density, and infall velocity for the Keto-Caselli model, the HDCRT model, and the physical model of HH211.	45
2.5	Synthetic spectra of CCH obtained with LOC.	50
2.6	Synthetic spectra of HCN, H ¹³ CN, and DCN obtained with LOC.	52
2.6	continued.	53
2.7	Synthetic spectra of HNC and isotopologues obtained with LOC.	55
2.8	Synthetic spectra of HCO ⁺ and isotopologues obtained with LOC.	56
2.9	C ¹³ CH/ ¹³ CCH ratio of starless, pre-stellar, and protostellar cores.	58
3.1	Comparison of different features for B68, L1521E, and L1544.	74
3.2	Clustering results for the starless core B68 for the dataset of Case 1, for feature combinations 2, 3, 4, 9, and 10.	77
3.2	Corresponding spatial distribution of each cluster across the core.	78
3.3	Clustering results for the starless core B68 for the dataset of Case 2, for feature combinations 2, 3, 4, 9, and 10.	80
3.3	Corresponding spatial distribution of each cluster across the core.	81
3.4	Comparison of the abundances of CH ₃ CCH and c-C ₃ H ₂ at the dust peak of different starless and prestellar cores.	86
3.5	Centroid velocities and linewidths of c-C ₃ H ₂ , CH ₃ OH, and CH ₃ CCH towards the prestellar core L1544.	90
3.6	continued.	91
4.1	Integrated intensity maps of the observed transitions.	98

4.1	continued.	99
4.2	Observed spectra of the CH_3CCH (5–4) and the CH_3CCD (6–5) $K=0$ and $K=1$ transitions.	100
4.3	Profiles of the gas temperature, H_2 number density, and infall velocity for the Keto-Caselli model of L1544.	101
4.4	Fractional abundance profiles of HC_3N of the best-fit results produced with LOC.	103
4.5	Comparison of the observed spectra of HC_3N , HCC^{13}CN , and CH_3CCH to the synthetic spectra produced with LOC.	104
4.6	Column density maps of HC_3N and CH_3CCH , and their deuterated isotopologues.	106
4.7	Deuteration maps of HC_3N and CH_3CCH	108
A.1	Ground-state rotational lines observed towards the prestellar core L1544.	120
A.1	continued.	122
A.2	Ground-state rotational lines observed towards the protostellar core HH211.	123
A.2	continued.	124
A.3	Temperature dependence of hyperfine resolved rate coefficients of HCN , DCN , HNC , DNC , and HN^{13}C	127
A.4	H_2 column density and dust temperature maps of HH211 derived from <i>Herschel</i> SPIRE data.	130
A.5	Radial H_2 column density and temperature profiles of HH211 derived from <i>Herschel</i> SPIRE data.	130
A.6	Fractional molecular abundance profiles of the best-fit results produced with LOC for the spectra observed towards L1544.	131
A.6	continued.	132
A.7	Fractional molecular abundance profiles of the best-fit results produced with LOC for the spectra observed towards HH211.	133
A.7	continued.	134
B.1	Integrated intensity maps of $\text{c-C}_3\text{H}_2$, CH_3OH , and CH_3CCH observed towards B68, L1521E, and L1544.	135
B.1	continued.	136
B.1	continued.	137
B.1	continued.	138
B.2	H_2 column density maps and the corresponding H_2 column density gradient maps for B68, L1521E, and L1544.	139
B.3	Clustering results for the starless core L1521E for the dataset of Case 1, for feature combinations 2, 3, 4, 9, and 10.	145
B.3	Corresponding spatial distribution of each cluster across the core.	146
B.4	Clustering results for the starless core L1521E for the dataset of Case 2, for feature combinations 2, 3, 4, 9, and 10.	147
B.4	Corresponding spatial distribution of each cluster across the core.	148

B.5	Clustering results for the prestellar core L1544 for the dataset of Case 1, for feature combinations 2, 3, 4, 9, and 10.	149
B.5	Corresponding spatial distribution of each cluster across the core.	150
B.6	Clustering results for the prestellar core L1544 for the dataset of Case 2, for feature combinations 2, 3, 4, 9, and 10.	151
B.6	Corresponding spatial distribution of each cluster across the core.	152
B.7	Spectra of CH ₃ CCH and c-C ₃ H ₂ at the dust peak of a sample of starless and prestellar cores.	153
B.7	continued.	154
C.1	Column density map and corresponding deuteration maps of CH ₃ CCH derived from the (5 ₁ – 4 ₁) transition.	156
C.2	Spectra of the observed transitions extracted at the dust peak, the CH ₃ OH peak, the CH ₃ CCH peak, and the c-C ₃ H ₂ peak of L1544.	158

List of Tables

1.1	Astrochemical gas-phase reactions.	9
2.1	Properties of the selected observed lines.	34
2.2	Parameters used in the derivation of the column densities.	41
2.3	Derived column densities of L1544 and HH211.	42
2.4	Column density ratios in both cores.	43
2.5	Summary of the simulation setups discussed in this work.	48
3.1	Source sample.	66
3.2	Spectroscopic parameters of the observed lines.	67
3.3	Feature combinations used in the clustering.	70
3.4	Molecular transitions used in the case studies.	72
4.1	Spectroscopic parameters of the observed lines.	97
4.2	Column densities of the best-fit models obtained with LOC.	102
A.1	Properties of the observed lines.	121
A.2	Best-fit parameters and column densities of the additional velocity component observed towards HH211.	125
A.3	Comparison of column density ratios in the main component and the additional velocity component observed towards HH211.	126
B.1	Cluster content for all Cases and feature combinations for the starless core B68.	140
B.1	continued.	141
B.2	Same as Table B.1, but for the starless core L1521E.	141
B.2	continued.	142
B.2	continued.	143
B.3	Same as Table B.1, but for the prestellar core L1544.	143
B.3	continued.	144
C.1	Comparison of the linewidths and velocities of the observed transitions at different locations in L1544.	157

Zusammenfassung

Die dichtesten Regionen im interstellaren Medium sind die Geburtsstätten von Sternen und Planeten. Die Analyse der chemischen und physikalischen Prozesse des Gases, aus dem das interstellare Medium besteht, hilft dabei, die Vererbung und Weiterentwicklung dieses Materials während des Prozesses der Stern- und Planetenentstehung zu verstehen. Moleküle sind nützliche Diagnosewerkzeuge, um die Entwicklung und die Kinematik in den frühesten Stadien der Sternentstehung zu untersuchen. Sie zeichnen ein breites Spektrum an physikalischen Bedingungen auf, darunter die Temperatur-, Dichte- oder Geschwindigkeitsprofile in Molekülwolken und dichten Kernen. Diese Dissertation widmet sich der Erforschung der chemischen und physikalischen Strukturen in massearmen Sternentstehungsgebieten, mit Schwerpunkt auf der chemischen Segregation von Molekülen in den frühesten Stadien der Sternentstehung. Unter Verwendung verschiedener molekularer Indikatoren untersuche ich die Emission und die Verteilung dieser Moleküle sowie den Einfluss der Umgebung auf die chemischen Prozesse in interstellaren Wolken und Kernen. Bei der Bearbeitung der großen Beobachtungsdatensätze wende ich verschiedene analytische Methoden an, darunter die Modellierung von Strahlungstransport und Techniken des maschinellen Lernens.

Das erste Projekt befasst sich mit der Frage, ob und wie Material entlang des Prozesses der Sternentstehung vererbt wird. Dazu vergleiche ich den Grad der Deuterierung einfacher Moleküle in den Hüllen des prästellaren Kernels L1544 und des jungen protostellaren Kernels HH211. Ich finde einen ähnlichen Grad an Deuterierung in L1544 und HH211, was zeigt, dass die Deuteriumfraktionierung vermutlich in beiden Kernen gleich effizient ist. Das wiederum deutet darauf hin, dass die protostellare Hülle von HH211 die chemische Zusammensetzung des ursprünglichen prästellaren Kernels beibehält. Die Tatsache, dass die beiden Kerne in unterschiedliche Molekülwolken eingebettet sind, aber dennoch ähnliche Deuterierungswerte aufweisen, deutet darauf hin, dass die Umgebungsbedingungen keinen signifikanten Einfluss auf die Deuteriumfraktionierung in dichten Kernen haben.

Im nächsten Projekt verwende ich dichte-basierte Clustering-Algorithmen, um die chemische Segregation von $c\text{-C}_3\text{H}_2$, CH_3OH und CH_3CCH in einer Auswahl von sternlosen und prästellaren Kernen zu analysieren. Ich finde eine signifikante chemische Differenzierung in den Kernen, die die bekannte Segregation zwischen $c\text{-C}_3\text{H}_2$ und CH_3OH reproduziert und eine nicht offensichtliche Differenzierung zwischen $c\text{-C}_3\text{H}_2$ und CH_3CCH identifiziert. Mit diesem Projekt zeige ich, dass dichte-basiertes Clustering wertvolle Einblicke in die chemischen und physikalischen Strukturen sternloser Kerne liefern kann, und dass bereits

kleine Datensätze, die nur zwei oder drei Moleküle umfassen, aussagekräftige Ergebnisse liefern können.

In meinem Abschlussprojekt verwende ich deuterierte Moleküle als Diagnosewerkzeuge, um die physikalischen und chemischen Prozesse des prästellaren Kerns L1544 zu untersuchen, wobei ich mich auf die beiden Kohlenstoffketten HC_3N und CH_3CCH konzentriere. Ich finde moderate Deuterierungsgrade für DC_3N und CH_3CCD , und signifikant höhere Werte für CH_2DCCH . Die Ergebnisse deuten darauf hin, dass die Bildung und Deuterierung von HC_3N und CH_3CCH in L1544 durch Gasphasenreaktionen dominiert wird, mit räumlichen Variationen durch die physikalische Struktur, die Dichte und externe Einstrahlung. Ähnlichkeiten zwischen der Deuterierung von CH_3CCH und CH_3OH deuten darauf hin, dass es individuelle Deuterierungsmechanismen für verschiedene funktionelle Gruppen geben könnte und weisen auf mögliche zusätzliche Deuterierungswege auf den Oberflächen von Staubkörnern hin.

Summary

Dense regions in the interstellar medium are the birthplaces of stars and planets. The analysis of the chemical and physical processes of the gas composing the interstellar medium helps to understand the inheritance and evolution of this material along the process of star and planet formation. Molecules are powerful diagnostic tools to probe the evolution and the kinematics in the earliest stages of star formation. They trace a wide range of physical conditions, including the temperature, density, or velocity structures in molecular clouds and dense cores. This thesis is dedicated to explore the chemical and physical structures in low-mass star-forming regions, focusing on the chemical segregation of molecules in the earliest stages of star formation. Using a variety of molecular tracers, I study the emission and distribution of those molecules and the influence of the environment on the chemical processes happening in interstellar clouds and cores. Handling large observational datasets, I apply different analytical methods, including radiative transfer modelling and machine learning techniques.

The first project deals with the question how or if material is inherited along the process of star formation. Therefore, I compare the level of deuteration of simple molecules in the envelopes of the evolved prestellar core L1544 and the young protostellar core HH211. I find similar levels of deuteration in L1544 and HH211, showing that the deuterium fractionation is most likely equally efficient towards both cores. This suggests that the protostellar envelope of HH211 still retains the chemical composition of the original prestellar core. In addition, the fact that the two cores are embedded in different molecular clouds but still show similar levels of deuteration indicates that the environmental conditions do not have a significant effect on the deuterium fractionation within dense cores.

In the next project, I use density-based clustering algorithms to analyse the chemical segregation of $c\text{-C}_3\text{H}_2$, CH_3OH , and CH_3CCH in a sample of starless and prestellar cores. I find significant chemical differentiation across the cores, reproducing the known segregation between $c\text{-C}_3\text{H}_2$ and CH_3OH , and identifying a non-apparent differentiation between $c\text{-C}_3\text{H}_2$ and CH_3CCH . With this study, I show that density-based clustering can provide valuable insights into the chemical and physical structures of starless cores, and demonstrate that already small datasets covering only two or three molecules can yield meaningful results.

In the final project, I use deuterated molecules as diagnostic tools to study the physical and chemical processes in the prestellar core L1544, focusing on the two carbon chains HC_3N and CH_3CCH . I find moderate levels of deuteration for DC_3N and CH_3CCD , and

significantly larger values for CH_2DCCH . The results suggest that gas-phase reactions dominate the formation and deuteration of both HC_3N and CH_3CCH in L1544, with spatial variations driven by physical structure, density, and external radiation. Similarities between the deuteration of CH_3CCH and CH_3OH suggest that there might be individual deuteration mechanisms for different functional groups and hint to possible additional deuteration pathways of CH_3CCH on the surfaces of dust grains.

Chapter 1

Introduction

In this introduction, I present the main concepts and methods used in this thesis for the analysis of the chemical and physical structures of low-mass star-forming regions. Section 1.1 focuses on the interstellar medium and molecular clouds and how they lead to the formation of stars and planets. Section 1.2 introduces the field of astrochemistry and how molecules can be used as diagnostic tools in the process of star formation. Section 1.3 describes the radiative processes in molecular clouds and how they are used to analyse astronomical objects, and introduces the concepts of radio astronomy. Section 1.4 provides a short overview of machine learning techniques in astrophysics, focusing specifically on unsupervised learning. Section 1.5 describes the content of this thesis.

1.1 The interstellar medium and low-mass star formation

Observations of the Planck satellite have shown that dark energy and dark matter make out approximately 95% of the total mass-energy content of the universe. In contrast to that, baryons, that is atoms and molecules, contribute only 5%. Despite this small fraction, baryons make up all of the visible universe, including galaxies, stars, or planets, and as such, we are able to study the universe through them. Of the baryonic mass, about 98% is composed of hydrogen and helium, while heavier elements such as carbon, nitrogen, oxygen, and sulfur contribute just about 2%. In the Milky Way, roughly 90% of the baryonic mass is locked in stars, while the remaining 10% forms the interstellar medium (ISM). In this thesis, I focus on the physical and chemical processes occurring in the densest regions of the ISM, specifically in cold cores embedded in molecular clouds.

Interstellar medium

The interstellar medium consists of the gas, dust, and radiation inbetween stars. The gas component of the ISM consists of ions, atoms, and molecules in the gas phase, with the biggest contribution made by atomic, molecular, and ionised hydrogen. The ISM has a

wide range of temperatures and densities that are categorised into roughly four different phases and can be differentiated depending on their temperature, density and chemical properties (McKee & Ostriker 1977; Draine 2011):

- Hot ionised medium: Low density gas ($n \sim 0.004 \text{ cm}^{-3}$) that gets heated by shocks from supernova blastwaves to temperatures $\geq 10^5 \text{ K}$ and is ionised by collisions.
- Warm ionised medium: Diffuse and warm ionised H ($n = 0.3 - 10^4 \text{ cm}^{-3}$), where the hydrogen gets ionised by UV photons from the interstellar radiation field.
- Warm and cold neutral medium: Mostly atomic gas that is heated by photoelectrons from dust, with temperatures of around 5000 K and 100 K and densities of approximately 0.6 cm^{-3} and 30 cm^{-3} for the warm and the cold component, respectively.
- Molecular gas: The densities are high enough ($n = 100 - 10^6 \text{ cm}^{-3}$) for self-shielding of H_2 molecules, which allows H_2 to be abundant, while cooling by CO leads to low temperatures of 10 – 50 K.

The different phases coexist in a pressure equilibrium, through the balance of heating and cooling processes.

The coldest and densest regions of the ISM, molecular clouds, serve as the birthplaces of stars and planets. Molecular clouds are mostly grouped together in big structures called Giant Molecular Clouds (GMCs), complexes that can have an expansion of around 50 pc, and contain up to 10^5 solar masses (Stahler & Palla 2004). In the Milky Way, approximately 22% of the interstellar gas is in molecular clouds. A well-studied example of a GMC complex in the Milky way is the Orion molecular cloud. Unlike the phases of the ISM, molecular clouds are regulated by gravity instead of pressure. Molecular clouds are known to show both clumpy and filamentary substructures. Filaments are elongated, thread-like structures of dense gas and dust that promote the formation of stars in them, spanning from 0.1 to several parsecs in length and with typical widths of around 0.1 pc, visible in dust emission (André et al. 2014). Clumps are overdense regions ($n \approx 10^3 \text{ cm}^{-3}$), characterising structures that may or may not form stars, with sizes of around 1-2 pc and 100 solar masses (Draine 2011). Local density peaks ($n \geq 10^4 \text{ cm}^{-3}$) within clumps are called cores, and describe individual star-forming sites for single- or multi-stellar systems. They are mostly found within filaments and have sizes of around 0.1 pc.

The dust component of the ISM contributes to about 1% of its total mass. Interstellar dust grains are solid particles that are mainly made of carbonaceous species and silicates. The grain cores have a typical size of $0.1 \mu\text{m}$, and mainly form from gases of refractory elements such as Mg, Si, O, and C around dying stars (Draine 2011). In molecular clouds, the grains can be covered by mantles of H_2O ice mixed with various molecules. Dust grains can scatter and absorb interstellar UV radiation, protecting molecules from photodissociation. The grains re-emit the absorbed background light as thermal radiation. This emission can provide important information about the physical properties of the ISM, for example the dust temperature or gas column density.

The ISM is penetrated by highly energetic cosmic-rays and UV photons. Cosmic rays are charged particles, mostly protons, helium nuclei and electrons, that are of solar, galactic, or extragalactic origin (Draine 2011). Due to their high energies, they can penetrate deep even into the denser parts of the ISM, ionising and dissociating interstellar atoms and molecules. Through collisions, they can exchange energies and eject energetic electrons into the gas, which makes them an important heating source, especially in dark and dense material.

Cloud stability

The stability of a molecular cloud can be described by the *virial theorem*, which relates the energy within a system of particles with the moment of inertia I of the system (following Stahler & Palla 2004):

$$\frac{1}{2} \frac{\partial^2 I}{\partial t^2} = 2\mathcal{U} + \mathcal{W} , \quad (1.1)$$

where $\mathcal{U}$ is the thermal energy (related to the internal pressure of the gas), and $\mathcal{W}$ the gravitational potential energy (i.e. the self-gravity of the cloud). This form of the virial theorem is an approximation only considering pressure and gravity, ignoring other forms of energy such as magnetic fields for simplicity. If the cloud is in virial equilibrium, the sum is equal to zero, which means the moment of inertia is not varying in time. This results in $|\mathcal{W}| = 2\mathcal{U}$. For a simple homogeneous and spherical molecular cloud of mass M , temperature T , and size R , the *virial parameter* α measures if the cloud is gravitationally bound or collapsing:

$$\alpha = \frac{2\mathcal{U}}{|\mathcal{W}|} \approx \left(\frac{M k_B T}{\mu m_H} \right) \cdot \left(\frac{G M^2}{R} \right)^{-1} \quad (1.2)$$

$$\approx 10^{-2} \left(\frac{M}{10^5 M_\odot} \right)^{-1} \left(\frac{R}{25 \text{ pc}} \right) \left(\frac{T}{15 \text{ K}} \right) , \quad (1.3)$$

where μ is the gas mean molecular weight, m_H the hydrogen atom mass, k_B the Boltzmann constant, and G the gravitational constant. Typical physical conditions of molecular clouds are around $10^4 - 10^5 M_\odot$, $10 - 20 \text{ K}$, and $10 - 50 \text{ pc}$ (Stahler & Palla 2004), which results in $\alpha < 1$. This means, the gravitational pressure is significantly larger than the internal thermal pressure, which triggers the collapse of the cloud.

When gravity dominates over thermal pressure, Eq. 1.1 reduces to:

$$\frac{\partial^2 I}{\partial t^2} \approx \mathcal{W} = \frac{G M^2}{R} . \quad (1.4)$$

If I is approximated by $M R^2$, the *free-fall time* can be derived as:

$$t_{\text{ff}} \approx \left(\frac{R^3}{G M} \right)^{1/2} = 7 \times 10^6 \text{ yr} \left(\frac{M}{10^5 M_\odot} \right)^{-1/2} \left(\frac{R}{25 \text{ pc}} \right)^{3/2} , \quad (1.5)$$

which describes the time a body needs to collapse under its own gravitational attraction, if no other forces such as pressure prevent the collapse. Since $M/R^3 \approx \rho$, the free-fall time can also be expressed in terms of density:

$$t_{\text{ff}} = \left(\frac{3\pi}{32G\rho} \right)^{1/2}. \quad (1.6)$$

The timescale is proportional to $\rho^{-1/2}$, which means the higher the density of a cloud or core, the earlier it collapses.

The maximum size of a spherical cloud before it collapses is described by the *Jeans mass*, given by the equilibrium between gravitational and thermal pressure (i.e. $\alpha = 1$):

$$M_J \approx 1M_{\odot} \left(\frac{T}{10 \text{ K}} \right)^{3/2} \cdot \left(\frac{n}{10^4 \text{ cm}^{-3}} \right)^{-1/2}. \quad (1.7)$$

If the mass of a cloud core is below the Jeans mass, the cloud is unbound and expands. If the cloud reaches the critical mass, the gravitational pressure exceeds the internal gas pressure. The core becomes unstable and collapses in a runaway contraction, until it is stopped by other forces, for example by thermal pressure when the first hydrostatic core forms. Clouds with a fixed given mass collapse if they are dense and cool enough.

Similarly, the *Jeans length* gives information about the maximum stable size of a cloud before collapse:

$$\lambda_J \propto M_J^{-1/3} \approx 0.2 \text{ pc} \left(\frac{T}{10 \text{ K}} \right)^{1/2} \left(\frac{n}{10^4 \text{ cm}^{-3}} \right)^{-1/2}. \quad (1.8)$$

The problem with this approximation is that roughly all observed molecular clouds are highly Jeans unstable. This would lead to a star formation rate of more than $30 M_{\odot} \text{ yr}^{-1}$ in the Milky Way (Zuckerman & Palmer 1974), which is an order of magnitude higher than the observed $1 M_{\odot} \text{ yr}^{-1}$ (Robitaille & Whitney 2010). This indicates that additional forces are present that oppose the gravitational pressure. A more general version of the virial theorem describes this equilibrium:

$$2\mathcal{U} + \mathcal{W} + 2\mathcal{K} + \mathcal{M} = 0, \quad (1.9)$$

which takes into account non-thermal motions ($\mathcal{K}$), such as rotation or turbulence, and magnetic field pressure ($\mathcal{M}$). In typical molecular cloud conditions, the gas motions related to non-thermal effects and magnetic fields can provide a balance to the gravitational pressure, and stabilise the cloud. In the ISM, turbulence can for example be induced by stellar outflows or supernovae, and is mostly supersonic (Larson 1981). The Planck mission showed that magnetic fields are ubiquitous in the ISM, penetrating and shaping structures at all scales (Planck Collaboration et al. 2016). However, the exact origin of both magnetic fields and turbulence in the ISM is still under debate.

Low-mass star formation

In this thesis, I focus on the formation of low-mass stars, which are stars with up to eight solar masses. Stars can either form in isolation, in small groups, or in massive clusters and associations. If a part of a cloud exceeds the critical Jeans mass, it can become unstable locally, and fragment into smaller, self-gravitating subregions. These fragments then can each collapse on their own and form protostars or fragment even further to become binary or multiple systems.

Figure 1.1 illustrates the evolution of a dense core into a low-mass star (following Öberg & Bergin 2021). The starting point of star formation in the densest regions of molecular clouds are dense cores, also referred to as *starless cores* (Fig. 1.1a), which are gravitationally bound cloud fragments. When such a dense core exceeds the critical Jeans mass, gravitational pressure dominates over turbulence, thermal and magnetic pressure, and the core becomes unstable and starts collapsing (Fig. 1.1b). The first collapse phase is isothermal, because the heating of the gas by the compression is balanced by cooling from molecular line emission and dust emission. Therefore, the released gravitational energy is essentially all radiated away, and the temperature stays constant at about 10 K (Bodenheimer 2011). Unfortunately, there is no analytical model that incorporates the combined effects of gravity, thermal pressure, rotation, magnetic fields, and cooling to explain the process of a core collapse self-consistently. The first analytical models were published by Shu (1977), and Larson (1969) and Penston (1969), and describe the collapse of a singular isothermal sphere with no rotation, magnetic fields, turbulence, or environmental influence. Later, Terebey et al. (1984) created a model that also includes the effects rotation by using perturbation theory. The timescale for a pressure-less isothermal collapse is given by the free-fall time (see Eq. 1.6), which means, denser parts collapse faster. This results in an inside-out collapse, where the infall starts at the centre of the core, and expands outwards with sound speed (= “collapse wave”).

When the density in the centre becomes high enough that the gas becomes optically thick, the adiabatic collapse phase begins (e.g. Young 2023). The released gravitational energy can no longer radiate away, and instead heats up the core. Pressure and temperature rise rapidly, and the collapse pauses when a quasi-hydrostatic equilibrium is reached, forming a pressure supported body, also referred to as the *first hydrostatic core*. The envelope material continues to fall onto the first core, further increasing mass and temperature. Once the central temperature exceeds about 2000 K, molecular hydrogen dissociates. This means the released gravitational energy is used as dissociation energy rather than thermal pressure, and leads to a second collapse. The collapse continues until all H_2 is dissociated at the centre, and the material is in hydrostatic equilibrium, forming a second stable core, also known as *protostar*.

In the following stage of evolution, known as the main accretion phase, the central object continues to accrete mass from the surrounding cloud core or envelope (Fig. 1.1c). To preserve the angular momentum of the core, some of the accreting material forms a circumstellar disk. Simultaneously, bipolar outflows and jets are launched perpendicular to the disk, which also remove angular momentum from the system. Jets consist of

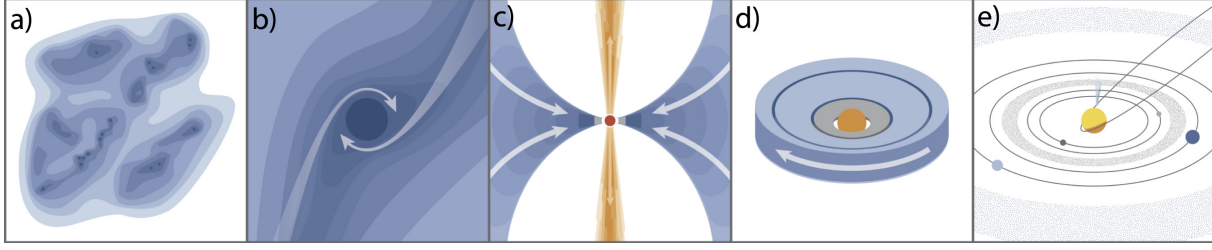


Figure 1.1: The different stages of low-mass star formation: a) clumps and dense cores in a molecular cloud, b) collapse of a dense core, c) protostellar core with accretion and outflow, d) pre-main sequence star with circumstellar disk and planet formation, e) main-sequence star with a planetary system. Figure credit: Öberg & Bergin (2021).

high-velocity particles and are highly collimated, while outflows describe low-velocity, expanding (molecular) gas. The properties of jets and outflows depend on the evolutionary stage of the protostar. When most of the envelope is dissipated or has been accreted, a *protoplanetary disk* remains (Fig. 1.1d), where planets form. In this phase, the disk material is further accreted onto the star, but also onto protoplanets, and is dispersed by interactions with stellar winds. The central protostar continues to contract, heating up and starting to burn deuterium, which is known as the *pre-main sequence* phase (Stahler & Palla 2004). With the ignition of hydrogen fusion at temperatures of 10^7 K, the object becomes a main-sequence star with a planetary system, and the star-formation process is completed (Fig. 1.1e).

In recent years, this classical picture of low-mass star formation has been challenged, as research suggests that a collapsing cloud core is in fact not spherically symmetric, but shows asymmetric infall structures (e.g. Pineda et al. 2020, 2023; Valdivia-Mena et al. 2023). In addition, observations indicate that planet formation may actually already begin much earlier, namely in the protostellar phase (e.g. ALMA Partnership et al. 2015; Segura-Cox et al. 2020; Sheehan et al. 2020).

In this thesis, I focus on the initial stages of star formation, covering the starless and the protostellar phase. In general, there are two classes of starless cores (following Keto & Caselli 2008) – starless and prestellar cores. Starless cores are thermally subcritical, which means their thermal pressure is strong enough to resist a gravitational collapse, creating stable, gravitationally bound objects with internal oscillations. Their structure can be described by a Bonnor-Ebert sphere (Ebert 1955; Bonnor 1956), a stable isothermal sphere in pressure equilibrium with its surroundings. Prestellar cores, on the other hand, are thermally supercritical and therefore dynamically unstable. Their masses are generally greater than the Bonnor-Ebert mass, which inevitably leads to a gravitational collapse. They show a compact, centrally peaked density distribution, and signs of contraction. Due to the lack of protostellar outflow and feedback in the starless phase, starless and prestellar cores are ideal laboratories to study the physical and chemical processes and structures in the early stages of (low-mass) star formation.

In the protostellar phase, I focus on the youngest objects, in particular Class 0 objects.

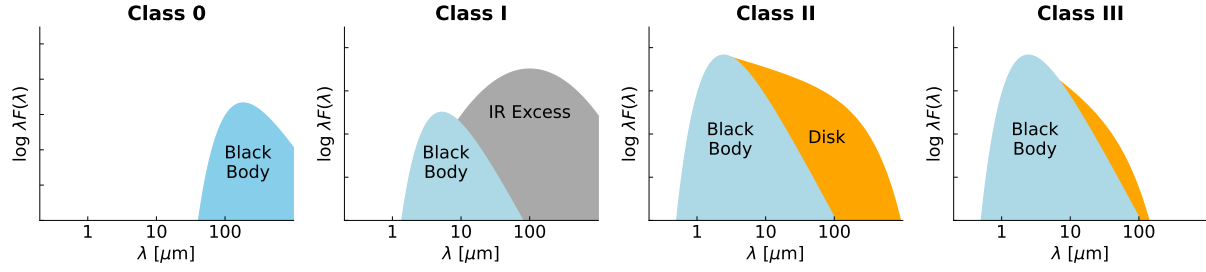


Figure 1.2: Expected spectral energy distribution at infrared (IR) and submillimeter wavelengths for each observational protostar class. The figure is based on Persson (2014).

The evolutionary stages of protostars are characterised by their emission in far infrared wavelengths. In practice, the shape of the spectral energy distribution across optical, infrared (IR), and submillimeter wavelengths is used to classify protostars into four different categories, as illustrated in Fig. 1.2 (McKee & Ostriker 2007):

- Class 0: the protostar is deeply embedded in a dense envelope. The emission comes mostly from the envelope and is only detectable above $10 \mu\text{m}$.
- Class I: less than 50% of the stellar mass is in the envelope. Emission from the star and disk are detectable, but are dominated by a large IR excess from the remaining envelope.
- Class II (or T Tauri star): the envelope is almost fully removed by accretion flows and stellar feedback. The detected signal is dominated by emission from the star and the disk.
- Class III: the disk is cleared from remnant material, there is no significant accretion onto the star. The spectrum shows the blackbody radiation from the star and almost no IR excess.

However, it has to be noted that this is an empirical categorisation correlated to time evolution. Therefore, it can also be affected by the environmental conditions present (e.g. Kuffmeier et al. 2023).

1.2 Astrochemistry

"Astrochemistry is the study of the formation, destruction, and excitation of molecules in astronomical environments and their influence on the structure, dynamics, and evolution of astronomical objects" – Alexander Dalgarno, 2008

In the 1930's, the discovery of the first astronomical radio source by Karl Jansky marked the beginning of a new area in astronomy. Using radio telescopes, the emerging field of radio astronomy observed astronomical objects in the centimeter, millimeter, and submillimeter

range of the frequency spectrum. In the following decades, the new technology grew to greater and greater importance, making groundbreaking and spectacular discoveries, such as the cosmic microwave background radiation in 1964. Astrochemistry, or molecular astrophysics, began in the late 1930's, with the discovery of CH, CH⁺, and CN in diffuse clouds (Dunham 1937; Adams 1941). Since those first detections, more than 300 molecules have been observed in space¹, in a wide range of astronomical objects: clouds, stars, supernovae, planets, comets, disks, jets, galaxies (for an overview see e.g. McGuire 2022).

The many detections of molecules show that molecules are omnipresent in space and that, in fact, chemical and physical structures are connected and cannot be understood separately. This gives rise to a number of interesting questions that drive current research in astrochemistry: How do molecules form in the ISM? How are the physical and chemical structures connected and how do they affect each other? How does the chemical composition of a dense core change during the process of star formation? Is the chemical complexity in protoplanetary disks and on exoplanets inherited from the early phases of star formation? Are complex molecules detected in space linked to the origin of life on Earth?

In the ISM, chemical reactions can occur either in the gas phase or on the surfaces and in the icy mantles of dust grains. Interstellar chemistry starts with the formation of the most abundant molecule in the universe, H₂. In the ISM, it is formed on the surface of dust grains via the reaction $\text{H} + \text{H} \rightarrow \text{H}_2$ (Watanabe & Kouchi 2008). The formation of other, more complex molecules is initiated by the formation of H₃⁺, the most important molecular ion in interstellar chemistry, by the ionisation of H₂ with cosmic rays followed by a reaction with H₂. Because H₃⁺ can easily donate a proton, it is very reactive and drives the formation of larger molecules (Yamamoto 2017). The most common elements forming interstellar molecules reflect the most abundant elements in space: hydrogen, carbon, oxygen, and nitrogen. On the one hand, these elements form ordinary molecules that are also common on Earth, such as ammonia (NH₃), water (H₂O), formaldehyde (H₂CO), or ethanol (CH₃CH₂OH). On the other hand, also more exotic molecules can be formed that are either rare or unstable under terrestrial conditions, for example ions like HCO⁺ and N₂H⁺, or long carbon chains such as HCCCCCN. In the extreme conditions of interstellar clouds (low density, low temperature), even reactive species such as free radicals or molecular ions can survive (Yamamoto 2017). A characteristic feature of interstellar chemistry is the high abundance of unsaturated hydrocarbons, for example carbon chains (e.g. HCCCN). The detection of so-called complex organic molecules (COMs), which contain at least six atoms including carbon (Herbst & van Dishoeck 2009), in interstellar environments such as dense cores demonstrates how complex the chemical evolution can get even under the extreme conditions in space (e.g. Jiménez-Serra et al. 2016; Scibelli et al. 2024).

In interstellar chemistry, mostly two-body reactions are relevant, because even the densest phases in star-forming regions are too diffuse for three-body reactions to be efficient².

¹See e.g. molecules-in.space for the latest detections

²Three-body reactions become relevant above number densities of 10^{13} cm^{-3} , which are for example

Table 1.1: Astrochemical gas-phase reactions.

Category	Type	Reaction
Formation of bonds	Radiative association	$X^+ + Y \rightarrow XY^+ + h\nu$
	Associative detachment	$X^- + Y \rightarrow XY + e^-$
Destruction of bonds	Photodissociation	$XY + h\nu \rightarrow X + Y$
	Dissociative recombination	$XY^+ + e^- \rightarrow X + Y$
	Collisional dissociation	$XY + M \rightarrow X + Y + M$
Rearrangement of bonds	Ion-molecule reactions	$X^+ + YZ \rightarrow XY^+ + Z$
	Charge-transfer reactions	$X^+ + YZ \rightarrow X + YZ^+$
	Neutral-neutral reactions	$X + YZ \rightarrow XY + Z$

This limits the types of reactions that can occur in the ISM and subsequently the formation pathways of molecules (Herbst & Klemperer 1973). An overview of the gas-phase chemical reactions possible in space is given in Table 1.1. Due to the low temperatures present in the early stages of star formation (~ 10 K; Bodenheimer 2011), the occurring reactions are mostly exothermic or with (nearly-) vanishing activation energies. Therefore, the most important pathways to form molecules in such environments are ion-molecule reactions (Puzzarini 2022). Here, the ion polarises a nearby neutral atom or molecule, which leads to a long-range attractive force that can overcome possible activation energies. Any excess reaction energy is then dissipated as kinetic or vibrational energy.

Molecules are destroyed by absorption of photons (photodissociation), collisions with electrons (dissociative recombination), and in shocks by high-velocity collisions (Tielens 2021). In very low density regions like the diffuse ISM, molecules only have a very short lifetime, as they are quickly destroyed by penetrating UV photons. In higher-density regions like molecular clouds or dense cores, however, molecules are shielded from the ambient radiation by H_2 molecules and dust grains and can survive.

In general, interstellar dust grains play a fundamental role in the formation of molecules and the regulation of chemistry. Via collisions, molecules can stick onto the surfaces of dust grains due to weak van der Waals or electrostatic forces (Yamamoto 2017). This leads to a depletion of the molecule from the gas phase and changes the observability of the species. On the surface, the adsorbed molecules build icy mantles around the grains, which is why this process is also called molecular freeze out. The grain surfaces themselves provide large reaction sites for various chemical processes and act as catalysts. The adsorbed molecules can move across the surface via thermal hopping or quantum-tunneling effects, which results in a complex solid-state chemistry and enriches the chemical complexity of molecular clouds and cores, for example with COMs. Molecules can be liberated into the gas phase via thermal and non-thermal desorption processes (Tielens 2021). Thermal desorption happens when the thermal energy of the molecule becomes larger than its binding energy

present in the atmospheres of stars or planets (Puzzarini 2022).

to the grain. The main non-thermal processes are desorption by excess energy of chemical reactions (chemical desorption) and desorption by cosmic rays or cosmic ray-induced UV radiation (photo-desorption).

The field of astrochemistry is a very interdisciplinary science, because the interpretation of observed molecular emission and abundances requires a profound understanding of the underlying astrochemical processes. In addition to telescope observations, both modelling and laboratory experiments are needed to obtain a complete picture of the physical and chemical structures present. The molecular information needed to interpret the astrochemical observations is provided by different databases, for example:

- Spectroscopic data:

Cologne Database for Molecular Spectroscopy (CDMS; Müller et al. 2001)

Jet Propulsion Laboratory molecular spectroscopic database (JPL; Pickett et al. 1998)

- Collision excitation data:

Leiden Atomic and Molecular Database (LAMDA; Schöier et al. 2005)

Excitation of Molecules and Atoms for Astrophysics database (EMAA; EMMA 2021)

- Chemical reaction data:

KInetic Database for Astrochemistry (KIDA; Wakelam et al. 2012)

UMIST database for astrochemistry (UDFA; Millar et al. 1991)

In this thesis, I focus on the observations and the modelling of the chemical (and physical) structures in the early stages of star formation, namely starless, prestellar, and protostellar cores. In the starless and prestellar phase, the cores consist mostly of cold and dense gas, which drives the formation of simple molecules (e.g. Caselli et al. 2012). Typical chemistry in these phases involves carbon and nitrogen chemistry, leading to abundant species such as CO, N_2H^+ , and NH_3 . In the central regions of these cores, the density increases above 10^5 cm^{-3} , and the temperature drops below 10 K (e.g. Crapsi et al. 2007). This leads to the freeze-out of atoms and molecules from the gas phase onto the surfaces of dust grains, forming icy mantles. In the centre of very evolved cores, almost 99% of all species heavier than Helium are frozen out (Caselli et al. 2022). Due to the cold temperatures and high visual extinctions in prestellar cores, the frozen-out species cannot desorb from the grains, leading to a significant drop of their gas-phase abundances and thick icy mantles on dust grains.

In the protostellar phase, the core collapses and the released gravitational energy is converted into radiation. This warms up the envelope, with the temperature increasing towards the centre. In the outer regions of the envelope, the temperature and density are still similar to the conditions in the prestellar phase, resulting in a similar chemical composition (see e.g. Chapter 2). In the inner regions, whenever the dust temperature

risers above their sublimation temperature, the molecules frozen-out on the grains sublimate and re-enter the gas phase, creating a rich chemistry. In the warm and dense gas, new and more complex molecules are formed. The sublimation of methane (CH_4), for example, drives the formation of carbon-chain molecules (Caselli et al. 2012).

1.2.1 Fractionation processes

Isotopic ratios are set by the primordial nucleosynthesis, right after the Big Bang, or by nuclear reactions in stars. In molecular clouds, though, there are no processes happening with enough energy to drive nuclear reactions. However, specific chemical reactions can change the isotopic ratio of individual molecular species, which is called *fractionation*. Molecules with different isotopic composition are known as *isotopologues*, for example H_2O and HDO , which are isotopologues of water. Elements typically have more than one stable isotope, which can lead to multiple isotopologues of one species. Isotopologues of the same molecule have different zero-point energies, because the isotopes differ in mass. This leads to a slightly different molecular geometry and creates an individual emission spectrum for each isotopologue.

In interstellar clouds, isotopic fractionation is an important aspect of the chemical processes present. Different fractionation processes happen in very specific conditions – therefore, isotopologues and isotopic ratios of molecules can be used as a diagnostic tool to study the different evolutionary stages of star formation (see e.g. Ceccarelli et al. 2014). Comparing isotopic ratios of molecules at different phases in the star-formation process can help to understand when and where those molecules formed and how (or if) chemical compositions are inherited from molecular clouds to planetary systems. In addition, variations in the isotopic ratios can give information about the link between Solar System objects and galactic interstellar environments (e.g. Caselli et al. 2012; Hily-Blant et al. 2013; Ceccarelli et al. 2014; Colzi et al. 2018).

A common usage of rarer isotopologues is the estimation of the abundance of their parent species. In some cases, the main isotopologue is very abundant and optically thick, which leads to an underestimation of the molecular column density. The rarer isotopologues, however, can still be optically thin. The column density of the main species is then approximated by multiplying the column density of the isotopologue with a standard isotope ratio measured in the medium. In the local ISM, for example, the $^{12}\text{C}/^{13}\text{C}$ ratio is assumed to be around 68 ± 15 (Milam et al. 2005), which was derived from the isotopic ratios of CO , CN , and H_2CO in Galactic molecular clouds. For individual cores, however, isotopic ratios can deviate significantly from those standard values, as they are molecule-dependent and vary with time, volume density, and temperature (e.g. Colzi et al. 2020).

In the following, I introduce the specific isotopic fractionation processes that are relevant in this thesis (see Chapters 2 and 4), specifically deuterium, carbon, and oxygen fractionation.

Deuterium fractionation

Deuteration is the process where hydrogen atoms in molecules get replaced by their next heavier isotope, deuterium. Deuterium has an abundance of around $1.6 \cdot 10^{-5}$ with respect to hydrogen in the Local Bubble³ (Linsky 2003), which was set by the nucleosynthesis after the Big Bang. However, the molecular deuterium abundance can be enhanced by the process of deuterium fractionation. Towards dense objects like prestellar cores, it can reach significantly higher levels than the elemental value (> 0.1 , e.g. Ceccarelli et al. 2014). In general, those cold and dense environments ($T \leq 15$ K, $n \geq 10^4$ cm⁻³) host ideal conditions for efficient deuteration. As illustrated in Fig. 1.3, the process of deuteration consists of three basic steps:

1. The formation of H_3^+ through the ionisation of H_2 and H by cosmic rays.
2. The formation of H_2D^+ via the isotope exchange reaction (e.g. Watson 1976; Dalgarno & Lepp 1984):



Similarly, D_2H^+ , and D_3^+ are formed by reactions with HD.

3. The formation of other D-bearing molecules by reaction of H_2D^+ , D_2H^+ , and D_3^+ with other molecules and atoms.

In the backwards direction, Reaction 1.10 is endothermic, which means at the low temperatures present in prestellar environments, it is suppressed. Over time, this leads to an enhancement of H_2D^+ with respect to H_3^+ , which drives the deuteration of other molecules. The H_2D^+ ion is mostly destroyed by CO, the second most abundant molecule in the ISM. However, in cold and dense regions like pre-stellar cores, CO is highly depleted from the gas phase and frozen out onto dust grains (e.g. Caselli et al. 2022), which further enhances the abundance of H_2D^+ and subsequently of deuterated molecules. In addition, the level of deuteration depends on the *ortho-to-para* ratio of H_2 . The higher this ratio, the more H_2D^+ is destroyed, as the endothermicity of the backwards reaction of Reaction 1.10 (~ 230 K) is close to the energy difference of the $(1 - 0)$ rotational transition of *ortho*- H_2 , ~ 170 K (Kong et al. 2015). However, the *ortho-to-para* ratio of H_2 in pre-stellar cores is usually around 10^{-4} (e.g. Sipilä et al. 2015), which further favours the deuterium fractionation.

On grain surfaces or in mantles, molecules are deuterated by addition or substitution of atomic D. The necessary deuterium atoms are created by the dissociative recombination of H_2D^+ , D_2H^+ , and D_3^+ , which leads to a significant increase of the local atomic D/H ratio in the gas phase (Ceccarelli et al. 2014). The D atoms then freeze out onto the dust grains, where they can diffuse very efficiently and deuterate molecules within the ice layers. A well-known example for this is methanol (CH_3OH), which is both formed and deuterated on the surfaces of dust grains (Watanabe & Kouchi 2008), by subsequent addition of H or

³The Local Bubble is a structure in the interstellar medium of the Orion arm of the Milky Way. It contains our Solar System and has an extension of roughly 300 pc.

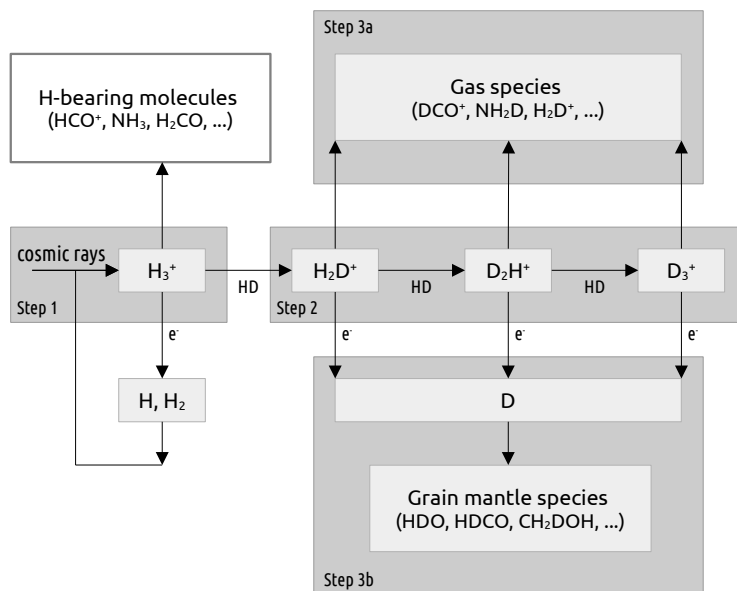


Figure 1.3: Scheme of the three basic steps of deuterium fractionation: 1) formation of H_3^+ , 2) formation of H_2D^+ (and D_2H^+ , D_3^+), 3) formation of other D-bearing molecules. The figure is mainly based on Ceccarelli et al. (2014).

D atoms to CO. In those cases, the deuterium fraction depends on the atomic D/H ratio present in the gas phase.

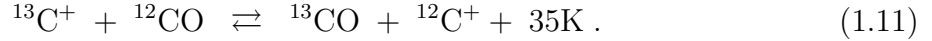
The deuterium fraction of a molecular tracer is measured by its D/H ratio, dividing the column density of the deuterated isotopologue by the column density of the normal species. The deuterium fraction can be used to trace the evolutionary stages of dense cores and protostars (e.g. Crapsi et al. 2005; Ceccarelli et al. 2014). In the cold and dense conditions of starless and prestellar cores, the level of deuteration of some molecules can build up with time, leading to characteristically high values, which was for example observed for N_2H^+ (Crapsi et al. 2005). Due to this correlation between the evolutionary stage of a starless/prestellar core and its deuteration level, deuterated species are very good tracers of the high-density central regions of these cores. They can be used to probe the kinematics of a core or to measure the elusive electron number density, and spatial variations in the D/H ratio can indicate temperature gradients or dynamical processes.

In protostellar cores, however, the cold envelope is heated by the new central object, which causes the dust temperatures to increase, and eventually, CO and other molecules to sublime back into the gas phase. Due to the increasing temperatures, the backwards reaction of Reaction 1.10 becomes more efficient, reducing the abundance of H_2D^+ . This eventually leads to a decrease of the total degree of deuteration. Therefore, the high levels of deuteration observed in protostellar cores and envelopes are inherited from the prestellar phase (see Chapter 2), and differences in the deuterium fractionation can tell about the

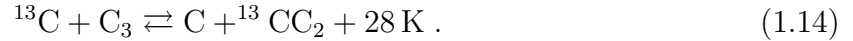
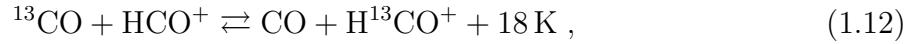
history of a core.

Carbon fractionation

In dense clouds, most of the carbon is locked up in stable molecules, such as CO or carbon chains (Colzi et al. 2020). In the gas phase, carbon isotope fractionation is driven by the isotope exchange reaction (Watson 1976):



Astrochemical models have shown that, in dark cloud conditions, this reaction divides carbon-bearing molecules into two groups (e.g. Langer et al. 1984): molecules formed from CO are enriched in ^{13}C , molecules formed from C^+ are poor in ^{13}C . However, recent studies have identified several other important isotope exchange reactions (Mladenović & Roueff 2014; Roueff et al. 2015; Colzi et al. 2020; Loison et al. 2020), such as:



Carbon fractionation is a useful tool to study the formation pathways of carbon-chain molecules. For this, isotopomers are especially useful, which are species with the same chemical formula, but different position of the ^{13}C , for example ^{13}CCH and C^{13}CH . Measurements of the $^{12}\text{C}/^{13}\text{C}$ ratio in such molecules showed that the abundances of the isotopomers depend on the position of the substituted ^{13}C , providing information about the specific formation pathways (e.g. Sakai et al. 2010; Taniguchi et al. 2019; Chapter 2 of this work).

Oxygen fractionation

Oxygen isotope fractionation is believed to be caused by isotope selective photodissociation of CO (Nomura et al. 2022). In dense regions, CO photodissociation is prevented by self-shielding, which means CO layers in the outer regions of cores absorb the dissociating UV radiation, protecting the inner CO layers. However, rarer isotopologues such as C^{18}O are much less abundant ($^{16}\text{O}/^{18}\text{O} = 557$ in the local ISM, Wilson 1999) and are therefore not self-shielded until much higher densities with respect to the main species. This results in a lower abundance of C^{18}O compared to CO or ^{13}CO with respect to the reference value (e.g. Shimajiri et al. 2014; Li et al. 2018; Yamagishi et al. 2019; Giers et al. 2023).

1.3 Observational studies

"Hier ist wahrhaftig ein Loch am Himmel!" - Friedrich Wilhelm Herschel, 1774

In this section, I introduce the theoretical concepts needed for the observation, processing, and analysis of the astronomical data used in this thesis. More precisely, this includes the theory of radiative transfer and spectral line emission, and their application in radiative transfer modelling and radioastronomy.

1.3.1 Radiative transfer

The theory of radiative transfer asks the question how the intensity of some radiation, for example emitted by a molecule, changes when it propagates through a gas like the interstellar medium. The *intensity* is defined as energy per unit area dS , time dt , frequency $d\nu$, and solid angle $d\Omega$ (following Yamamoto 2017):

$$I_\nu = \frac{dE_\nu}{dS dt d\nu d\Omega} . \quad (1.15)$$

When crossing a small volume, the intensity can increase through emission and decrease by absorption. The resulting change in the intensity can be expressed as:

$$dI_\nu = -\alpha_\nu I_\nu ds + j_\nu ds , \quad (1.16)$$

where α_ν is the absorption coefficient, j_ν the emission coefficient, and s the path of propagation along the line of sight. From this follows the *radiative transfer equation* (RTE):

$$\frac{dI_\nu}{ds} = -\alpha_\nu I_\nu + j_\nu . \quad (1.17)$$

The absorption can also be expressed via the *opacity* τ_ν , which describes how transparent a medium is for radiation:

$$d\tau_\nu = \alpha_\nu ds , \quad (1.18)$$

or alternatively via the total opacity between s_0 and s :

$$\tau_\nu(s) = \int_{s_0}^s \alpha_\nu(s) ds . \quad (1.19)$$

Dividing the RTE by absorption α_ν then leads to:

$$\frac{dI_\nu}{d\tau} = -I_\nu + S_\nu , \quad (1.20)$$

where the *source function* S_ν is given by:

$$S_\nu \equiv \frac{j_\nu}{\alpha_\nu} . \quad (1.21)$$

If the source S_ν is known, the intensity can be derived at any position by solving the radiative transfer equation. The general solution of the RTE is given by:

$$I_\nu = I_{\nu,0} e^{-\tau_\nu} + \int_0^{\tau_\nu} S_\nu(\tau'_\nu) e^{-(\tau_\nu - \tau'_\nu)} d\tau'_\nu . \quad (1.22)$$

This means, the intensity measured by the observer (I_ν) is the sum of the background intensity ($I_{\nu,0}$) attenuated by the intervening medium and the integrated emission from the medium that is affected by absorption due to the ISM between the point of emission and the observer.

In *thermodynamic equilibrium*, the radiation is in complete equilibrium with its surrounding, and matter and radiation are perfectly coupled. There, the intensity is described by *Planck's law*, which only depends on the temperature T of its surroundings:

$$\frac{dI_\nu}{ds} = 0 \quad \Rightarrow \quad I_\nu = S_\nu = B_\nu(T) = \frac{2h\nu^3}{c^2} [e^{h\nu/k_B T} - 1]^{-1} . \quad (1.23)$$

At radio frequencies, $h\nu \ll k_B T$ holds⁴, so the Planck function can be approximated by the *Rayleigh-Jeans approximation*:

$$B_\nu(T) \simeq \frac{2\nu^2}{c^2} k_B T . \quad (1.24)$$

If a radio source can be approximated by a black body, and $h\nu \ll k_B T$ holds, the *brightness temperature* T_B is given by Eq. 1.24:

$$T_B = \frac{c^2}{2k_B \nu^2} I_\nu , \quad (1.25)$$

and can therefore directly be derived from the observed intensity I_ν .

For the analysis of observational data, the background intensity $I_{\nu,0}$ is subtracted from the observed intensity I_ν to extract the information coming from the astronomical source (see Eq. 1.22):

$$\Delta I_\nu = I_\nu - I_{\nu,0} = [B_\nu(T) - B_\nu(T_{\text{bg}})] \cdot (1 - e^{-\tau_\nu}) . \quad (1.26)$$

The background intensity is usually assumed to be equal to the cosmic microwave background radiation, which is a black-body emission with temperature $T_{\text{bg}} = 2.73$ K.

If $h\nu/k_B T \ll 1$, the expression above simplifies to:

$$\Delta I_\nu = [T - T_{\text{bg}}] \cdot (1 - e^{-\tau_\nu}) , \quad (1.27)$$

which results in a brightness temperature of:

$$T_B = \frac{c^2}{2k_B \nu^2} \Delta I_\nu . \quad (1.28)$$

⁴At a temperature of around 15 K, the approximation is ~ 0.3 for $\nu \approx 100$ GHz.

Substituting this into Eq. 1.26 leads to the radiative transfer equation expressed in terms of the observable brightness temperature:

$$T_B = [J_\nu(T) - J_\nu(T_{\text{bg}})] \cdot (1 - e^{-\tau_\nu}) , \quad (1.29)$$

which describes how the measured quantity is composed of attenuated background radiation and emission from the astronomical source, depending only on temperature T and optical depth τ_ν . To simplify the above equation, the *Rayleigh-Jeans equivalent temperature* J_ν was introduced, which is valid at radio frequencies:

$$J_\nu(T) = \frac{h\nu/k}{e^{h\nu/kT} - 1} . \quad (1.30)$$

1.3.2 Spectral line emission

There are three types of transitions producing molecular emission (e.g. Yamamoto 2017):

1. electronic: $1 - 10 \text{ eV}$, $10^4 - 10^5 \text{ K}$, visual or UV wavelengths
2. vibrational: $0.1 - 1 \text{ eV}$, $10^3 - 10^4 \text{ K}$, IR wavelengths
3. rotational: $10^{-3} - 10^{-4} \text{ eV}$, $10 - 100 \text{ K}$, cm or mm wavelengths

In the low-temperature regime of dense cores ($T < 15 \text{ K}$; Bodenheimer 2011), molecular emission is mainly caused by rotational transitions, and typically, a two-level system is used to describe the interaction of matter and radiation through emission and absorption. In this system, the upper level u and the lower level l are separated by the energy $\Delta E = E_u - E_l = h\nu$, where ν is the frequency of an emitted or absorbed photon in the case of a transition. The number of molecules populating the upper and lower level are given by the number densities n_u and n_l , respectively. Possible transitions between the two levels are:

- spontaneous emission of a photon ($n_u \rightarrow n_l$), with the transition rate described by the Einstein coefficient A_{ul} ,
- stimulated emission of a photon ($n_u \rightarrow n_l$), described by the Einstein coefficient B_{ul} ,
- spontaneous absorption of photon ($n_l \rightarrow n_u$), described by the Einstein coefficient B_{lu} ,
- emission and absorption of a photon caused by collision with other particles, described by the collisional rate coefficients γ_{ul} and γ_{lu} .

The emissivity j_ν of the spontaneous emission of photons is given by:

$$j_\nu = \frac{h\nu}{4\pi} n_u A_{ul} \phi(\nu) , \quad (1.31)$$

where $\phi(\nu)$ is the line profile function and gives the probability per frequency interval that a photon is emitted at ν . The function is normalised ($\int \phi(\nu) d\nu = 1$) and describes the shape of the broadening of an emission line around its central frequency (ν_0).

Furthermore, the Einstein coefficients can be used to express the absorption coefficient α_ν (see Eq. 1.17) in a two-level system (e.g. Mangum & Shirley 2015):

$$\alpha_\nu = \frac{h\nu}{4\pi} \phi(\nu) (n_l B_{lu} - n_u B_{ul}) = \frac{c^2}{8\pi\nu^2} \phi(\nu) \frac{g_u}{g_l} \left(1 - \frac{g_l n_u}{g_u n_l}\right) n_l A_{ul} , \quad (1.32)$$

where

$$g_l B_{lu} = g_u B_{ul} ,$$

$$B_{ul} = \frac{c^2}{2h\nu_{ul}^3} A_{ul} .$$

If we assume that the temperature is constant in a medium (= thermal equilibrium), the upper and lower level populations, n_u and n_l , are described by the Boltzmann distribution (e.g. Wilson 1999):

$$\frac{n_u}{n_l} = \frac{g_u}{g_l} \exp\left(-\frac{h\nu}{kT_{\text{ex}}}\right) , \quad (1.33)$$

where g_i is the statistical weight of each level i , and T_{ex} is the excitation temperature of the transition between the two levels. If the temperature of the gas changes on scales larger than the mean free path of its particles, it is in so-called *local thermodynamical equilibrium* (LTE). In this condition, the excitation temperature is equal to the kinetic temperature of the gas, $T_{\text{ex}} = T$, for all transitions (i.e. combinations of upper and lower levels).

Depending on its mass, geometry, and electronic structure, every molecule has a unique collection of transitions defining its emission. The column density is a common quantity used for the analysis and interpretation of this molecular emission. It is defined as the number of molecules per unit area along the line of sight. In a two-level system, the column density of molecules in the upper energy state is given by (e.g. Mangum & Shirley 2015):

$$N_u = \int n_u ds . \quad (1.34)$$

Using the definitions of the optical depth (Eq. 1.18), the Einstein coefficient A_{ul} (Eq. 1.32), and the Boltzmann distribution (Eq. 1.33), N_u can be expressed as⁵:

$$N_u = \frac{8\pi\nu^3}{A_{ul}c^3} (e^{h\nu/k_B T_{\text{ex}}} - 1)^{-1} \int \tau_\nu d\nu . \quad (1.35)$$

By taking into account the total population of all energy levels in the molecule, the total column density of the molecule N_{tot} can be derived from N_u :

$$\frac{N_{\text{tot}}}{N_u} = \frac{Q_{\text{rot}}}{g_u} e^{E_u/k_B T_{\text{ex}}} , \quad (1.36)$$

⁵A detailed derivation of this can for example be found in Mangum & Shirley (2015).

where the rotational partition function $Q_{\text{rot}} = \sum g_i e^{-E_i/k_B T}$ describes the statistical sum over all rotational energy levels of the molecule (see e.g. Mangum & Shirley 2015).

For optically thin emission ($\tau \ll 1$), the radiative transfer equation (Eq. 1.29) becomes:

$$T_B = [J(T_{\text{ex}}) - J(T_{\text{bg}})] \cdot \tau_\nu \quad (1.37)$$

With the assumption that all excitation states have the same T_{ex} , the total column density can then directly be derived from the observed emission:

$$N_{\text{tot}}^{\text{thin}} = \frac{8\pi\nu^3}{A_{ul}c^3} \frac{Q_{\text{rot}}}{g_u} [J_\nu(T_{\text{ex}}) - J_\nu(T_{\text{bg}})]^{-1} \frac{e^{E_u/k_B T_{\text{ex}}}}{e^{h\nu/k_B T_{\text{ex}}} - 1} \int T_B d\nu. \quad (1.38)$$

However, in practice, many abundant molecules like carbon monoxide are optically thick ($\tau \gg 1$), which means the line emission is absorbed or reprocessed by the material along the line of sight. This leads to a saturation of the line, where more molecules do not lead to a further increase of the line brightness, and an underestimation of the molecular column density. Therefore, astronomers often use isotopologues, where one or more atoms of a molecule are replaced by their rarer isotopes, and the respective isotopic ratio to derive column densities.

Observations of a given molecular transition are most sensitive to gas with a density near the so-called *critical density*. This quantity describes the limit where the probability of a downward transition via radiative decay is equal to the probability of a decay via collision:

$$n_{\text{crit}} = \frac{A_{ul}}{\gamma_{ul}n_{\text{col}}}, \quad (1.39)$$

where n_{col} is the density of the collision partner. In interstellar clouds, it is mostly H_2 molecules that excite and de-excite other molecules. If the volume density in a region is above the critical density ($n > n_{\text{crit}}$), collisions dominate the de-excitation of a transition. The populations of the upper and lower level are thermalised at the kinetic temperature and can be described by the Boltzmann distribution, so $T_{\text{ex}} = T_{\text{kin}}$. The density is high enough that collisions dominate, so the molecules have time to exchange energy and come into LTE. If the volume density is below the critical density ($n < n_{\text{crit}}$), spontaneous emission rates are much faster than collisions, so each collision causing a $l \rightarrow u$ transition leads to the emission of a photon. The densities are too low to thermalise the molecular levels, as there are not enough collisions to redistribute the energy efficiently. Therefore, the excitation temperature is lower than the kinetic temperature (= subthermal excitation), and the system is not in LTE. In practice, the critical density is used to estimate the density at which a particular transition is excited and observed.

Radiative transfer modelling

In general, the interstellar medium is not in thermodynamic equilibrium. Even in the denser environments of molecular clouds or dark cores, the H_2 number density is often too low to attain LTE. Therefore, to derive the level populations, and subsequently determine

the absorption and emission of the gas, the radiative transfer has to be solved in the non-LTE regime. In a two-level system, the statistical equilibrium (SE) of the level population n_i of each level i is assumed to be (following Mangum & Shirley 2015):

$$n_i \sum_j (R_{ij} + C_{ij}) = \sum_j n_j (R_{ji} + C_{ji}) , \quad (1.40)$$

where the rate coefficients R_{ij} are given by the Einstein coefficients for spontaneous and stimulated transitions:

$$R_{ij} = \begin{cases} A_{ij} + B_{ij}\bar{J}_\nu & (i > j) \\ B_{ij}\bar{J}_\nu & (i < j) \end{cases} . \quad (1.41)$$

The integrated mean intensity $\bar{J}_\nu$ is determined by integrating the intensity over the frequency range and the solid angle, to take into account the effects of non-local radiation.

The collisional coefficients C are determined by:

$$C_{ij} = \gamma_{ij} n_{\text{col}} . \quad (1.42)$$

The rate coefficients γ_{ul} (see also Sect. 1.3.2) for a transition from upper level u to lower level l have to be determined for each species separately via laboratory measurements or quantum mechanical calculations, and can for example be taken from the Leiden Atomic and Molecular Database (LAMDA; Schöier et al. 2005).

The SE equations (often) do not have an analytic solution. Therefore, they require a numerical approach, which is usually done with radiative transfer simulations, where the astronomical object is divided into a grid of cells. Two common approaches to solve the complex problem are stochastic and deterministic methods. Stochastic methods such as Monte Carlo Markov Chains (e.g. used in RADMC-3D; Dullemond et al. 2012) use random sampling of the physical space or model space to improve the individual population levels, which is particularly useful for highly complex or multi-dimensional problems. Deterministic methods such as ray tracing (e.g. used in LOC; Juvela 2020) compute the local contribution of the radiation field at each cell by solving the equations directly. With this, radiative transfer models iteratively solve the SE equations, considering collisional processes and background radiation. This allows to interpret the observed data, constrain physical properties and make predictions used for future observations.

In the context of this thesis, I focus on the modelling of spectral-line observations. Spectral lines are sensitive to local physical conditions such as the velocity profile of a cloud or core. Therefore, synthetic observations are a useful tool to understand the relations between the physical conditions in a source and the properties of the observed radiation. In order to model and interpret observed molecular emission, the physical and chemical structures of the emitting source have to be taken into account. Full three-dimensional models, however, are very computationally expensive and do not exist for most objects. Therefore, cores or environments are usually approximated by zero- or one-dimensional models. In a one-dimensional model assuming spherical symmetry, the physical structure of a core is characterised by volume density, kinetic temperature and infall velocity. Prestellar cores,

for example, are commonly described by a Bonnor-Ebert sphere, where the gas is in hydrostatic equilibrium (see e.g. Keto et al. 2015). The chemical structure of a core is described by the fractional abundances of the present molecules. To derive the abundance profile for a specific molecule, chemical models are used, where the chemistry is portrayed by large networks containing thousands of reactions both in the gas phase and on the surfaces of dust grains (e.g. Sipilä et al. 2015). These models divide the physical model of a core in concentric shells and solve the chemical evolution separately in each shell. The molecular abundance profiles are extracted at different time steps of the chemical simulation. In the radiative transfer modelling, the best-fit time step is then determined by comparing the simulated emission to an observed spectrum. This gives the opportunity to both estimate physical parameters of the observations and constrain molecular abundances and formation pathways in the chemical models. Unfortunately, radiative transfer calculations are not only limited by computational complexity and expensiveness, but also by the availability of the molecular data: spectroscopic information such as frequencies or Einstein coefficients, and collisional rate coefficients. In general, the spectroscopic data are available in large public databases (e.g. on CDMS, Müller et al. 2001), while collisional coefficients often do not exist (yet). This highlights the interdisciplinary nature of astrochemistry that requires collaboration between astronomical observations, laboratory measurements, and theoretical models.

1.3.3 Radioastronomy

Rotational transitions of molecules lie in the centimeter, millimeter and submillimeter wavelength regime and are therefore observed with radio telescopes. Usually, observations with ground-based telescopes are affected by the absorption and scattering of radiation in the Earth’s atmosphere. Luckily, in the wavelength range 1 mm - 30 m⁶, the atmosphere has windows of much lower opacity, allowing for large radio observation facilities on the ground (Wilson 1999). However, in the millimeter-wave regime, the atmosphere gets increasingly opaque because of water vapour and oxygen absorption features. Therefore, radio telescopes are built at locations with high altitude and dry weather conditions, such as the Atacama Desert in Chile.

The observations used within this thesis were obtained with two single-dish radio telescopes, the IRAM 30 m radio telescope on Pico Veleta, Spain, and the Onsala 20 m radio telescope in Onsala, Sweden. Single-dish telescopes typically consist of an antenna and a receiver. The signal from the sky is collected by a parabolic reflector, focused onto a secondary reflector, and then fed into a receiver. The receiver converts the detected electromagnetic waves into guided waves, and then amplifies and digitises the signal. Optionally, different polarisations are separated, and high frequencies are down-sampled to mid or lower frequency ranges to enable the electronics to deal with the signal. Following the receiver, a suitable backend then processes the signal and extracts the desired scientific information, for example via a spectral or polarisation analysis.

⁶Corresponding to frequencies between 10 MHz and 300 GHz.

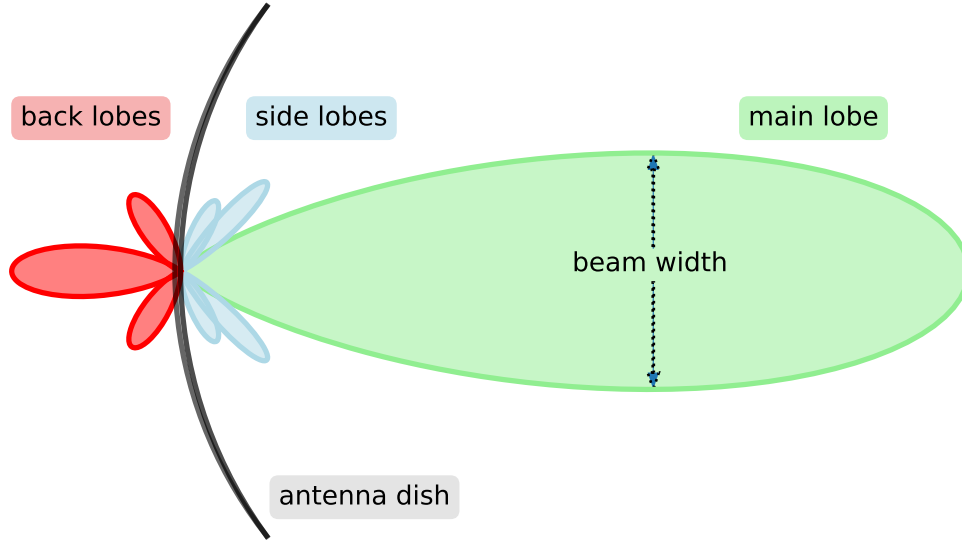


Figure 1.4: Sketch of an antenna beam pattern, showing the main lobe (green), the side lobes (blue), and the back lobes (red), and indicating the half power beam width. The black curve shows the antenna dish.

The beam pattern of a telescope describes the power received by the antenna as a function of direction, with respect to the position the antenna is pointing at (see e.g. Wilson 1999). Figure 1.4 illustrates a typical antenna beam pattern: a sharp maximum in the forward direction, the so-called *main lobe* (also known as main beam), and several smaller, secondary maxima in other directions, the so-called *side lobes*. The main beam contains the highest sensitivity, and its angular extent is defined as the angle between its half-power points (*half power beam width*, HPBW). The HPBW defines the angular resolution of the telescope (θ) at a given diameter of the main reflector (D) and wavelength (λ):

$$\theta \approx 1.2 \frac{\lambda}{D} , \quad (1.43)$$

which is analogous to the relation for optical telescopes. The exact proportionality factor depends on the shape of an antenna and is therefore specific for each telescope⁷. The angle θ determines the minimum angular distance on the projected sky where two objects can be distinguished as separate entities. Thus, compared to optical telescopes, antennas detecting millimeter wavelengths require a collecting dish with a larger diameter to achieve high angular resolution.

The total noise level received by an antenna is described by the *system temperature* (e.g. Wilson 1999):

$$T_{\text{sys}} = T_{\text{A}} + T_{\text{Rx}} . \quad (1.44)$$

⁷For the IRAM 30 m telescope, the HPBW is given by $1.19 \lambda/D$ (IRAM 2025).

The *antenna temperature* T_A contains the signal of the observed astronomical source, and is defined as the temperature of a blackbody or resistor that would lead to the equivalent power received by the antenna. The *receiver temperature* T_{Rx} considers the internal noise from the receiver components, ground radiation, and atmospheric emission. To recover the small contribution of the astronomic signal, different observing techniques can be used to determine the background noise level. Two commonly used techniques to determine the atmospheric contribution are frequency switching, where the operating frequency of the receiver is changed to a range without signal, and position switching, where the telescope is moved to a nearby position in the sky to record an off-source signal.

In the next step, the antenna temperature is corrected for atmospheric losses to obtain the real signal from the source (e.g. Wilson 1999):

$$T_A^* = T_A e^{-\tau_{\text{atm}}} , \quad (1.45)$$

where τ_{atm} is the opacity of the atmosphere along the line of sight, and depends on the elevation of the observed source with respect to the sky horizon.

And finally, the beam pattern of the observing telescope has to be taken into account. For the data analysis, only the signal collected by the main lobe (which is pointing to the astronomical source) is relevant. Therefore, the detected antenna temperature is corrected by the ratio of the *forward efficiency* F_{eff} to the *main beam efficiency* B_{eff} , which leads to the *main beam temperature* (e.g. Wilson 1999):

$$T_{\text{MB}} = \frac{F_{\text{eff}}}{B_{\text{eff}}} T_A^* . \quad (1.46)$$

The main beam efficiency is defined as the fraction of power observed by the main beam of the telescope relative to the total collected power (see Fig. 1.4). A higher B_{eff} indicates that most of the telescope's sensitivity is concentrated in the main beam, and less signal is detected by the side lobes. The forward efficiency, on the other hand, gives the ratio of the signal detected by the front of the antenna to the total detected power. Basically, the main-beam temperature is the beam-averaged brightness temperature of the source, or, more precisely, the convolution of the true brightness distribution with the antenna beam pattern.

In this thesis, I use single-pointing spectra (see Chapter 2) and On-The-Fly (OTF) maps (see Chapters 3 and 4) of molecular emission lines. A single-pointing spectrum contains all scans recorded towards one specific location in the sky. In contrast to that, the OTF mapping technique is an imaging mode of single-dish radio telescopes that enables them to obtain maps of small areas of the sky (see e.g. Mangum et al. 2007). The maps consist of a three-dimensional data cube, containing two positional axes and one spectral axis. During OTF mode, the telescope continuously records data while it moves across the desired region of the sky, which is much more efficient and accurate than by moving step by step and integrating at each position separately. The field of interest is covered rapidly and smoothly, with only a minimum change in atmosphere and system. To create the final map, the data are placed on the coordinates they were observed, followed by the

convolution with a Gaussian kernel to simulate the shape of the telescope beam, and a re-sampling onto an image grid with a user-defined pixel size. The telescope acts as a low pass filter, where spatial information smaller than the beam size is lost. Therefore, the beam size is the lowest spatial scale resolvable with an OTF map. To avoid both over- and undersampling, the data is sampled with roughly twice the Nyquist (spatial) frequency, which is about 4-5 points per HPBW (Wilson 1999).

1.4 Machine learning in astrophysics

“Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed” – Arthur Samuel, 1959

Machine learning is a branch of artificial intelligence (AI) that uses algorithms and statistical models to identify underlying patterns in data and automate decision-making to predict future events and scenarios. The modern-time machine learning revolution started roughly 10 years ago with the public releases of **Tensorflow** (Abadi et al. 2016) and **PyTorch** (Paszke et al. 2019). These Python interfaces made machine learning tools easily accessible to a broad audience. Today, machine learning is a widespread technology that finds usage in web search, social media, healthcare, facial recognition, chatbots, scientific research, and many more applications. The public release of generative artificial intelligence models, such as ChatGPT in 2022, has significantly accelerated the growth of machine learning and AI, revolutionising numerous industries and transforming the way humans interact with technology.

With the opportunity for acquiring large amounts of data via telescope observations and the need for processing those complex, enormous datasets, astronomy is predestined for the usage of machine learning techniques. Following the boom of AI, also the possibilities for applications of machine learning techniques in the various subfields of astrophysics have increased. The applied frameworks and pipelines are becoming increasingly complex, enabling scientific experiments and discoveries.

The two main types of machine learning mechanisms are supervised and unsupervised learning. Supervised learning techniques perform a mapping between an input space and a known truth. More precisely, they use a training dataset to develop a prediction model by analysing the underlying relationship between input data and corresponding output values. The derived model can then make predictions for the output when a new dataset is given as input. The performance of the model hereby highly depends on the size and the variance of the used training set. Typically, supervised learning is used for classification, where the category an object belongs to is predicted, and regression problems, where an attribute associated with an object is predicted.

Unsupervised learning techniques group data based on similar attributes or characteristics without any information specified a priori. Unsupervised learning mechanisms are sorted into two different branches, namely dimensionality reduction and clustering. The process of applying an unsupervised machine learning to a specific dataset includes the

following steps (following Fotopoulou 2024):

1. Data gathering and calibration. In an astrophysical context that can contain telescope observations or data from numerical models or simulations.
2. Pre-processing. This can include for example data cleaning and standardisation. Datasets are commonly standardised to improve the behaviour of features in the learning algorithms, as many of them treat larger numerical numbers as more important. Effective pre-processing ensures that the model learns from the true underlying patterns in the data rather than from artifacts caused by inconsistent scaling or noise.
3. Dimensionality reduction. If multiple features of a dataset are correlated, the data can be mapped from a high-dimensional space to a lower-dimensional space, the so-called latent space, that preserves the majority of the information in the data. A prominent example of this is the principal component analysis (PCA, Hotelling 1936), which maximises the variance of the data. By selecting a subset of these components, a PCA effectively reduces dimensionality while preserving the most informative aspects of the original dataset.
4. Hyperparameter tuning. Hyperparameters are parameters of the learning algorithm itself, independent of the dataset. A common approach to optimise these values is a grid search, which systematically explores a predefined set of hyperparameter combinations to identify the configuration that yields the best performance based on cross-validation scores.
5. Accuracy estimation. This is done to evaluate the performance of the algorithm and the accuracy of its predictions. Common practice is for example cross-validation, where a dataset gets divided into a training set that trains the machine learning model and a test set that evaluates the found solution.

In this work, I focus on unsupervised machine learning techniques, specifically density-based clustering. Clustering is a process of knowledge discovery, where data is organised into a set of categories by learning their underlying patterns. This helps to uncover hidden structures or attributes in complex datasets. There exist multiple clustering techniques that provide different approaches to solve the problem – for example centroid-based clustering, hierarchical clustering, tree-based clustering, density-based clustering, or probabilistic clustering (see e.g. Awad & Khanna 2015 for an overview). The most common and highly-cited clustering technique is the density-based spatial clustering of applications with noise (DBSCAN; Ester et al. 1996). This algorithm is based on the concept that clusters have a typical density of points that is considerably higher than the density outside of the clusters. In the mechanism, points that are closely packed (= within a predefined distance) are grouped together and form a cluster. The number of clusters found, however, is not predefined by the user but determined by the data itself. Additionally, the cluster shapes can be arbitrary and are not fixed as in other clustering algorithms. Another big advantage DBSCAN is that outliers located in low-density regions are marked as noise points instead

of being assigned to a cluster, which allows for a more realistic clustering of noisy data. In the context of this thesis, I use density-based clustering to find patterns in molecular emission and to identify correlations between physical and chemical parameters of dense cores and different molecular species.

1.5 This thesis

After introducing the theoretical background of this thesis, the next chapters will present the results and analysis of three projects combining astronomical observations, theoretical modelling, and machine learning.

In the following, I give a short summary of each chapter:

- In Chapter 2, I compare the level of deuteration of four simple molecules in the envelopes of the prestellar core L1544 and the protostellar core HH211. I use single-pointing spectra of CCH, HCN, HNC, HCO^+ , and their ^{13}C -, ^{18}O -, and D-bearing isotopologues, observed with the Onsala 20 m single-dish radio telescope. With a combination of non-LTE radiative transfer modelling and abundances from chemical simulations, I reproduce the observed molecular lines. I derive the molecular column densities and deuteration levels by using both ^{13}C - and ^{18}O -bearing isotopologues of the main species. I find similar levels of deuteration in L1544 and HH211, showing that the deuterium fractionation is most likely equally efficient towards both cores. I discuss the similarities and differences among the deuterium fractionation of the individual molecules.
- In Chapter 3, I analyse the chemical segregation of $\text{c-C}_3\text{H}_2$, CH_3OH , and CH_3CCH in a sample of starless and prestellar cores with unsupervised machine learning. I use on-the-fly maps observed with the IRAM 30 m single-dish radio telescope. I apply the density-based clustering algorithms DBSCAN and HDBSCAN to identify chemical and physical structures within these cores. With the clustering analysis, I find significant chemical differentiation across the cores, reproducing the known segregation between $\text{c-C}_3\text{H}_2$ and CH_3OH and identifying a non-apparent differentiation between $\text{c-C}_3\text{H}_2$ and CH_3CCH . I discuss the insight into the chemical and physical structures provided by density-based clustering and the benefits compared to other statistical methods of data analysis.
- In Chapter 4, I study the deuteration of the two carbon chains HC_3N and CH_3CCH in the prestellar core L1544, using emission maps observed with the IRAM 30 m radio telescope. With non-LTE radiative transfer modelling, I constrain the physical parameters of the observed species to determine the molecular column densities. I derive the deuteration maps of DC_3N , CH_3CCD , and CH_2DCCH , finding that the molecules trace different physical conditions. I discuss the differences in the deuterium fractionation of each isotopologue and the effect of the environment of the core on it.

- In Chapter 5, I give a brief summary of this work, and discuss future perspectives of the topics in this thesis.

Chapter 2

Similar levels of deuteration in the prestellar core L1544 and the protostellar core HH211

The content of this chapter was published in the *Astronomy & Astrophysics* journal.
Credit: Giers, K., Spezzano, S., Caselli, P., Sipilä, O., Pineda, J. E., Redaelli, E., Bop, C. T., and Lique, F., 2023, A&A, 676, A78.

Abstract

Context. In the centre of pre-stellar cores, deuterium fractionation is enhanced due to low temperatures and high densities. Therefore, the chemistry of deuterated molecules can be used to probe the evolution and the kinematics in the earliest stages of star formation.

Aims. We analyse the deuterium fractionation of simple molecules, comparing the level of deuteration in the envelopes of the prototypical pre-stellar core L1544 in Taurus and the young protostellar core HH211 in Perseus.

Methods. We used single-dish observations of CCH, HCN, HNC, and HCO^+ and their ^{13}C -, ^{18}O -, and D-bearing isotopologues, detected with the 20 m telescope at the Onsala Space Observatory. We derived the column densities, and subsequently the carbon isotopic ratios and deuterium fractions of the molecules. Additionally, we used radiative transfer simulations and results from chemical modelling to reproduce the observed molecular lines. We used new collisional rate coefficients for HNC, HN^{13}C , DNC, and DCN that consider the hyperfine structure of these molecules.

Results. For CCH, we find high levels of deuteration (10%) in both sources, consistent with other carbon chains. We find moderate deuteration of HCN (5-7%), with a slight enhancement towards the protostellar core. Equal levels of deuteration for HNC towards both cores ($\sim 8\%$) indicate that HNC is tracing slightly different layers compared to HCN. We find that the deuterium fraction of HCO^+ is enhanced towards HH211, most likely caused by isotope-selective photodissociation of C^{18}O . With radiative transfer, we were

able to reproduce the observed lines of CCH, HCN, H^{13}CN , HNC, HN^{13}C , and DNC towards L1544 as well as CCH, H^{13}CN , HN^{13}C , DNC, H^{13}CO^+ , HC^{18}O^+ , and DCO^+ towards HH211.

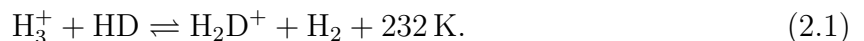
Conclusions. Similar levels of deuteration show that the deuterium fractionation is most probably equally efficient towards both cores, suggesting that the protostellar envelope still retains the chemical composition of the original pre-stellar core. The fact that the two cores are embedded in different molecular clouds also suggests that environmental conditions do not have a significant effect on the deuterium fractionation within dense cores. Our results highlight the uncertainties when dealing with ^{13}C isotopologues and the influence of the applied carbon isotopic ratio. Radiative transfer modelling shows that it is crucial to include the effects of the hyperfine structure to reproduce the observed line shapes. In addition, to correctly model emission lines from pre-stellar cores, it is necessary to include the outer layers of the core to consider the effects of extended structures. In addition to HCO^+ observations, HCN observations towards L1544 also require the presence of an outer diffuse layer where the molecules are relatively abundant.

Keywords: astrochemistry – ISM: clouds – ISM: molecules – ISM: abundances – stars: formation – radiative transfer

2.1 Introduction

Deuterated molecules are important diagnostic tools of the earliest phases of star formation, allowing astronomers to study the central regions of pre-stellar cores in detail (e.g. Caselli et al. 2002b; Pineda et al. 2022), where CO and other CO-bearing molecules are heavily frozen onto dust grains (e.g. Caselli et al. 1999; Bergin & Tafalla 2007). High levels of deuteration are also found in the cold envelopes of newly formed protostars (van Dishoeck et al. 1995; Hatchell et al. 1998; Roberts et al. 2002; Parise et al. 2004; Emprechtinger et al. 2009), bringing up the question of how much of this material can survive during the star and planet formation process. High levels of deuteration have been found in planet-forming disks (e.g. Mathews et al. 2013), comets (e.g. Altwegg et al. 2015), and carbonaceous chondrites (e.g. Robert 2003; Busemann et al. 2006; Ceccarelli et al. 2014). Our oceans are also enriched in heavy water, partially inherited from the solar pre-stellar phase (Cleeves et al. 2014). It is therefore important to understand the chemical processes that regulate deuterium fractionation in the early phases of star formation in detail and investigate possible differences between various evolutionary stages. In particular, observational constraints are needed for the chemical models to discriminate among the processes of deuteration happening in the gas phase and on the surface of dust grains.

At low temperatures, deuterium fractionation is driven by the gas phase reaction



Due to its exothermicity, the reaction proceeds more efficiently from left to right with decreasing temperature (assuming a low ortho-to-para H_2 ratio; e.g. Pagani et al. 1992).

This leads to an enhancement of the abundance of H_2D^+ and an associated increase in the abundances of D_2H^+ and D_3^+ , and subsequently to efficient deuteration of other molecules. In the cold and dense conditions of the earliest stages of star formation, the main destructor of H_2D^+ , CO, is highly depleted from the gas phase and frozen out onto the surfaces of dust grains (e.g. Caselli et al. 1999), and this further increases the deuterium fractionation (Dalgarno & Lepp 1984). The deuterium fraction of a molecular tracer is measured by dividing the column density of the deuterated molecule by the column density of the hydrogenated molecule.

In this paper, we focus on deuterated molecules present in the well-known pre-stellar core L1544 in Taurus, and in one of the youngest (and most highly deuterated) Class 0 sources, HH211 in Perseus. L1544 is located at a distance of 170 pc (Galli et al. 2019), with clear evidence of gravitational contraction (Caselli et al. 2012). Its central temperature approaches 6 K (Crapsi et al. 2007) and its central density reaches 10^7 cm^{-3} (Keto & Caselli 2010; Caselli et al. 2019). Towards its centre, L1544 exhibits a high level of deuteration (Crapsi et al. 2005; Redaelli et al. 2019). Within the central 2000 au, recent Atacama Large Millimeter/submillimeter Array observations have been found consistent with almost complete (99.999%) depletions of elements heavier than He (Caselli et al. 2022). L1544 is on the verge of star formation, thus shedding light on the initial conditions in the process of star formation. HH211 is a newly born Class 0 protostar (e.g. Enoch et al. 2006; Lee et al. 2018) at a distance of 321 pc (Ortiz-León et al. 2018), which presents high degrees of deuterium fractionation (Emprechtinger et al. 2009). It hosts a jet-driven molecular outflow (McCaughrean et al. 1994; Gueth & Guilloteau 1999), and has a Keplerian disk (Segura-Cox et al. 2016). The surrounding envelope is elongated in a direction roughly perpendicular to the jet and outflow axis (Gueth & Guilloteau 1999). In Chantzos et al. (2018), it is shown that the deuteration of cyclopropenylidene, $\text{c-C}_3\text{H}_2$, is more efficient towards HH211 ($\text{N}(\text{c-C}_3\text{HD})/\text{N}(\text{c-C}_3\text{H}_2)=20\%$) than in L1544 ($\text{N}(\text{c-C}_3\text{HD})/\text{N}(\text{c-C}_3\text{H}_2)=10\%$). Gas phase deuteration processes are sufficient to reproduce the $\text{N}(\text{c-C}_3\text{HD})/\text{N}(\text{c-C}_3\text{H}_2)$ ratio observed in L1544, so the enhancement of deuteration in HH211 is assumed to be due to efficient deuteration happening on the surface of dust grains in the pre-stellar phase, followed by the release from the grains in the protostellar stage.

We present a survey of ground state-rotational lines of deuterated molecules towards L1544 and HH211, observed with the single-dish 20 m radio telescope at the Onsala Space Observatory. Our observations cover CCH, HCN, HNC, and HCO^+ , and their deuterated isotopologues. To avoid optical depth limitations with the main isotopologues, we also observed ^{13}CCH , C^{13}CH , H^{13}CN , HN^{13}C , H^{13}CO^+ , and HC^{18}O^+ . This data set allows for a unique comparison between the deuteration in a very dynamically evolved pre-stellar core on the verge of forming a low-mass star, and a young Class 0 protostellar core, and thus, to study the influence of the evolutionary stage on the deuterium fractionation. Such a comparison is important to understand how the deuterated molecules are inherited in the earlier stages of formation of a low-mass star, and eventually in its planetary system.

In Sect. 2.2, we describe the observations, followed by the results in Sect. 2.3. The analysis in Sect. 2.4 covers the derivation of the column densities and deuterium fraction assuming local thermal equilibrium (LTE) conditions. In Sect. 2.5, we use the non-LTE

radiative transfer code LOC to model our observations. We discuss the results of the LTE and non-LTE analysis in Sect. 2.6 and present our conclusions in Sect. 2.7.

2.2 Observations

The data presented in this paper were obtained with the 20 m radio telescope at the Onsala Space Observatory. The observed spectra are centred on the dust emission peaks of the sources (L1544: $\alpha_{2000} = 05^{\text{h}}04^{\text{m}}17^{\text{s}}.21$, $\delta_{2000} = +25^{\circ}10'42''.8$; HH211: $\alpha_{2000} = 03^{\text{h}}43^{\text{m}}56^{\text{s}}.80$, $\delta_{2000} = +32^{\circ}00'50''.0$). The first part of the observations was done in May 2020 (project ID: O2019b-04). A follow-up project observed the second part from January to May 2021 (project ID: O2020b-05). For the observations, we used the 3 mm (Belitsky et al. 2015) and the 4 mm receiver (Walker et al. 2016) in combination with the OSA spectrometer at 19 kHz resolution, with 625 MHz bandwidth and dual polarisation. We applied the dual beam-switched mode, with the off position shifted by 11.8 arcmin in the azimuth direction. The intensity calibration was done using standard Dicke-switched type calibration, switching between a hot load and a sky direction close to the source. The 1σ pointing accuracy is estimated to be 3 arcsec. The frequency resolution of the data corresponds to a velocity resolution of about 0.07 km s^{-1} at 84 GHz. The weather conditions during the observations were variable, with the system temperature ranging between 120 – 300 K.

The observed transitions are summarised in Table A.1. The corresponding spectra towards L1544 and HH211 are presented in Fig. A.1 and Fig. A.2, respectively. A selection of the observed transitions is given in Table 2.1 and Fig. 2.1. The average beam size of the observations is about 50 arcsec, which corresponds to a physical size of 8500 au for L1544 and 16000 au for HH211.

The data processing was done using the GILDAS software (Pety 2005) and the Python package PYSPECKIT (Ginsburg & Mirocha 2011; Ginsburg et al. 2022). The antenna temperature T_A^* was converted to the main beam temperature T_{mb} using the relation $T_{\text{mb}} = T_A^*/B_{\text{eff}}$. The corresponding values for the main beam efficiencies (B_{eff}) of the 20 m telescope are given in Table 2.1.

2.3 Results

Many species observed in this survey show either resolved (CCH, ^{13}CCH , C^{13}CH , CCD, HCN, H^{13}CN , DCN), or unresolved (HNC, HN^{13}C , DNC, H^{13}CO^+ , DCO^+) hyperfine structure. Due to the amount of hyperfine lines in the resolved case, we decided to select one hyperfine component per species for the LTE analysis, based on optical depth and signal-to-noise ratio (S/N). The properties of the selected lines are given in Table 2.1, determined by fitting a Gaussian profile to the lines (see Fig. A.1 and Fig. A.2). The corresponding spectra are shown in Fig. 2.1, where lines observed towards L1544 and HH211 are plotted in black and red, respectively. In the following, we describe the observed spectra and

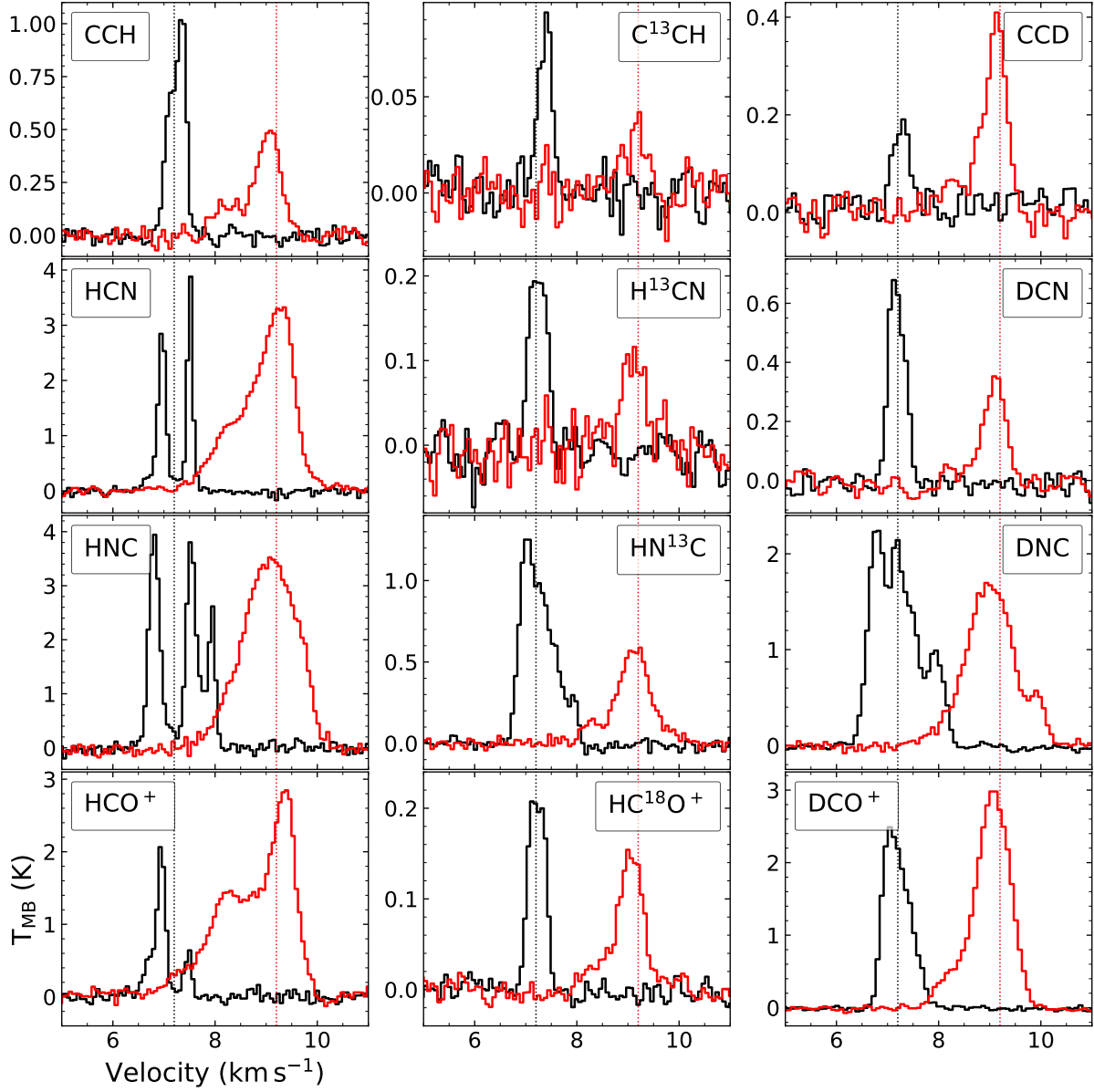


Figure 2.1: Selection of the spectra observed towards the pre-stellar core L1544 (black) and the protostellar core HH211 (red). The dotted vertical lines indicate the respective rest velocities of the cores. For HCN, the $F = 2 - 1$ transition is shown. For CCH, $C^{13}\text{CH}$, CCD, $H^{13}\text{CN}$, and DCN, the transitions listed in Table 2.1 are shown.

Table 2.1: Properties of the selected observed lines. Numbers in brackets give the uncertainty on the last digit.

Species	Transition	Frequency ^a (MHz)	B_{eff} (K)	$T_{\text{mb,peak}}$ (mK)	rms (mK)	V_{LSR}^b (km s ⁻¹)	Δv^b (km s ⁻¹)	$\int T_{\text{mb}} dv$ (K km s ⁻¹)
L1544								
CCH	$N=1-0, J=1/2-1/2, F=1-0$	87446.47(1)	0.52	0.96(3)	27	7.24	0.39	0.40(2)
¹³ CCH	$N=1-0, J=3/2-1/2, F_1=2-1, F=5/2-3/2$	84119.33(2)	0.53	0.039(5)	7	7.31	0.32	0.013(2)
C ¹³ CH	$N=1-0, J=3/2-1/2, F_1=2-1, F=5/2-3/2$	85229.335(4)	0.53	0.089(6)	9	7.35	0.32	0.030(3)
CCD	$N=1-0, J=1-0, F_1=3/2-1/2, F=3/2-3/2$	72101.811(5)	0.54	0.19(1)	24	7.23	0.39	0.078(9)
H ¹³ CN	$J=1-0, F=0-1$	86342.2543(3)	0.53	0.22(1)	18	7.21	0.39	0.091(7)
DCN	$J=1-0, F=0-1$	72417.03(1)	0.54	0.69(2)	23	7.14	0.37	0.27(1)
H ¹³ CO ⁺	$J=1-0$	86754.288(5)	0.53	1.38(4)	20	7.13	0.52	0.77(4)
HC ¹⁸ O ⁺	$J=1-0$	85162.223(5)	0.53	0.229(6)	9	7.18	0.38	0.093(4)
DCO ⁺	$J=1-0$	72039.3124(8)	0.54	2.44(5)	25	7.12	0.60	1.55(4)
HN ¹³ C	$J=1-0, F_1=0-1, F_2=1-2, F=1-2$	87090.675(3)	0.52	0.27(3)	25	7.92	0.24	0.07(2)
DNC	$J, I, F=1, 2, 1-0, 2, 2; 1, 2, 1-0, 0, 0$	76305.513(1)	0.54	0.99(2)	26	7.90	0.40	0.42(2)
HH211								
CCH	$N=1-0, J=3/2-1/2, F=1-1$	87284.11(1)	0.52	0.50(1)	27	9.02	0.54	0.29(1)
¹³ CCH	$N=1-0, J=3/2-1/2, F_1=1-0, F=3/2-1/2$	84153.31(2)	0.53	0.016(4)	6	9.07	0.4	0.007(2)
C ¹³ CH	$N=1-0, J=3/2-1/2, F_1=2-1, F=5/2-3/2$	85229.335(4)	0.53	0.032(4)	8	9.10	0.53	0.018(4)
CCD	$N=1-0, J=1-0, F_1=3/2-1/2, F=5/2-3/2$	72107.721(3)	0.54	0.37(5)	28	9.10	0.45	0.18(3)
H ¹³ CN	$J=1-0, F=2-1$	86340.1666(1)	0.53	0.45(2)	23	9.05	0.61	0.29(2)
DCN	$J=1-0, F=0-1$	72417.03(1)	0.54	0.35(2)	22	9.06	0.50	0.18(4)
H ¹³ CO ⁺	$J=1-0$	86754.288(5)	0.53	1.95(1)	24	8.97	0.60	1.24(2)
HC ¹⁸ O ⁺	$J=1-0$	85162.223(5)	0.53	0.153(4)	8	9.03	0.53	0.087(5)
DCO ⁺	$J=1-0$	72039.3124(8)	0.54	3.01(1)	22	9.04	0.74	2.38(2)
HN ¹³ C	$J=1-0$	87090.825(4)	0.52	0.58(1)	27	9.09	0.71	0.46(1)
HN ¹³ C	$J=1-0, F_1=0-1, F_2=1-2, F=1-2$	87090.675(3)	0.52	0.090(5)	27	9.78	0.35	0.034(7)
DNC	$J, I, F=1, 2, 1-0, 2, 2; 1, 2, 1-0, 0, 0$	76305.513(1)	0.54	0.56(3)	31	9.90	0.38	0.22(4)

Notes. Numbers in brackets give the uncertainty on the last digit. ^(a) Extracted from Cologne Database for Molecular

Spectroscopy (Müller et al. 2001). ^(b) The uncertainties of the fits are smaller than the velocity resolution. Therefore, the error on

V_{LSR} and Δv are given by the observed spectral resolution, 0.07 km s⁻¹.

discuss the line selection.

2.3.1 L1544

The radical CCH and its isotopologues present lines with multiple hyperfine (hf) components. In L1544, we detected all six hf components of CCH, where four out of six likely show a dip due to self-absorption. This is caused by absorption in the outer layers of the core and dilutes the signal from the high-density central regions. We selected the line with the lowest optical depth ($N=1-0$, $J=1/2-1/2$, $F=1-0$), to minimise any issues with optical depth effects. In L1544, we successfully observed the components of the ^{13}C isotopologues of CCH: For ^{13}CCH , we detected three of the eight hyperfine components with a signal-to-noise ratio > 3 ; C^{13}CH shows five of the seven components with $\text{S/N} > 4$. For both, we selected the transition with the highest S/N , ($N=1-0$, $J=3/2-1/2$, $F_1=2-1$, $F=5/2-3/2$). The deuterated isotopologue, CCD, is much brighter, and seven of the nine hyperfine components are detected ($\text{S/N} > 8$). Some of them show self-absorption, thus, we selected the line with the lowest S/N (≈ 8), $N=1-0$, $J=3/2-1/2$, $F=5/2-3/2$, to avoid optical depth issues. The Gaussian fit parameters of the ^{13}CCH and C^{13}CH hf components show that the lines are shifted by approximately 0.1 km s^{-1} ($V_{\text{LSR}} = 7.3 \text{ km s}^{-1}$) compared to CCH and CCD ($7.2-7.25 \text{ km s}^{-1}$), which are located at the typical system velocity observed for L1544 (7.2 km s^{-1}). However, this shift is close to the spectral resolution of the data (0.07 km s^{-1}) and might therefore not be significant.

HCN is optically thick, showing a double-peaked line profile in each hyperfine component with a dip reaching down to almost zero level, caused by self-absorption in the outer layers of the core. Due to its high optical depth, the line profile provides information on the kinematics in the outer layers of the core. A blue asymmetry in a self-absorbed peak is a typical signature of infall motions - extended inward velocities in the outer regions of a core cause enhanced self-absorption at redshifted velocities (e.g. Evans 1999). The intensity ratio between the two peaks can give information on the infall rates itself (e.g. Keto et al. 2015), because with increasing infall motions, the redder peak decreases in intensity and evolves to a shoulder of the blue peak (like it is seen with HCO^+ in Fig. 2.1 and for example in Redaelli et al. 2022). However, the peak of the central HCN hf component shows a red asymmetry. Following the argumentation above, this is caused by an expansion motion of the core in its outer regions, more specifically, in the layers traced by the HCN transition.

It has to be noted that the red asymmetry of HCN has a different origin than the spectra of HC_3N ($1-0$) towards L1544 observed by Bianchi et al. (2023). In the case of the optically thin HC_3N , the red asymmetry is caused by the molecule tracing the southern region of the core, which is redshifted with respect to the rest velocity of L1544 (see the V_{LSR} map in Fig. B.1 in Spezzano et al. 2016). By overplotting the central hyperfines of HCN and HC_3N , Fig. 2.2 displays the large line width of HCN. This additionally shows the optical thickness of the transition, as in L1544 the line width observed for optically thin species is approximately 0.4 km s^{-1} (e.g. Caselli et al. 2002a). The red asymmetry is not observed for the other two hyperfine components of HCN. They show a rather symmetric double-peak profile, possibly tracing a more static or non-moving layer, which was also

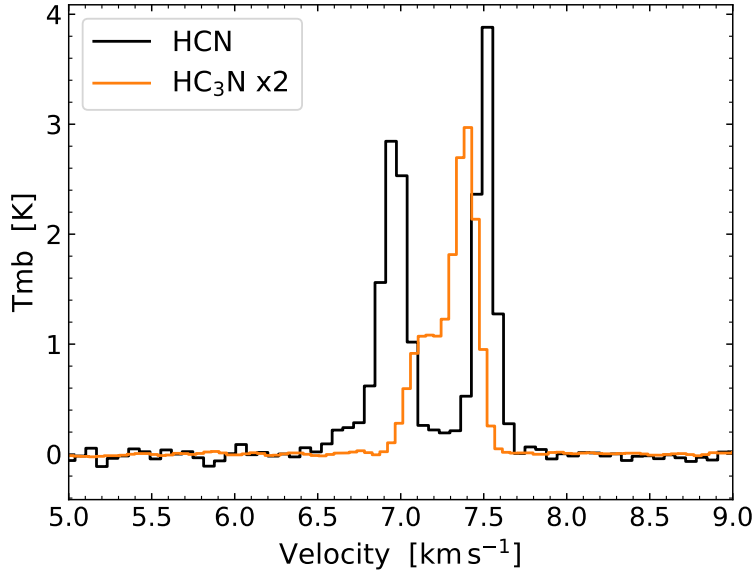


Figure 2.2: Central hyperfine ($F = 2 - 1$) of the $(1 - 0)$ transitions of HCN (black) and HC_3N (orange), observed towards L1544. The data for HC_3N are taken from Bianchi et al. (2023)

observed in CS (2-1) by Tafalla et al. (1998). This shows that the hyperfine components clearly trace different layers, depending on their optical depth. In Sect. 2.5, we analyse this behaviour by applying non-LTE radiative transfer modelling. Due to its optical thickness, the main species cannot be used to derive the column density and deuterium fraction of HCN, so we focus instead on the ^{13}C isotopologue.

The hyperfine structures of both the ^{13}C - and the deuterated isotopologue of HCN show some self-absorption. A hyperfine structure fitting, using the HFS method provided by the GILDAS package CLASS, reveals that in both cases the weakest component ($F=0-1$) is moderately optically thin, with an optical depth of $\tau = 0.48(9)$ and $\tau = 0.9(2)$ for H^{13}CN and DCN , respectively (numbers in brackets give the uncertainty on the last digit). Therefore, we selected these lines for the further analysis.

Our observations cannot resolve the hyperfine structure of HNC, but only show blended components. In addition, the lines are expected to be heavily self-absorbed like in the case of HCN, which makes it impossible to reconstruct the components. Therefore, we included this molecule in the non-LTE analysis in Sect. 2.5 but do not consider it in the derivation of column densities. The lines of the isotopologues, HN^{13}C and DNC , are split into multiple hyperfine components, mainly due to the non-zero nuclear spins of the D and N nuclei. For HN^{13}C , van der Tak et al. (2009) identified four effective hyperfine components, with relative intensities of 1.00, 4.04, 6.63, and 3.66, from low to high frequency. In Fig. 2.3 we fitted the four effective components of HN^{13}C to our observed spectrum by applying a multi-component Gaussian. The amplitudes of the fits result in relative intensities of

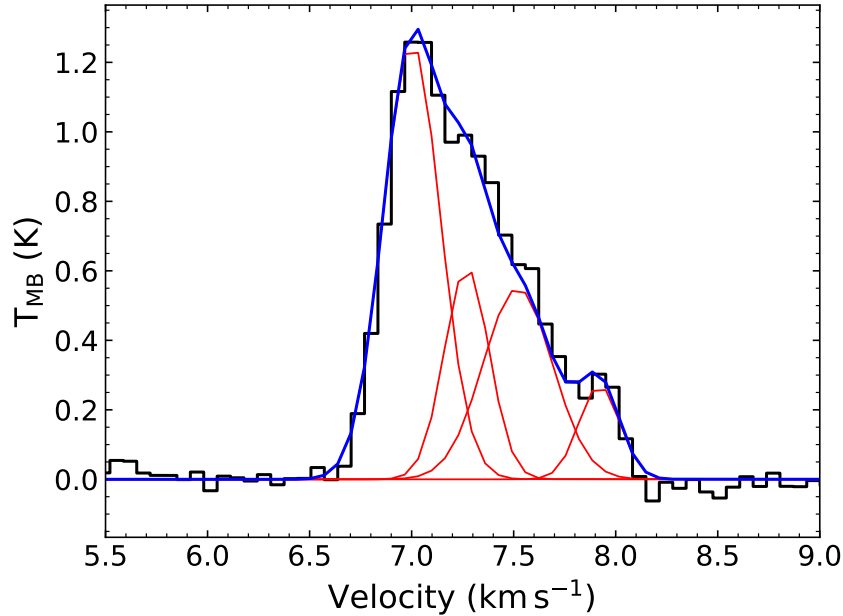


Figure 2.3: Multi-component Gaussian fit of the four effective hyperfine components found by van der Tak et al. (2009) to the spectrum of HN^{13}C observed towards L1544.

1.00, 2.18, 2.41, 4.95 (from high to low velocity components), corresponding to optical depths of 0.3, 0.7, 0.8, 4.5, respectively. Therefore, we conclude that all components of HN^{13}C , except for the weakest component at the reddest velocity, are optically thick and most probably self-absorbed. Hence, we only used the hf component at 87090.675 MHz to derive the column density of HN^{13}C . For DNC, van der Tak et al. (2009) identified six effective hyperfine components, reporting relative intensities of 0.33, 2.67, 1.67, 2.33, 0.99, 1.00, from low to high frequency. A CLASS HFS fitting gives out optical depths of 0.24, 1.94, 1.21, 1.69, 0.72, 0.72, from low to high frequency. Therefore, we selected the lowest frequency component of the DNC hyperfine structure for our further analysis.

The $J=1-0$ transition of HCO^+ shows strong self-absorption in L1544, with a blue asymmetry caused by infall motions in the outer layers of the core. Therefore, we excluded it in the LTE analysis. A detailed discussion and non-LTE modelling of this line can be found in Redaelli et al. (2022). The ^{13}C isotopologue of HCO^+ also shows asymmetric self-absorption. However, we used the line to derive a lower limit for the integrated intensity and column density of HCO^+ . The spectrum of the ^{18}O isotopologue shows a flattened top. With higher spectral resolution (20 kHz) and a smaller telescope beam ($31''$), a small dip is visible (e.g. Caselli et al. 2002a; Redaelli et al. 2019) that is caused by kinematics and CO depletion in the central regions of the core. Due to the lower spectral resolution and larger telescope beam of our data, however, the emission is diluted. The line does not show any asymmetry, as the abundance of the molecule in the outer envelope is too low to be affected by self-absorption. The deuterated isotopologue consists of three blended hyperfine components that we cannot resolve with our spectral resolution. Therefore, we

treated the line as one component.

2.3.2 HH211

The spectra observed towards HH211 show an additional velocity component centred at around 8.22 km s^{-1} , which is significantly blue-shifted with respect to the main emission peak ($\sim 9.2 \text{ km s}^{-1}$). This emission likely originates from the outskirts of a lower velocity plateau which might be part of or influenced by the outflow. It is located to the south-east of the core (Pineda & Friesen et al. in prep.), which is partly covered by the telescope beam ($40\text{--}60''$, depending on the frequency of the transition) of the Onsala 20 m telescope. To account for the extra emission of the additional velocity component, we apply a two-component Gaussian fit to the spectral lines observed towards HH211, wherever possible. In the analysis, we used the fit parameters of the main component emission to derive the column density and deuterium fraction of the molecular tracers. The only exception are the two ^{13}C isotopologues of CCH, where the signal-to-noise ratio is too low to resolve the second component. In this case, a one-component Gaussian fit is sufficient to obtain the line properties. In Table A.2, we list the best-fit parameters and the derived column densities for the additional velocity component. There is no clear trend visible in the FWHM of the additional component compared to the main component. For CCH, CCD and HCO^+ , the second component appears to be broader than the main component, for the rest it is narrower. However, the line width has a rather large error and in some cases, the two line widths are, within the error bars, the same.

The main emission components of the observed lines all show a Gaussian line shape and do not exhibit any dips or asymmetries caused by self-absorption. This is possibly caused by a combination of infall motions and turbulent motions induced by the protostellar outflow in the less dense regions of the core, which is enlarging the velocity range in the core and thus, decreasing the possibility for self-absorption. In addition, the lines are also tracing outflow material, for example HCO^+ or HNC (e.g. Arce & Sargent 2004). In general, the increased level of turbulence in the core causes a broadening of the line widths (on average 0.2 km s^{-1} larger than in L1544).

For CCH, we detected all six hf components, and selected the transition with the lowest optical depth to calculate the column density ($N = 1-0$, $J = 3/2-1/2$, $F = 1/1$). Towards the protostellar core, the ^{13}C isotopologues of CCH are less bright than towards L1544, and we detected only two hyperfine components with $S/N > 3$ for each of them. We chose the components with the higher S/N , $N=1-0$, $J=3/2-1/2$, $F_1=1-0$, $F=3/2-1/2$ for ^{13}CCH and $N=1-0$, $J=3/2-1/2$, $F_1=2-1$, $F=5/2-3/2$ for C^{13}CH . For the deuterated isotopologue of CCH, we detected seven out of nine hyperfine transitions. Among them, the weaker lines do not show the additional velocity component. We selected the transition with the highest signal-to-noise ratio ($S/N > 13$), which is $N=1-0$, $J=3/2-1/2$, $F=5/2-3/2$. The lines of the main isotopologue, CCH, seem to be shifted to lower velocity ($\sim 9.0 \text{ km s}^{-1}$), while the lines of the rarer isotopologues are located around the rest velocity of the system ($\sim 9.1 \text{ km s}^{-1}$).

For HCN, H^{13}CN and DCN, we applied the hyperfine-fitting method of CLASS to check

the optical depth. As in L1544, all components of HCN are optically thick. We therefore excluded the molecule from the LTE analysis, and only considered it in the radiative transfer modelling in Sect. 2.5. The optical depths of DCN are similar to those towards L1544, only the weakest hyperfine is optically thin, with $\tau = 0.35$. For H^{13}CN , however, all hyperfine components turn out to be optically thin, with $\tau = 0.08 - 0.4$. Therefore, we selected the brightest hyperfine ($F=2 - 1$) for the further analysis.

In the spectra of HNC and isotopologues, the hyperfine components are completely blended. There is no clear evidence of self-absorption as it was seen in L1544. However, the main species is expected to be optically thick, thus, we focus on the ^{13}C isotopologue. Towards HH211, the hf components of HN^{13}C and DNC are less distinct than towards L1544, and the reddest component is less pronounced. Due to this and the additional velocity component at lower velocity ($\sim 8.2 \text{ km s}^{-1}$), we were not able to fit the four effective components of HN^{13}C and the six effective components of DNC reported by van der Tak et al. (2009). Therefore, we applied a three-component Gaussian fit to the observed spectrum to account for the additional velocity component and the high velocity hf component separately from the remaining, optically thick emission.

The observed emission of HCO^+ shows a very pronounced additional velocity component, with almost half of the intensity of the main emission. To avoid any optical depth issues, we did not use HCO^+ for the LTE analysis. The main velocity components of the ^{13}C - and D-bearing isotopologues have an optical depth of 0.9, and therefore have to be interpreted with caution.

2.4 Analysis

2.4.1 Excitation temperature and column density

In this work, all column densities were calculated under the assumption of optically thin emission, using the formula presented in Mangum & Shirley (2015). Furthermore, we applied the approximation of a constant excitation temperature throughout the core (CTex), following Caselli et al. (2002b) and Redaelli et al. (2019):

$$N_{\text{tot}}^{\text{thin}} = \frac{8\pi\nu^3}{c^3} \frac{Q_{\text{rot}}(T_{\text{ex}})}{g_u A_{\text{ul}}} [J_\nu(T_{\text{ex}}) - J_\nu(T_{\text{bg}})]^{-1} \frac{e^{\frac{E_u}{kT_{\text{ex}}}}}{e^{\frac{h\nu}{kT_{\text{ex}}}} - 1} \int T_{\text{mb}} dv, \quad (2.2)$$

where $Q_{\text{rot}}(T_{\text{ex}})$ is the partition function of the molecule at an excitation temperature T_{ex} , g_u and E_u are the degeneracy and energy of the upper level of the transition, respectively, A_{ul} the Einstein coefficient for spontaneous emission, $T_{\text{bg}} = 2.73 \text{ K}$ the temperature of the cosmic microwave background, $J(T)$ the Rayleigh-Jeans equivalent temperature and T_{mb} the main beam temperature. Assuming a Gaussian line profile, the integrated main beam temperature is calculated by $\frac{1}{2}\sqrt{\pi/\ln 2} \cdot \Delta V \cdot T_{\text{mb,peak}}$. The corresponding parameters for each transition are listed in Table 2.1 and Table 2.2. We derived the partition function from the energies and degeneracies of the rotational levels (see Mangum & Shirley 2015), taken from the Cologne Database for Molecular Spectroscopy (CDMS Müller et al. 2001).

However, as some of our transitions are only moderately optically thin, we applied the optical depth correction factor (Goldsmith & Langer 1999) to all of our lines to derive the total column density:

$$N_{\text{tot}} = N_{\text{tot}}^{\text{thin}} \frac{\tau}{1 - e^{-\tau}}. \quad (2.3)$$

The optical depth is derived using:

$$\tau_\nu = -\ln \left[1 - \frac{T_{\text{mb}}}{f [J(T_{\text{ex}}) - J(T_{\text{bg}})]} \right], \quad (2.4)$$

where we assumed a filling factor $f = 1$.

For CCH and isotopologues, we assumed an excitation temperature of 6 K in both sources. This temperature has been applied to similar carbon chain molecules such as $c\text{-C}_3\text{H}_2$ in Giers et al. (2022) and Chantzios et al. (2018), following Crapsi et al. (2005) and Emprechtinger et al. (2009). It is also consistent with Taniguchi et al. (2019) who derived $T_{\text{ex}} \approx 6$ K for CCH in the starless cores L1521B and L134N. To estimate the excitation temperature of H^{13}CN , DCN, HN^{13}C and DNC, we used the results of the hyperfine-fitting done earlier with the HFS method in CLASS. The derived values are listed in Table 2.2 and are consistent with excitation temperatures used by Padovani et al. (2011), Hily-Blant et al. (2013) and Quénard et al. (2017). For HC^{18}O^+ and DCO^+ , we used the excitation temperatures derived by Redaelli et al. (2019) using radiative transfer. For H^{13}CO^+ , we assumed an excitation temperature of 6 K, which is an intermediate value and similar to what we used for CCH and isotopologues. In general, we assumed the same or similar excitation temperatures for both the pre-stellar and the protostellar core, as our observations mainly trace the cold outer envelopes of the two cores, where similar conditions hold. Indeed, the effect of the excitation temperature on the derived column density was found to be moderately small. Changing T_{ex} by one Kelvin introduces an error of about 20 %.

The derived column densities of all species observed towards the two sources are presented in Table 2.3. To analyse the nature of the additional velocity component in HH211, we derived the column density of it separately. The results are presented in the last column of Table A.2. The uncertainties for the column densities were derived by propagating the 1σ errors on the integrated intensity and the optical depth, and adding an additional error of 15% to account for the uncertainties in the flux calibration.

For CCH, the derived column densities are consistent with previous measurements within a factor of two (Sakai et al. 2008; Zhang et al. 2021). The ^{13}C isotopologues of CCH are very weak and difficult to observe. Therefore, they have only been detected towards the sources L1521B, TMC-1, L134N/L183, L483 and L1527 so far (Taniguchi et al. 2019; Sakai et al. 2010; Agúndez et al. 2019; Yoshida et al. 2019).

HCN, HNC and isotopologues have been observed multiple times towards L1544 (e.g. Hirota et al. 2003; Hily-Blant et al. 2010; Quénard et al. 2017; Spezzano et al. 2022a) and HH211 (Imai et al. 2018) in previous studies, reporting results which are consistent with our measurements. Previous measurements towards HH211 by Roberts et al. (2002) reported column densities for H^{13}CN and DCN that are smaller by a factor of two compared to our

Table 2.2: Parameters used in the derivation of the column densities.

Molecule	Frequency	T_{ex} (K)	$Q(T_{\text{ex}})$	E_{u}^a (K)	A^a (s^{-1})	n_{crit} (cm^{-3})	g_{u}^a	τ
L1544								
CCH	87446.47(1)	6	12.88	4.2	2.61e-7	9e+4	3	0.348(8)
^{13}CCH	84119.33(2)	6	26.40	4.08	1.37e-6	5e+5	6	0.012(1)
C^{13}CH	85229.335(4)	6	26.30	4.1	1.42e-6	5e+5	6	0.028(2)
CCD	72101.811(5)	6	22.92	3.5	3.28e-7	1e+5	4	0.058(9)
H^{13}CN	86342.2543(3)	3.3	5.92	4.1	2.22e-5	7e+5	1	0.48(9)
DCN	72417.03(1)	3.7	7.49	3.5	1.32e-5	4e+5	1	0.9(2)
H^{13}CO^+	86754.288(5)	6	3.24	4.2	3.85e-5	2e+5	3	0.55(1)
HC^{18}O^+	85162.223(5)	5.5	3.05	4.1	3.65e-5	1e+5	3	0.086(2)
DCO^+	72039.3124(8)	7.8	4.86	3.5	2.21e-5	1e+5	3	0.657(7)
HN^{13}C	87090.675(3)	4	27.49	4.2	1.58e-5	1e+5	7	0.24(3)
DNC	76305.513(1)	5	27.81	3.7	1.60e-5	2e+5	6	0.57(2)
HH211								
CCH	87284.11(1)	6	12.88	4.2	2.60e-7	9e+4	3	0.165(5)
^{13}CCH	84153.31(2)	6	26.40	4.08	1.37e-6	5e+5	4	0.005(1)
C^{13}CH	85229.335(4)	6	26.30	4.1	1.42e-6	5e+5	6	0.010(1)
CCD	72107.721(3)	6	22.92	3.5	8.60e-7	3e+5	6	0.12(2)
H^{13}CN	86340.1666(1)	4.1	7.05	4.1	2.22e-5	7e+5	5	0.39(2)
DCN	72417.03(1)	3.7	7.49	3.5	1.32e-5	4e+5	1	0.44(7)
H^{13}CO^+	86754.288(5)	6	3.24	4.2	3.85e-5	2e+5	3	0.91(1)
HC^{18}O^+	85162.223(5)	5.5	3.05	4.1	3.65e-5	1e+5	3	0.057(2)
DCO^+	72039.3124(8)	7.8	4.86	3.5	2.21e-5	1e+5	3	0.899(7)
HN^{13}C	87090.675(3)	4	27.49	4.2	1.58e-5	1e+5	7	0.073(4)
DNC	76305.513(1)	5	27.81	3.7	1.60e-5	2e+5	6	0.28(2)

Notes. ^(a) Extracted from the Cologne Database for Molecular Spectroscopy (Müller et al. 2001)

results. These differences are likely due to different approximations used in the derivation of the column densities, as well as a different telescope beam.

By assuming the carbon isotopic ratio for the local interstellar medium, $^{12}\text{C}/^{13}\text{C} = 68$ (Milam et al. 2005), we derived a HNC/HCN ratio of 0.3(1) and 0.4(1) for L1544 and HH211, respectively. Previous measurements of this ratio towards L1544 and other cores have revealed a value revolving around 1 (L1544: 1.0(5), Quénard et al. 2017; L1498: 0.9(1), L1521E: 1.1(1), TMC-2: 1.1(1), Padovani et al. 2011), suggesting that HNC and HCN have similar abundances. However, as this is derived using the ^{13}C isotopologues, it implies that the fractionation of the two molecules is the same. In Colzi et al. (2020) it is shown that the carbon fractionation of HCN and the carbon fractionation of HNC do behave similarly with a ratio close to 68. However, most of the time, the isotopic ratio for HNC is slightly larger, which can change the resulting HNC/HCN ratio by up to 40%. Therefore, using the same $^{12}\text{C}/^{13}\text{C}$ ratio for the two molecules might not be appropriate.

In L1544, HCO^+ and its deuterated and non-deuterated isotopologues have been the

Table 2.3: Derived column densities of both sources, given in units of $\times 10^{12} \text{ cm}^{-2}$.

Molecule	L1544	HH211
CCH	320(40)	210(30)
^{13}CCH	1.6(4)	1.3(5)
C^{13}CH	3.6(7)	2.2(5)
CCD	37(10)	22(5)
H^{13}CN	5(1)	1.7(2)
DCN	18(4)	9(2)
H^{13}CO^+	1.2(1)	2.2(2)
HC^{18}O^+	0.12(2)	0.11(2)
DCO^+	3.1(3)	5.4(5)
HN^{13}C	1.6(4)	0.7(2)
DNC	9(1)	4(1)

subject of multiple studies (e.g. Caselli et al. 2002b; Jørgensen et al. 2004; Vastel et al. 2006; Redaelli et al. 2019). Recently, Redaelli et al. (2019) analysed high-sensitivity maps of HC^{18}O^+ and DCO^+ , where the peak column densities are consistent with our measurements, taking into account that our observations are taken towards the dust peak and are averaged over the whole beam. To our knowledge, there is no literature data of HCO^+ and isotopologues towards HH211 so far.

2.4.2 Deuterium fraction

The deuterium fractions are derived by dividing the column density of the deuterated isotopologues by the column density of the main species. As the main species of all observed molecules, except CCH, are optically thick, we used the ^{13}C and ^{18}O isotopologues to estimate their column densities. For this, we multiplied the column densities of the ^{13}C isotopologues by $^{12}\text{C}/^{13}\text{C} = 68$ (Milam et al. 2005). Similarly, the column density of HC^{18}O^+ was multiplied by 557 (Wilson 1999) to obtain an estimate for $N(\text{HCO}^+)$. The resulting deuterium fractions are summarised in Table 2.4. Furthermore, we derived the deuteration level in the additional velocity component of HH211. The results are presented in Table A.3. Despite the high uncertainties in some cases, the additional component and the main emission generally show quite similar deuteration levels. This likely suggests that they both trace the same gas, with similar properties along the line-of-sight.

Colzi et al. (2020) show that in the local interstellar medium the carbon isotopic ratio is molecule-dependent and varies with time, volume density, and temperature, and can deviate significantly from 68. For example, for a fixed density and temperature, it may change by a factor of two between 10^5 and 10^6 years of core evolution, depending on the molecule. In this work, we applied the canonical value, taking note of the uncertainties it might introduce.

Table 2.4: Column density ratios in both cores.

	L1544	HH211
$N(\text{CCH})/N(^{13}\text{CCH})$	200(50)	160(60)
$N(\text{CCH})/N(\text{C}^{13}\text{CH})$	90(20)	100(30)
$N(\text{C}^{13}\text{CH})/N(^{13}\text{CCH})$	2.2(7)	1.7(8)
$N(\text{HN}^{13}\text{C})/N(\text{H}^{13}\text{CN})$	0.3(1)	0.4(1)
$N(\text{H}^{13}\text{CO}^+)/N(\text{HC}^{18}\text{O}^+)$	10(1)	20(4)
$N(\text{CCD})/N(\text{CCH})$	0.11(3)	0.10(3)
$N(\text{CCD})/[N(^{13}\text{CCH})\times 68]$	0.3(1)	0.2(1)
$N(\text{CCD})/[N(\text{C}^{13}\text{CH})\times 68]$	0.15(5)	0.14(5)
$N(\text{DCN})/[N(\text{H}^{13}\text{CN})\times 68]$	0.05(2)	0.07(2)
$N(\text{DCO}^+)/[N(\text{H}^{13}\text{CO}^+)\times 68]$	0.040(7)	0.037(5)
$N(\text{DCO}^+)/[N(\text{HC}^{18}\text{O}^+)\times 557]$	0.048(9)	0.09(2)
$N(\text{DNC})/[N(\text{HN}^{13}\text{C})\times 68]$	0.08(3)	0.08(3)

When interpreting the deuterium fraction, one has to keep in mind that, first, with single-dish telescopes the observations of deuterated and non-deuterated species correspond to a different beam size and therefore can cover different regions of the core. In our case, the beam sizes for the deuterated isotopologues on average are generally $10''$ larger than the telescope beams of the normal and ^{13}C isotopologues (see Table A.1). This might dilute the real deuterium fraction of a molecule and cause systematic errors in the ratios. Moreover, the species themselves are abundant in different regions or shells of the cores. In pre-stellar cores, deuterated species are expected to trace higher-density regions close to the centre of the core, where the deuteration process is most efficient (e.g. Giers et al. 2022). Non-deuterated isotopologues, on the other hand, are additionally also abundant in the outer parts of a core. In L1544, this is observed for example in HCO^+ , CH_3OH , H_2CO , and $\text{c-C}_3\text{H}_2$ (Redaelli et al. 2019; Chacón-Tanarro et al. 2019; Giers et al. 2022). Towards protostellar cores, various single dish observations have shown high levels of deuteration (van Dishoeck et al. 1995; Parise et al. 2004, e.g.). However, Persson et al. (2018) show that with interferometric observations the D/H ratio is actually lower in the inner regions, as they are less affected by optical depth effects and beam dilution. These uncertainties are important to keep in mind when interpreting single dish observations of the inner warm regions of protostellar cores. However, due to the size of the telescope beam, these regions are very much diluted, and with our data, we are mainly sensitive to the outer (and cold) protostellar envelope.

2.5 Radiative transfer modelling

In this section, we use non-LTE radiative transfer simulations to model the observed molecular lines. We applied the software LOC (Line Transfer with OpenCL, Juvela 2020), which is based on ray tracing. To reduce the complexity of the simulations, we made the simplifying assumption of a one dimensional, spherically symmetric grid for both of our cores.

For CCH, HCN, H^{13}CN , HCO^+ and DCO^+ , we used the latest collision rate coefficients available on the Leiden Atomic and Molecular Database (LAMDA, Schöier et al. 2005; van der Tak et al. 2020; Dagdigan 2018; Hernández Vera et al. 2017; Denis-Alpizar et al. 2020; Pagani et al. 2012). In the case of HCN, H^{13}CN , and DCO^+ , we applied the hyperfine-resolved rate coefficients. For the collisional coefficients of H^{13}CO^+ and HC^{18}O^+ , we used the rates of HCO^+ from LAMDA and scaled them with the corresponding reduced mass of the isotopologues. For the ^{13}C - and the deuterated species of CCH, a scaling of the CCH rates is not possible, as the hyperfine structures are different due to the different nuclear spins. Therefore, we excluded these molecules from the non-LTE modelling.

To the best of our knowledge, hyperfine resolved rate coefficients for DCN, HNC, HN^{13}C and DNC do not exist¹. A simple approach to simulating lines of these species using radiative transfer would be to use the rate coefficients of the dominant isotopologues (HCN and HNC), neglecting the effect of isotopic substitution. However, it has been shown recently (Navarro-Almaida et al. 2022) that isotope effects can be important in the case of hydrogen cyanides and isocyanides. Indeed, the comparison of the different sets of collisional data revealed that isotopic substitution can lead to substantial changes (up to $\sim 50\%$) at 10 K (the typical temperature of cold cores).

Accurate hyperfine-resolved rate coefficients for the excitation of DCN, HNC, HN^{13}C and DNC induced by collisions with H_2 are computed in this work. The scattering calculations are based on the HCN- H_2 (Denis-Alpizar et al. 2013) and HNC- H_2 (Dumouchel et al. 2011) interaction potentials corrected to consider the effect of isotopic substitution. The quantum mechanical close-coupling approach (Green 1975) and the almost exact recoupling method (Lanza & Lique 2014) was used to compute the collisional data for temperature ranging between 5 and 30 K. In the calculations, only the excitation by para- H_2 , the strongly dominant form of H_2 in cold ISM, has been considered. Details on the scattering calculations are provided in Appendix A.3.

The data underlying this article will be made available through the EMAA², LAMDA (Schöier et al. 2005; van der Tak et al. 2020), and BASECOL (Dubernet et al. 2013) data bases. They are also available on request.

2.5.1 Physical structure of L1544

The physical structure of L1544 is well-studied (e.g. Keto et al. 2015, and references therein). To model the physical structure of the core, we applied the physical model pre-

¹Hyperfine resolved rate coefficients for HNC do not exist but data of reasonable accuracy can be obtained from pure rotational rate coefficients as described in Goicoechea et al. (2022).

²<https://emaa.osug.fr/>

sented by Keto et al. (2015, hereafter Keto-Caselli model). It describes an unstable quasi-equilibrium Bonnor-Ebert sphere with a peak central H_2 volume density of $n_0 \approx 10^7 \text{ cm}^{-3}$ and a central gas temperature of 6 K. The model provides the infall velocity, density and gas temperature structure as a function of distance from the centre of the core (see Fig. 2.4).

Recent studies prove the Keto-Caselli model to be unable to explain double-peaked profiles of optically thin lines (e.g. Redaelli et al. 2019; Ferrer Asensio et al. 2022). On top of that, Redaelli et al. (2022) show that to reproduce the blue asymmetry and high level of self-absorption of HCO^+ (1-0), it is necessary to add a low-density and contracting envelope around the Keto-Caselli model. To test the effect of velocity variations in the outer core, we additionally used the HydroDynamics with Chemistry and Radiative Transfer model introduced by Sipilä et al. (2022, hereafter HDCRT model). The corresponding infall velocity, density and gas temperature structure are shown alongside the Keto-Caselli model in Fig. 2.4. The HDCRT model was designed to recreate the conditions of the Keto-Caselli model in the inner regions of the core where the gas-dust thermal coupling is strong. Unlike the Keto-Caselli model, the HDCRT model allows for expansion motions of the gas in the outer regions, caused by photoelectric heating.

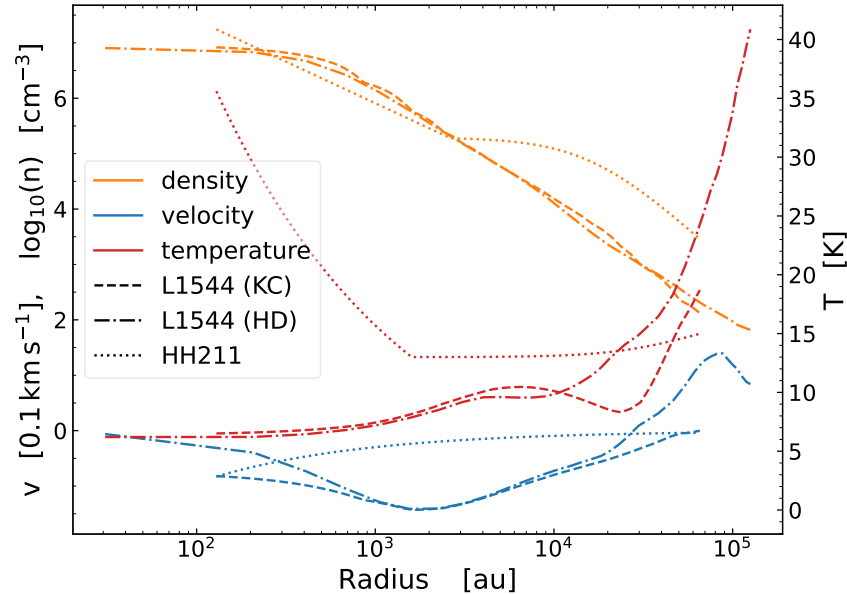


Figure 2.4: Profiles of the gas temperature (red), H_2 number density (orange, in logarithmic scale) and infall velocity (blue, in units of 0.1 km s^{-1}) for the Keto-Caselli model (dashed, KC Keto et al. 2015), the HDCRT model (dotdashed, HD Sipilä et al. 2022) of L1544, and the physical model of HH211 (dotted) derived in this work.

2.5.2 Physical structure of HH211

There is no model of the physical structure of HH211 available in the literature. We computed it using data from the *Herschel* Science Archive (HSA). Following Harju et al. (2017), we used the archive *Herschel*/SPIRE dust continuum emission maps to derive the temperature and column density maps. From those, we extracted radial temperature and volume density profiles for the outer regions of the core. For the inner regions, we used power-laws following Crimier et al. (2010) and Motte & André (2001). For the velocity, we assumed free fall with the literature mass of the protostar (Lee et al. 2018). A detailed derivation of the physical structure of HH211 can be found in Appendix A.4. The used physical model of HH211 is presented in Fig. 2.4.

2.5.3 Molecular abundances

We used the state-of-the-art gas-grain chemical model of Sipilä et al. (2019) to obtain an estimate of the fractional molecular abundances in the two cores. We applied the two-phase model of the chemical code, where the gas phase chemistry and the entire ice on the grains are active.

To account for the fact that L1544 is embedded in a molecular cloud, we assumed an external visual extinction of $A_V = 2$ mag. We employed the same initial abundances and chemical networks as in Sipilä et al. (2019) and the same standard values for various model parameters as in Giers et al. (2022), not recounted here for brevity. The abundance profiles are obtained by using the Keto-Caselli model and the HDCRT model as a static physical model of the core. We ran the chemical simulation, extracting the abundance profiles for various evolutionary times. In addition, we tested the abundances at the time of the best match between the Keto-Caselli model and the fiducial hydrodynamical simulation, as presented in Sipilä et al. (2022).

To derive the molecular abundances for HH211, we first adopted the physical conditions that roughly correspond to the edge of the HH211 model ($n_H = 5 \times 10^3 \text{ cm}^{-3}$, $T = 15 \text{ K}$, $A_V = 6 \text{ mag}$) and let the chemistry evolve over 10^6 yr in those conditions, extracting the abundances at that time to use as the initial abundances for the protostellar core model.

The deuterium chemical model does not consider the isotope chemistry of carbon, and therefore we scaled the abundance profiles by the carbon isotopic ratio of the local ISM (68, Milam et al. 2005) to obtain estimates for the ^{13}C -species. The same applies for the ^{18}O isotopologue of HCO^+ (557, Wilson 1999).

In addition to the molecular abundance profiles derived by chemical modelling, we applied a simple approach by assuming a constant abundance of the molecules throughout the core. This simplification allows to estimate the abundance levels of the molecules, especially in the case of the HDCRT model, and can be used for parameter-space explorations in future studies. However, in the case of L1544, Caselli et al. (2022) find evidence of almost complete freeze out within the central 1800 au. We took this into account by setting the abundance to zero for radii within 1500 au.

2.5.4 Results

For the comparison with the observed spectra, the synthetic ones were convolved with the respective corresponding telescope beam size (see Table A.1). To find the best fits to the observed spectra, we tested the different timesteps of the abundance profiles (10 logarithmically spaced steps between 10^5 and 10^6 years of chemical evolution) and run through a range of constant abundances in log scale. The best fits were determined by optimising the line shape, line width and peak intensity of the synthetic spectra in accordance with the observations. A selection of the best-fit synthetic spectra derived with radiative transfer modelling using LOC are presented in Figs. 2.5-2.8. The respective labelling of the applied models is explained in Table 2.5.

For L1544, the results are split between the Keto-Caselli model and the HDCRT model. The labels KC (red solid line) and HD (blue solid line) denote the results of a constant abundance applied to the respective model, while STKC (cyan dashed line) and STHD (orange dashed line) refer to the abundance profiles derived with chemical modelling. The results produced by the abundance profile from the best match between Keto-Caselli and fiducial hydrodynamical model are described by the label HD0 (green dot-dashed line). For HH211, we plot the best-fit results of applying a constant abundance (blue solid line) and an abundance profile derived with the two-phase chemical model (red dotdashed line). The observed spectra are overplotted in black. The corresponding best-fitting abundances are plotted in Fig. A.6 and Fig. A.7 for L1544 and HH211, respectively.

To account for the level of turbulence in the cores, we used the average of the non-thermal components of the observed velocity dispersions. The values that best reproduce the line widths of the optically thin species are $\sigma_{\text{turb}} = 0.1 \text{ km s}^{-1}$ (L1544, Keto-Caselli model), $\sigma_{\text{turb}} = 0.17 \text{ km s}^{-1}$ (L1544, HDCRT model) and $\sigma_{\text{turb}} = 0.23 \text{ km s}^{-1}$ (HH211). Separate values for the two different physical models of L1544 are necessary, because the outer radius of the HDCRT model is a factor of two larger compared to the Keto-Caselli model. Therefore, the model covers more turbulent gas.

To account for the additional velocity component at 8.22 km s^{-1} observed towards HH211, we approximated it by a static Gaussian that is added to the modelling results afterwards. For this, we used the results of the two-component Gaussian fits listed in Table A.2. In the case of HNC, where a two-component Gaussian fit was not successful, we estimated the static Gaussian by eye to improve the readability of the modelled spectra.

L1544

The synthetic spectra of CCH are shown in Fig. 2.5, with each hyperfine component presented in a separate plot (numbered from low to high frequency). Towards L1544, the constant abundances applied to the two different physical models are able to reproduce the intensities of the more optically thin hyperfines. The line profiles of the more optically thick components are not reproduced by the Keto-Caselli model. The observed lines show rather symmetric double peaks, whereas the synthetic lines display a blue asymmetry. This

Table 2.5: Summary of the simulation setups discussed in this work.

Model	Description
L1544	
KC	Constant abundance throughout the core with a drop to zero in the central 1500 au, applied to the Keto-Caselli model for L1544 (Keto et al. 2015)
STKC	Abundance profile derived with a static physical model adopting the parameters and physical structure of the Keto-Caselli model for L1544
HD	Constant abundance throughout the core with a drop to zero in the central 1500 au, applied to the HDCRT model for L1544
STHD	Abundance profile derived with a static physical model adopting the parameters and physical structure of the HDCRT model for L1544
HD0	Abundance profile derived from the best match between the Keto-Caselli and the fiducial hydrodynamical model presented in Sipilä et al. (2022)
HH211	
const	Constant abundance throughout the core, applied to the physical model for HH211 derived in this work
2-phase	Abundance profile derived with a static physical model adopting the parameters and physical structure of the physical model for HH211 derived in this work, applying the 2-phase model of the chemical code (= gas phase chemistry and entire ice on the grains are active)

behaviour was also observed for CS ($2 - 1$) by Tafalla et al. (1998) and might be an indication that these transitions trace a static envelope. Using the HDCRT model and a constant abundance, we are indeed able to reproduce the symmetric double-peak profiles, except for hyperfine component 5. The abundance profiles derived with chemical modelling largely underestimate the amount of CCH, and therefore fail to reproduce the lines in both the STKC and the STHD case. To reproduce the same intensities as in the KC case, the STKC abundance profile has to be multiplied by a factor of 20. The discrepancy between the simulated and observed CCH abundances is most likely related to the uncertainties in the chemistry of CCH, though the use of a simple 1D physical model may contribute as well. Future studies of the formation and destruction pathways of CCH are required to improve the chemical simulations, as well as the adoption of 3D models (Jensen et al. 2023).

The synthetic spectra of HCN, H^{13}CN , and DCN are shown in Fig. 2.6. The Keto-Caselli model fails at reproducing HCN with both a constant abundance and an abundance profile. The synthetic spectra show blue asymmetries in all hyperfines, indicating infall motion, and do not recreate the red asymmetry observed in the central hyperfine. This shows that the applied abundance profiles overestimate the amount of HCN in regions of infall and indicates that HCN might have a higher abundance in the low-density regions of the core. In addition, the modelled dip caused by the self-absorption does not reach down to zero level. This might be solved by adding an extended low-density envelope to the Keto-Caselli model, such as in the case of HCO^+ (Redaelli et al. 2022). The HDCRT model, on the other hand, is able to produce the red asymmetry profiles for HCN by applying the STHD abundance profile. However, it results in red asymmetries for all hyperfines, reproducing the intensity of two hyperfines, but underestimating the weakest component by a factor of two. Applying a constant abundance results in a rather symmetric line profile for all components. The self-absorption dip of the synthetic spectrum goes down to almost zero level, as observed. A comparison of the STHD and HD0 abundance profiles (see Fig. A.6) shows that, in the case of HD0, HCN is more abundant at larger radii than in the case of STHD. This results in a deeper self-absorption dip. The distance between the two peaks is not reproduced (see Fig. 2.6), indicating that for HCN the turbulence level of 0.17 km s^{-1} is overestimated in the hydrodynamical case.

In the case of H^{13}CN (see Fig. 2.6), the Keto-Caselli model works very well and is able to reproduce the spectrum with both the KC and STKC abundance profile. The HDCRT model, however, struggles to reproduce the correct shapes of the self-absorbed hyperfines. Both the HD and STHD abundance profiles result in symmetric rather than blue asymmetric lines, which might hint at an overestimation of H^{13}CN in the outer regions and/or a radially changing carbon isotopic ratio. In the case of HD0, the spectrum is severely underestimated. This might be a hint that at this specific timestep chosen by Sipilä et al. (2022), the carbon isotopic ratio is actually below 68. For DCN (see Fig. 2.6), the Keto-Caselli model can reproduce the intensities of the hyperfines within a factor of two. However, it fails to reproduce the shapes of the lines. Both the KC and the STKC abundance profiles predict a strong red asymmetry that is not observed. In addition, the synthetic spectra show higher self-absorption than the observed spectrum. This shows that

the abundance profiles are actually overestimating the amount of DCN and that in reality there is less material available to trace the infall motions. The HDCRT model cannot reproduce the spectrum, neither the line shapes nor the intensities match the observations.

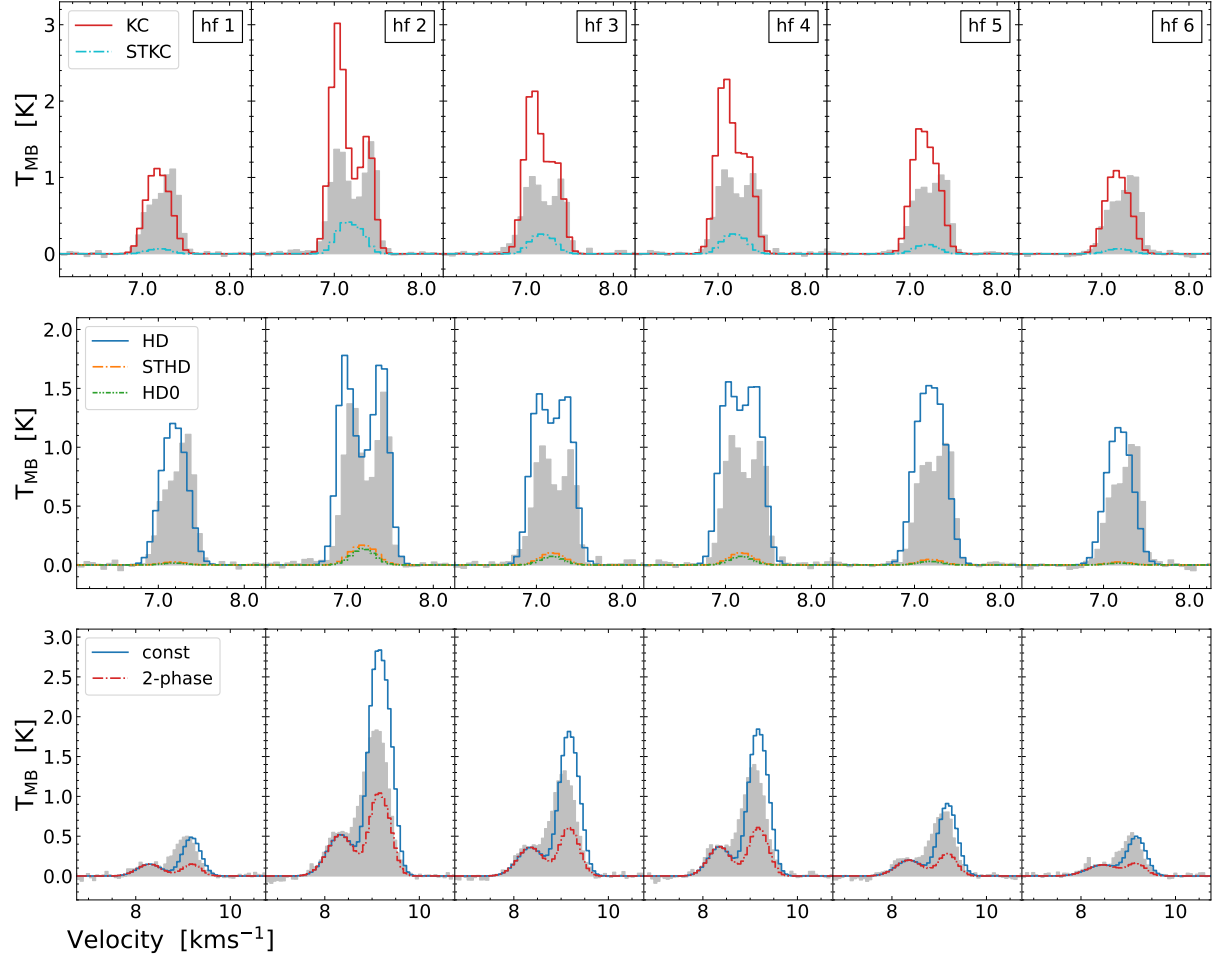


Figure 2.5: Synthetic spectra of CCH obtained with LOC (top rows: L1544, bottom row: HH211). Each hyperfine line is plotted separately, numbered from low to high frequency. For L1544, the results are splitted between using the Keto-Caselli model (top) and the HDCRT model (centre) as approximation for the physical structure of the core. Towards L1544, the best-fitting abundances and timesteps are 2×10^{-8} (KC), $1e5$ yr (STKC), 3×10^{-8} (HD), $1e5$ yr (STHD). Towards HH211, the best-fitting abundances and timesteps are 2×10^{-9} (const), $1e6$ yr (2-phase).

The synthetic spectra of HNC, HN^{13}C , and DNC are shown in Fig. 2.7. The collision rate coefficients for HN^{13}C and DNC only consider the spin of the ^{14}N for the hyperfine splitting. Therefore, the modelling fails to reproduce the lowest frequency component of these two molecules. However, work is underway to address this issue, where the collisional rate coefficients are computed by also including the effect of ^{13}C and D, respectively, which

will allow to reproduce the full spectrum. Towards L1544, HNC can be reproduced quite well, though the Keto-Caselli model seems to perform better than the HDCRT model at reproducing the observed intensity peaks. However, due to the greater physical size of the structure in the HDCRT model, the dip at 7.2 km s^{-1} is better reproduced than in the case of the Keto-Caselli model. At high velocity there seems to be an issue with the hyperfine component that has the lowest statistical weight. This might be caused by an underestimation of the general abundance of the molecule and subsequently of the depth of self-absorption in the lower velocity hyperfine components.

In the case of HN^{13}C , the Keto-Caselli model is able to reproduce the spectrum with both constant abundance and abundance profile, except for the high velocity component. The HDCRT model, however, results in a stronger self-absorption and therefore underestimates the observed intensity. In the case of DNC, the STKC abundance profile, applied to the Keto-Caselli model, results in a good match with the observed line shape and intensity (excluding the high velocity component). The synthetic spectrum produced by the constant abundance largely overestimates the amount of DNC in the outer regions of the model, resulting in a deep dip caused by self-absorption of the strongest hyperfines that is not observed. The HDCRT model is not able to reproduce DNC, it underestimates both intensity and line width. A comparison of the abundance profiles (see Fig. A.6) shows that the peaks of the STHD and the HD0 profiles are at least one magnitude lower than for the STKC profile, which leads to an underestimation of the molecular abundance.

The synthetic spectra of HCO^+ are shown in Fig. 2.8. Towards L1544, our models fail to reproduce the strong self-absorption and asymmetry of HCO^+ . However, the case of this molecule is extensively studied in Redaelli et al. (2022). By adding a low-density (27 cm^{-3}), contracting envelope extending out to 1 pc, they manage to reproduce the line shape of the transition. In the case of H^{13}CO^+ , a constant abundance profile applied to the Keto-Caselli model can reproduce the observed spectrum. The modelled line profile shows a stronger blueshift and a higher intensity than observed. However, a constant abundance of H^{13}CO^+ with complete freeze out in the central 1500 au is clearly too simplistic for this molecule, present not only in the core, but also in the surrounding cloud. In the case of the HDCRT model, a constant abundance of H^{13}CO^+ results in an almost symmetric double-peak profile. The abundance profiles derived by chemical modelling severely underestimate the observed intensity and cannot reproduce the spectrum. This might indicate that in this case the carbon isotopic ratio is less than 68.

For HC^{18}O^+ , the line cannot be reproduced by either of the models. The HD abundance profile roughly reproduces the line width, but fails at the flattened top caused by CO depletion in the central regions of the core. Like in the case of H^{13}CO^+ , the synthetic spectra derived with the abundance profiles underestimate the intensity of the line and indicate that the assumption of the isotopic ratio to be 557 might be not correct in this case. For DCO^+ , the observed line shape cannot be reproduced. The simulation is considering the overlapping hyperfine structure of the molecule and the modelled spectra show a rather flat peak. This is likely caused by individual, self-absorbed hyperfine components.

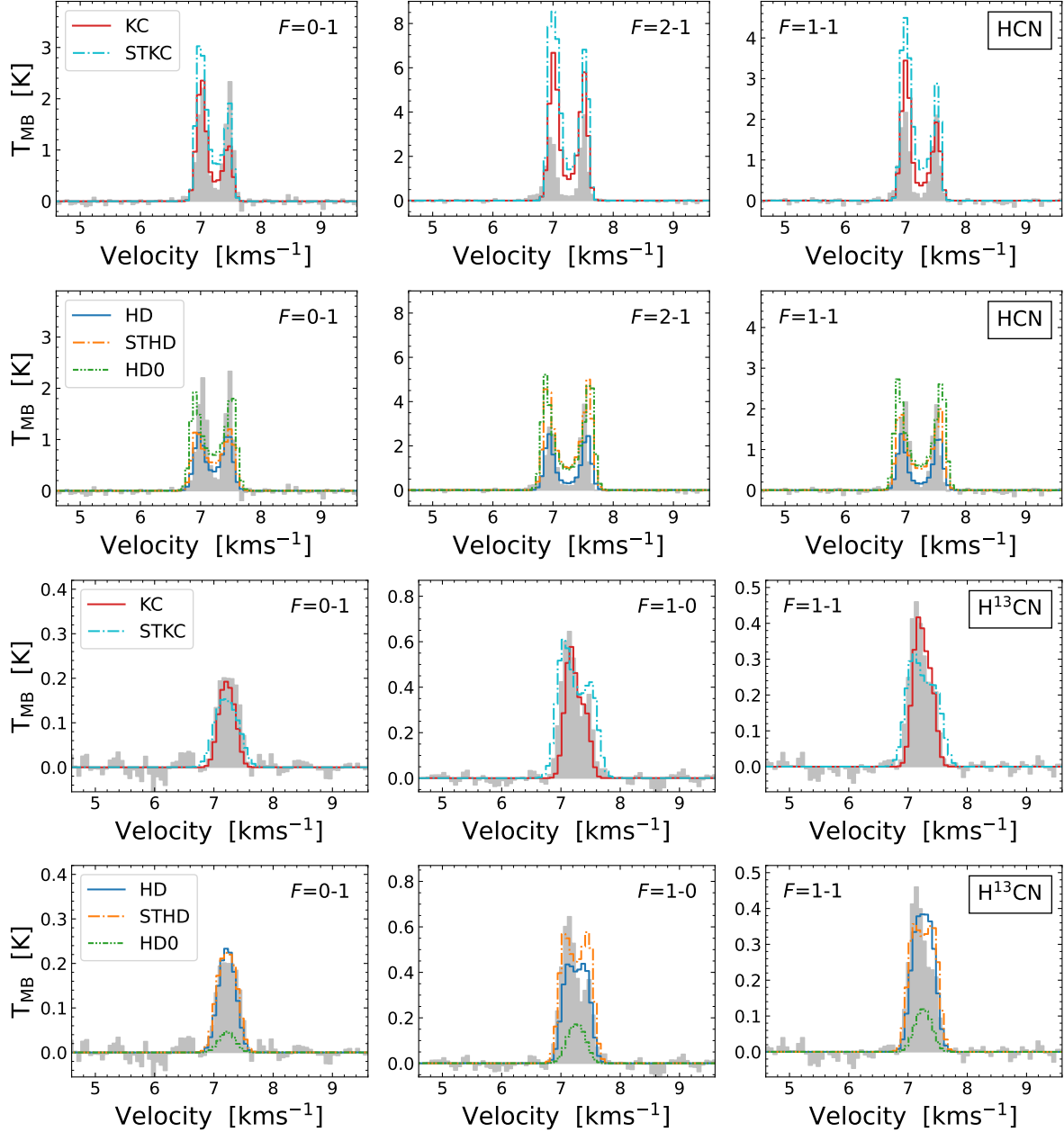


Figure 2.6: Synthetic spectra of HCN, H^{13}CN , and DCN obtained with LOC (top: L1544, bottom: HH211). For L1544, the results are splitted between using the Keto-Caselli model (KC, STKC) and the HDCRT model (HD, STHD, HD0) as approximation for the physical structure of the core. Towards L1544, the best-fitting abundances and timesteps are 2×10^{-8} (KC), 7.9e5 yr (STKC), 1×10^{-8} (HD), 1.3e5 yr (STHD) for HCN; 2×10^{-10} (KC), 6e5 yr (STKC), 3×10^{-10} (HD), 1e5 yr (STHD) for H^{13}CN ; 1×10^{-9} (KC), 1.3e5 yr (STKC), 1×10^{-9} (HD), 9.8e5 yr (STHD) for DCN. Towards HH211, the best-fitting abundances and timesteps are 2×10^{-9} (const), 1e5 yr (2-phase) for HCN; 3×10^{-11} (const), 1.3e5 yr (2-phase) for H^{13}CN ; 4×10^{-10} (const), 1e5 yr (2-phase) for DCN.

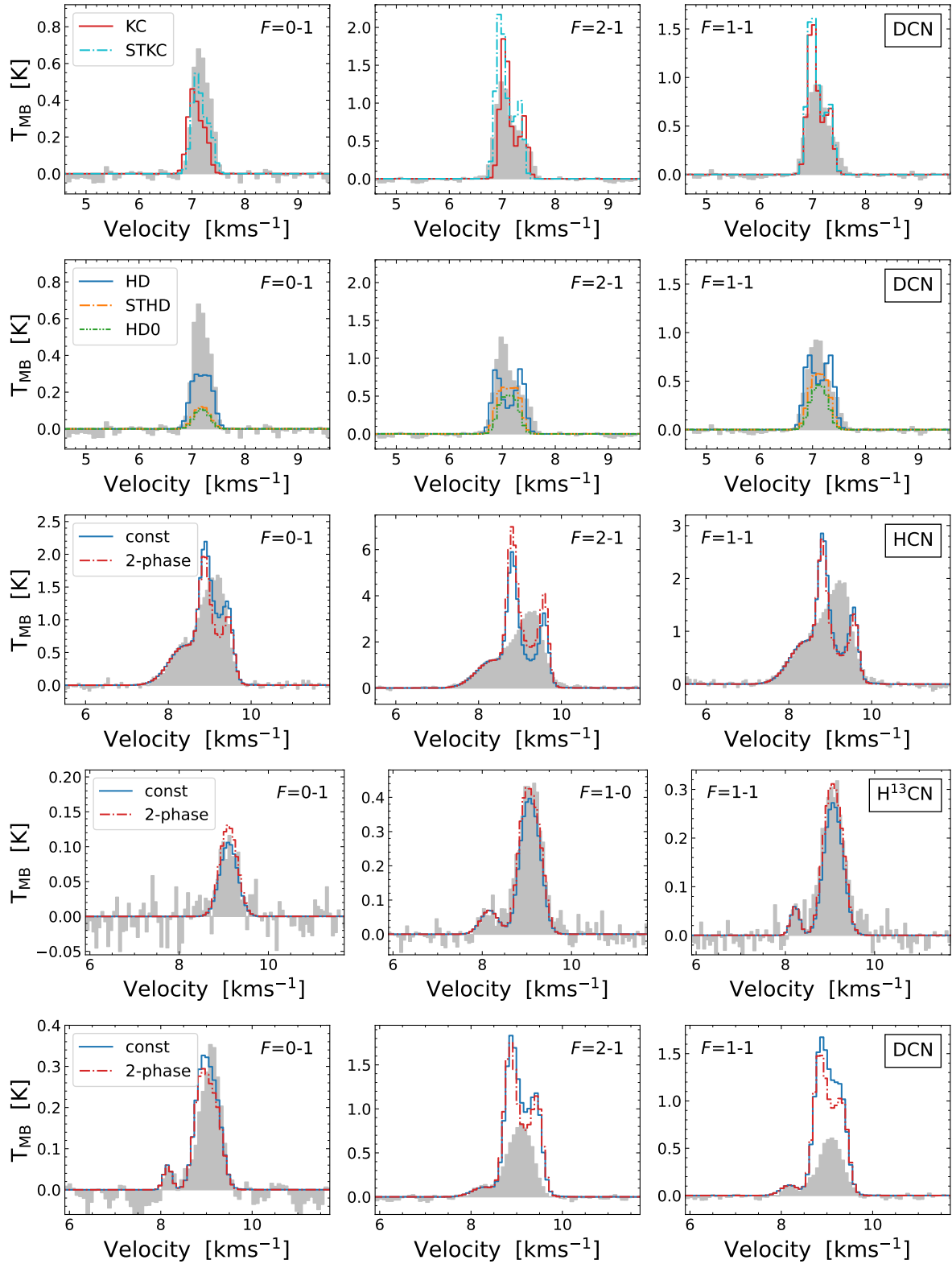


Figure 2.6: continued.

HH211

The synthetic spectra of CCH are shown in Fig. 2.5, with each hyperfine component presented in a separate plot (numbered from low to high frequency). The observed spectrum can be reproduced assuming a constant molecular abundance of 2×10^{-9} with respect to H_2 throughout the core. The intensity ratios between the hyperfine components are reproduced within a factor of two in the abundance. The synthetic spectra resulting from abundance profiles derived with the 2-phase chemical model, where gas phase and grain chemistry are considered, underestimate the observed lines by a factor of two to five. This supports the need to constrain the reactions involved in the chemistry of CCH to improve the chemical simulations.

The synthetic spectra of HCN, H^{13}CN , and DCN are shown in Fig. 2.6. Towards HH211, we are not able to reproduce the observed spectrum of HCN. Applying the derived physical model of the protostellar core produces a self-absorbed line profile with blue asymmetry for HCN that is not observed. The ^{13}C isotopologue, however, can be reproduced with the correct intensities, line shapes and line widths with both the constant abundance and the abundance profile. In the case of DCN, only the weakest hyperfine component is reproduced. The more optically thick hyperfines, however, are overestimated by the model by a factor of two to three. This might indicate that DCN is more extended than what the model covers currently, resulting in more self-absorption than produced by the model. This should be addressed in future studies.

The synthetic spectra of HNC, HN^{13}C , and DNC are shown in Fig. 2.7. The modelled HNC shows an asymmetric self-absorption that is not observed. Also here, the linewidth towards high velocity is underestimated, similar to HNC in L1544. Taking into account the missing hyperfine component at high velocity ($\sim 10 \text{ km s}^{-1}$), HN^{13}C and DNC are well reproduced by a constant abundance of 1×10^{-11} and 7×10^{-11} , respectively. For HN^{13}C , the synthetic spectrum derived with an abundance profile shows that also in this case, the assumption of $^{12}\text{C}/^{13}\text{C} = 68$ is underestimating the amount of HN^{13}C present. For DNC, the use of an abundance profile results in a self-absorbed line profile, most likely due to an overestimation of the molecular abundance in the outer regions.

The synthetic spectra of HCO^+ and isotopologues are shown in Fig. 2.8. Towards HH211, all isotopologues of HCO^+ can be reproduced in shape and intensity, using a constant abundance profile. The line shape of the main species can be modelled, however, the constant abundance used to achieve this is of the same magnitude as for DCO^+ . This indicates that the spectrum of HCO^+ is heavily self-absorbed. When applying an abundance profile, this self-absorption is shown in the synthetic spectrum. The same happens in the case of DCO^+ , indicating that this molecule might be slightly optically thick as well. The ^{13}C and ^{18}O isotopologues cannot be reproduced with the HCO^+ abundance profile scaled down by the respective isotopic ratio. This highlights the necessity of new constraints on the isotopic abundance ratios in varying physical conditions.

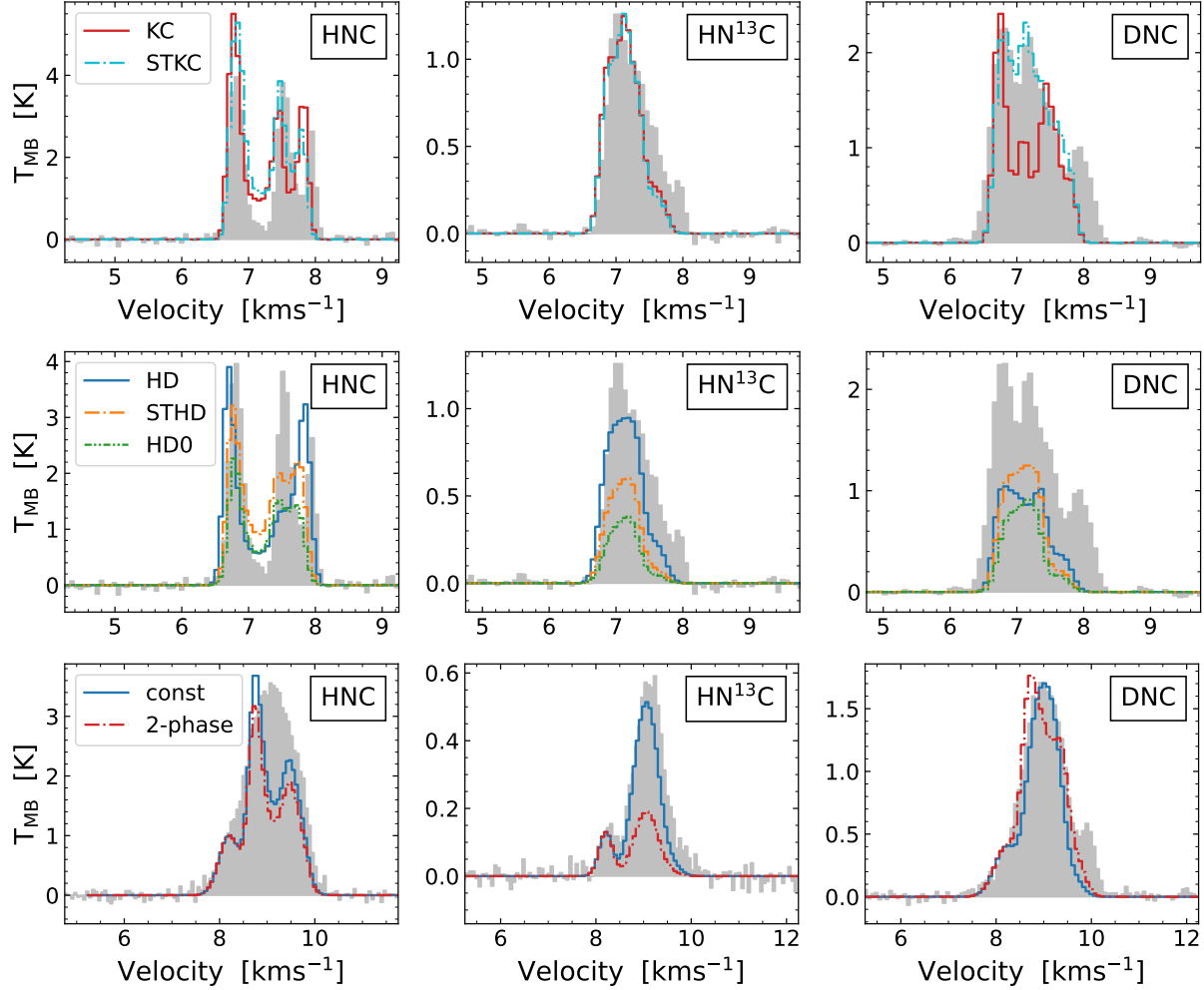


Figure 2.7: Synthetic spectra of HNC and isotopologues obtained with LOC (top rows: L1544, bottom row: HH211). For L1544, the results are splitted between using the Keto-Caselli model (top) and the HDCRT model (centre) as approximation for the physical structure of the core. Towards L1544, the best-fitting abundances and timesteps are 2×10^{-8} (KC), 1e5 yr (STKC), 5×10^{-8} (HD), 1e5 yr (STHD) for HNC; 2×10^{-10} (KC), 1.3e5 yr (STKC), 2×10^{-10} (HD), 9.8e5 yr (STHD) for HN^{13}C ; 3×10^{-9} (KC), 1e5 yr (STKC), 8×10^{-10} (HD), 9.8e5 yr (STHD) for DNC. Towards HH211, the best-fitting abundances and timesteps are 5×10^{-10} (const), 1e5 yr (2-phase) for HNC 1×10^{-11} (const), 1e5 yr (2-phase) for HN^{13}C ; 7×10^{-11} (const), 1e6 yr (2-phase) for DNC.

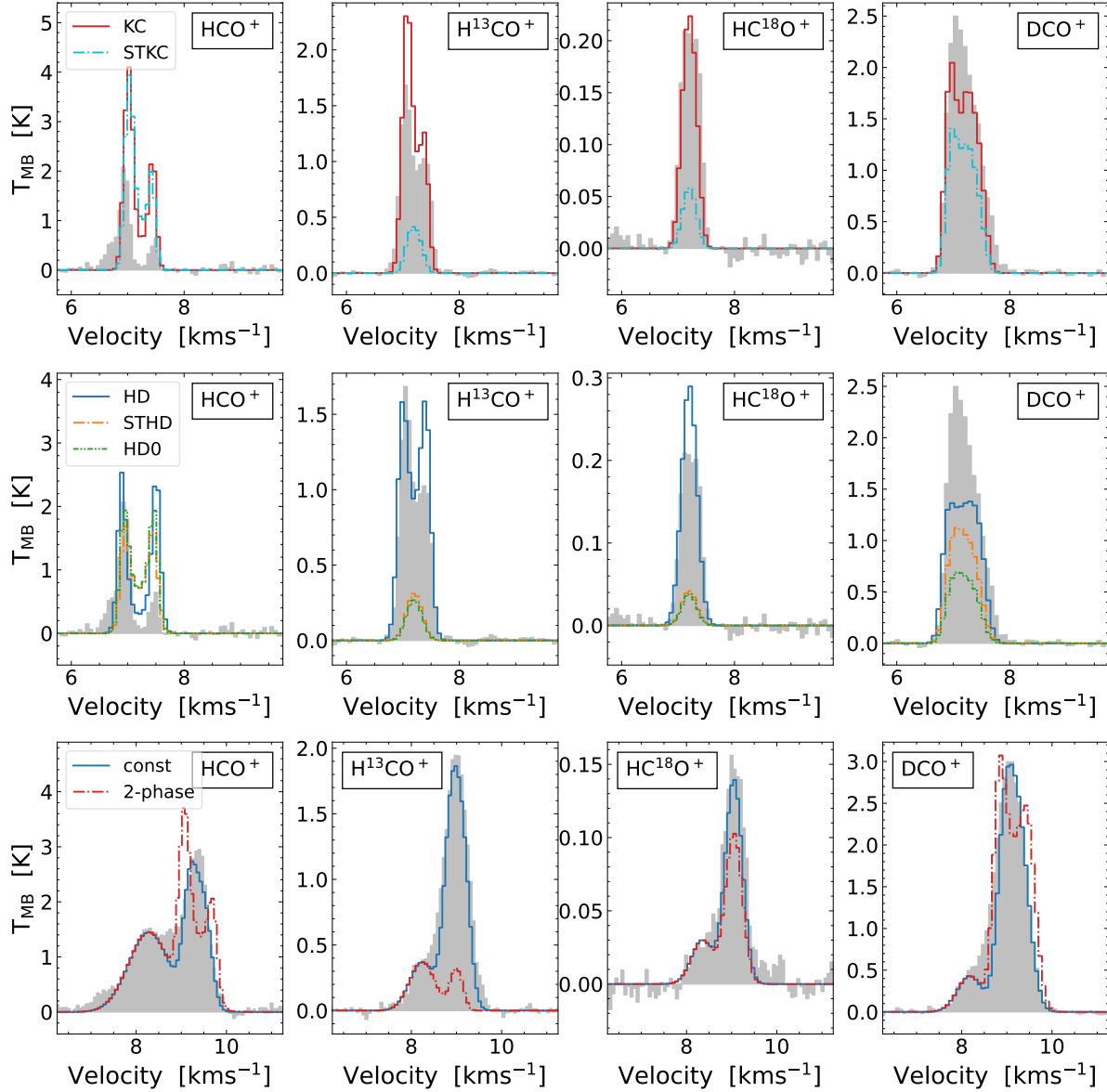


Figure 2.8: Synthetic spectra of HCO^+ and isotopologues obtained with LOC (top rows: L1544, bottom row: HH211). For L1544, the results are splitted between using the Keto-Caselli model (top) and the HDCRT model (centre) as approximation for the physical structure of the core. Towards L1544, the best-fitting abundances and timesteps are 7×10^{-10} (KC), $9.8\text{e}5$ yr (STKC), 3×10^{-9} (HD), $9.8\text{e}5$ yr (STHD) for HCO^+ ; 2×10^{-10} (KC), $1\text{e}5$ yr (STKC), 4×10^{-10} (HD), $1\text{e}5$ yr (STHD) for H^{13}CO^+ ; 6×10^{-12} (KC), $1\text{e}5$ yr (STKC), 1×10^{-11} (HD), $1\text{e}5$ yr (STHD) for HC^{18}O^+ ; 2×10^{-10} (KC), $1.7\text{e}5$ yr (STKC), 4×10^{-10} (HD), $1.7\text{e}5$ yr (STHD) for DCO^+ . Towards HH211, the best-fitting abundances and timesteps are 1×10^{-11} (const), $1\text{e}6$ yr (2-phase) for HCO^+ ; 7×10^{-12} (const), $1\text{e}5$ yr (2-phase) for H^{13}CO^+ ; 4×10^{-13} (const), $1\text{e}5$ yr (2-phase) for HC^{18}O^+ ; 2×10^{-11} (const), $1\text{e}6$ yr (2-phase) for DCO^+ .

2.6 Discussion

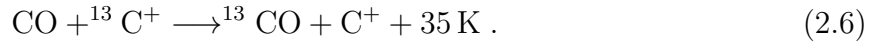
2.6.1 Dilution of ^{13}C in CCH

We find that C^{13}CH is more abundant than ^{13}CCH in L1544 and in HH211 (see Table 2.4). This characteristic appears to be common in cold dark clouds, and was observed already in the starless cores TMC-1, L1521B and L134N (Turner 2001; Sakai et al. 2010; Taniguchi et al. 2019), and in the protostellar cores L1527 and L483 (Sakai et al. 2010; Agúndez et al. 2019). Taniguchi et al. (2019) conclude that this difference in abundances is caused during the formation pathway of the molecule and the isotopomer-exchange reaction after it is formed. This reaction is exothermic:



thus, at low temperatures the forward reaction is more efficient, increasing the abundance of the C^{13}CH isotopomer. Chemical models of Furuya et al. (2011) predicted an increase of the $\text{C}^{13}\text{CH}/^{13}\text{CCH}$ ratio with pre-stellar evolution. Literature values of this ratio are collected in Fig. 2.9, which shows the ratios observed towards the starless cores L1521B, TMC-1, L134N, the very evolved pre-stellar core L1544, and the protostellar cores HH211 and L1527. However, due to the large uncertainties, no conclusive statement can be made. More observations are needed to confirm any trends.

The carbon isotopic ratios observed for CCH are 200 ± 50 and 90 ± 20 (L1544), and 160 ± 60 and 100 ± 30 (HH211) for ^{13}CCH and C^{13}CH , respectively. This is significantly higher than the canonical value for the ISM, 68, and is caused by the dilution of ^{13}C in carbon-chain molecules in the local ISM. In dark clouds, ^{13}C is mainly locked in ^{13}CO due to the reaction



As this reaction is exothermic, the forward reaction is more efficient at low temperatures. Therefore, molecules such as CCH or $\text{c-C}_3\text{H}_2$ that are produced from C^+ show a significantly higher $^{12}\text{C}/^{13}\text{C}$ ratio in molecular clouds. This dilution of ^{13}C species was predicted by Langer et al. (1984), and confirmed by previous studies (Sakai et al. 2010; Agúndez et al. 2019; Taniguchi et al. 2019; Yoshida et al. 2019). In Colzi et al. (2020), this behaviour is also predicted for CH.

2.6.2 Deuterium fractionation

Due to the dilution of the ^{13}C isotopologues of CCH, the derived deuterium fractions have to be interpreted as upper limits, as the applied $^{12}\text{C}/^{13}\text{C}$ ratio of 68 is underestimating the real value. However, as it was possible to derive the column density of CCH using emission of the main species directly, we use these results for a comparison of the deuteration in the two cores. CCH shows a high and similar level of deuteration in both cores ($\approx 10\%$) and is therefore consistent with other carbon chains such as $\text{c-C}_3\text{H}_2$ (e.g. Chantzos et al. 2018). Measurements towards other cores report lower deuteration levels (L183: 0.06(4), Turner

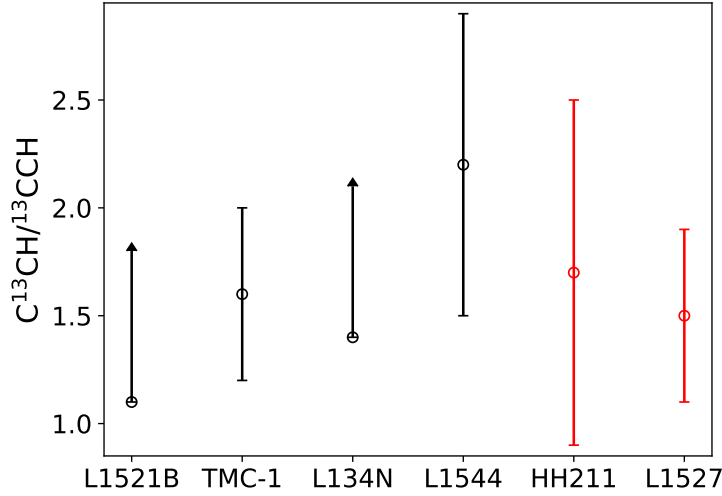


Figure 2.9: $C^{13}CH/^{13}CCH$ ratio of starless/pre-stellar (black) and protostellar (red) cores. L1521B: Taniguchi et al. (2019); L134N: Taniguchi et al. (2019); TMC-1: Sakai et al. (2010); L1527: Yoshida et al. (2019).

2001; TMC-1: 0.05(2), Turner 2001; L1527: 0.04(1), Yoshida et al. 2019). Evolutionary and environmental effects might play a role in this difference, as well as differences in the excitation temperatures used to derive the column densities.

The deuterium fraction of HCN appears to be slightly larger in HH211 with respect to L1544. However, due to the uncertainties in the carbon isotopic ratio of HCN, these numbers have to be interpreted with caution. Therefore, no conclusive statement can be made based on our errorbars, and further studies are necessary to eventually confirm this trend. The D/H ratio for HCN of the cold dark cloud L183 (0.05(2), Turner 2001) is consistent with our measurements towards L1544. Roberts et al. (2002) derived slightly lower values in HH211, $N(DCN)/N(HCN) = 0.038(8)$, which can be explained by different approximations used in the calculations. The measurements towards the protostellar core L1527 are consistent with our observations (0.05(1), Yoshida et al. 2019).

For HNC, we derive a deuterium fraction of 0.08(2) for both L1544 and HH211. Previous measurements of the deuteration in the two cores show moderate deuterium fractions (L1544: 0.03, Hirota et al. 2003; HH211: 0.07, Imai et al. 2018). However, the authors considered the emission of $HN^{13}C$ and DNC to be optically thin and apply simple one-component Gaussian fits to obtain the line parameters. Especially in the case of DNC this can lead to an underestimation of the column density, and subsequently of the deuterium fraction. Observations towards other cores show moderate deuterium fractions (L183: 0.05(2), Turner 2001; L1527: 0.045, Yoshida et al. 2019).

The deuterium fractions of HCO^+ derived towards L1544 and HH211 are consistent with previous measurements in similar objects (see e.g. Turner 2001; Jørgensen et al. 2004; Koumpia et al. 2017; Yoshida et al. 2019). For L1544, the D/H ratios we derived using $H^{13}CO^+$ and $HC^{18}O^+$ are in agreement with previous observations by Caselli et al. (2002a)

and Redaelli et al. (2019), who report a ratio of around 4%.

Towards HH211, the result derived using the ^{18}O isotopologue is increased by a factor of two compared to the D/H ratio derived from H^{13}CO^+ . This is most likely caused by an overestimation of the elemental carbon isotopic ratio. Chemical models predict that the $\text{HCO}^+/\text{H}^{13}\text{CO}^+$ ratio is actually lower than the canonical $^{12}\text{C}/^{13}\text{C}$ ratio (Colzi et al. 2020), because of isotopic exchange reactions that are important at low temperatures. The canonical ratio of $^{13}\text{C}/^{18}\text{O}$ is 8.2, whereas our observations give 10(1) and 20(4) for L1544 and HH211, respectively. This shows that especially towards HH211, the carbon fractionation of HCO^+ is higher than in the local ISM, resulting in $\text{HCO}^+/\text{H}^{13}\text{CO}^+ < 68$. Hence, by using 68, the deuterium fraction derived from H^{13}CO^+ is underestimated.

The comparison of the D/H ratio derived from HC^{18}O^+ in the two cores shows that the ratio towards the protostellar core is higher by a factor of almost two. This difference is caused by some effect that is either increasing or decreasing the observed column densities of the isotopologues. The latter one could be an effect of isotope-selective photodissociation which is impacting the ^{18}O more than the D isotopologue. The higher abundance of CO leads to a stronger self-shielding of the molecule compared to C^{18}O . This results in a local enhancement of $\text{CO}/\text{C}^{18}\text{O}$, that is likely larger in HH211 than in L1544. This is reflected in an increased D/H ratio towards HH211, as DCO^+ is created directly from CO, and therefore it is more efficiently formed than HC^{18}O^+ .

In conclusion, we do not see a general trend in the deuteration efficiency in simple molecules going from a pre-stellar to a protostellar core, as it was indicated previously by observations of $\text{c-C}_3\text{H}_2$ (Chantzios et al. 2018). This might be due to a more efficient deuteration of $\text{c-C}_3\text{H}_2$ happening on the surface of the dust grains in the pre-stellar phase that is not applicable for the simple molecules studied in this work. Additionally, the fact that the deuterium fraction is similar in two cores in different environments seems to suggest that the deuterium fractionation is not sensitive to the initial conditions of its surroundings. Similar results have been found for N_2H^+ (Crapsi et al. 2005; Emprechtinger et al. 2009) and $\text{c-C}_3\text{H}_2$ (Chantzios et al. 2018). Nevertheless, more comprehensive surveys including multiple pre-stellar and protostellar cores and covering more molecules are necessary to provide larger statistics and confirm any trends.

2.6.3 Radiative transfer modelling

L1544

Our radiative transfer simulations show that the Keto-Caselli model works fine to explain moderately optically thin lines such as H^{13}CN , HN^{13}C , DNC , and H^{13}CO^+ that trace regions with moderate to high density (few 10^5 cm^{-3}). The model shows problems when it comes to optically thin lines like HC^{18}O^+ , where the line shape is affected by the freeze-out of CO in addition to contraction motions. In the case of CCH, the model fails to reproduce the line shapes of the optically thin hf components, which show a red asymmetry. As the lines are optically thin, they do not trace any expansion motion like it was seen for HCN. Instead, this is likely an effect of the asymmetric distribution of the molecule across the

core. In L1544, carbon chains are known to peak towards the south-east of the core (Spezzano et al. 2017; Giers et al. 2022). Due to the velocity gradient across the core (Spezzano et al. 2016), the spectra observed towards the south are redshifted with respect to the rest velocity of the system. As CCH is most probably peaking in the south, where the gas motion is redshifted, the contribution of the redshifted lines is more prominent, resulting in an asymmetric line shape. This behaviour is also observed for cyanopolyynes by Bianchi et al. (2023). As we assume a spherically symmetric core in the simulations, and therefore a symmetric distribution of the molecule, this line shape is not reproducible.

The HDCRT model, on the other hand, struggles to explain the blue asymmetry observed in the moderately optically thin lines. This is most likely caused by an overestimation of the molecular abundances in the outer regions, resulting in symmetric double-peak profiles that trace more static rather than infalling layers.

In most cases, the simulations fail to reproduce the observed intensities of the rarer isotopologues when applying the abundance profiles of the main isotopologues scaled by the respective isotope ratio. This highlights the importance of considering the carbon and oxygen isotope chemistry in the chemical modelling, which will be addressed in future work.

In the case of the optically thick and more widespread main isotopologues, the results are better reproduced when applying the HDCRT model, while the Keto-Caselli model struggles to explain the lines. The only exception here is HNC, where the Keto-Caselli model is sufficient to reproduce the observed intensities. This might be an indication that HNC is actually tracing slightly different layers compared to HCN. In the case of CCH, the HDCRT model works fine for the optically thick components that trace a static inbetween layer.

In the case of HCN, the radiative transfer modelling in this work showed that the Keto-Caselli model is not sufficient to explain the line shapes and intensities of the molecule’s hyperfine components observed towards L1544. Instead, it is necessary to use a physical model that shows expansion motions in the outer layers to reproduce the red asymmetry in the central hyperfine component of HCN. Contrary to this, Redaelli et al. (2022) show that for the blue asymmetry and strong self-absorption observed in HCO^+ , it is necessary to add an extended low-density and contracting envelope to the Keto-Caselli model. One possibility to solve this discrepancy could be that HCN is tracing an even outer layer than HCO^+ . As HCN can be efficiently formed also without CO, the layers where this process occurs could be warm enough to be associated with expansion motions.

HH211

The physical model derived in this work struggles to explain the very optically thick main isotopologues (HCN, HNC, HCO^+). Both a constant abundance across the core and the abundance profiles predict high molecular abundances in the outer regions of the core. This leads to a strong self-absorption in the synthetic spectra that is not observed. To get a more realistic estimate of the molecular abundances across the core, it would be necessary to run more tests of the physical conditions and the parameter space used in the chemical

modelling. However, this is beyond the scope of this paper.

Furthermore, very abundant and widespread molecules like HCO^+ , HCN , and possibly also HNC also trace the outflow of the protostar and the extended structures of the core's environment. These kind of large-scale structures are not considered in the simplistic and spherically symmetric physical model derived in this work and will likely be different for the sources studied here. L1544 is located in an isolated region in Taurus. In contrast, HH211 is forming in an active region within Perseus, with low density material traced by ^{13}CO spreading over several km s^{-1} around HH211 (Sun et al. 2006). The complex and turbulent environment towards HH211 is reflected in the line profiles for the more abundant species. For instance, our observations detect a second velocity component blueshifted by 1 km s^{-1} , which is associated with large-scale emission extending to the north-east of HH211 (Sun et al. 2006). To correctly model the observed line shapes of molecular tracers like HCO^+ , HCN , and HNC , these structures have to be taken into account in the density, velocity, and temperature profiles of the core.

On the other hand, we are able to reproduce all optically thin and moderately optically thick molecular lines observed towards the protostellar core. For these spectra, the simplifying assumptions of spherical symmetry and constant abundance throughout the core seem to be sufficient to reproduce the molecular emission. This is most likely due to the fact that these molecules are less abundant and trace higher-density regions, and are therefore less affected by the kinematics of the large-scale structures surrounding the core as well as the extended outflow emission.

2.7 Conclusions

We presented a survey of ground-state rotational lines of simple molecules, which allow us to compare the levels of deuteration in the very evolved pre-stellar core L1544 and the very young protostellar core HH211. In a non-LTE approach, we used radiative transfer simulations and molecular abundance profiles derived from chemical modelling to reproduce the observed molecular spectra. By applying new hyperfine-resolved collisional rate coefficients, we take into account the specific spectroscopy of the respective isotopologues.

Our main results can be summarised as follows:

- The similar levels of deuteration show that the deuterium fractionation seems to be equally efficient towards both cores. We do not see a general trend in the level of deuteration when going from a pre-stellar core to a protostellar core, suggesting that the protostellar envelope still retains the chemical composition of the original pre-stellar core. In addition, the deuterium fraction seems to be independent of the initial conditions present in the molecular clouds where the two cores are embedded.
- Towards HH211, the D/H ratio of HCO^+ derived using the ^{18}O isotopologue is higher by a factor of two compared to L1544. This is likely an effect of isotope-selective photodissociation, creating a local enhancement of $\text{CO}/\text{C}^{18}\text{O}$ in the protostellar core.

- The ^{13}C dilution of carbon-chain species in dark clouds leads to an increased $^{12}\text{C}/^{13}\text{C}$ ratio for CCH. On the other hand, the increased $^{13}\text{C}/^{18}\text{O}$ ratio of HCO^+ indicates a decreased $^{12}\text{C}/^{13}\text{C}$ ratio for the molecule, caused by the isotopic exchange reaction being more effective at low temperatures. This highlights the uncertainties when dealing with ^{13}C isotopologues and the influence of the applied carbon isotopic ratio.
- The central hyperfine component of HCN (1-0) observed towards L1544 shows a red asymmetry, tracing expansion motions caused by external heating in the outer layers of the core. This stands in contrast to HCO^+ that has been shown to trace infall motion in the outer layers and low-density envelope of L1544. This discrepancy could be explained by HCN tracing an even outer layer than HCO^+ .
- The radiative transfer modelling showed that the new collisional rate coefficients for HN^{13}C and DNC are incomplete because they only consider the splitting by ^{14}N and therefore miss a part of the hyperfine structure. This issue is addressed by work underway that includes the effect of ^{13}C and D.
- The radiative transfer modelling of HH211 using the physical structure derived in this work is successful for optically thin emission lines, but shows problems with optically thick lines that also trace more complex structures and are influenced by them.
- The radiative transfer modelling of L1544 applying the Keto-Caselli model works well for moderately optically thin lines that trace inner layers of the core, but shows problems with optically thin lines influenced by CO depletion (HC^{18}O^+) and asymmetric distribution across the core (CCH).
- The radiative transfer modelling of L1544 applying the HDCRT model is successful at reproducing the optically thick hyperfine components of CCH, which show symmetric double-peak profiles, and the red asymmetry observed in the central hyperfine component of HCN (1-0).
- The modelling results of both cores show that to correctly model emission lines, it is crucial to include the outer layers of the cores to consider the effects of extended structures.

Future projects with more comprehensive surveys, including multiple pre-stellar and protostellar cores and covering more molecules, will be able to provide larger statistics and confirm the trends reported in this work. New detailed chemical models including the time-dependent variations in $^{12}\text{C}/^{13}\text{C}$ ratios (Colzi et al. 2020; Sipilä et al. 2023) will aid in further constraining the lines of C-containing molecules, including isotopologues. Additionally, modelling of molecules such as CS or H_2CO will help to further constrain the diffuse envelope surrounding the pre-stellar core L1544.

Chapter 3

Chemical segregation analysed with unsupervised clustering

The content of this chapter was published in the *Astronomy & Astrophysics* journal.
Credit: Giers, K., Spezzano, S., Lin, Y., Valdivia-Mena, M. T., Caselli, P., Sipilä, O. , 2025, A&A, 699, A103.

Abstract

Context. Molecular emission is a powerful tool for studying the physical and chemical structures in cold and dense cores. The distribution and abundance of different molecular species provide information on the chemical composition and physical properties in these cores.

Aims. We study the chemical segregation of three molecules – $\text{c-C}_3\text{H}_2$, CH_3OH , and CH_3CCH – in the two starless cores B68 and L1521E, and the prestellar core L1544.

Methods. We applied the density-based clustering algorithms DBSCAN and HDBSCAN to identify chemical and physical structures within these cores. To enable cross-core comparisons, the clustering input samples were characterised based on their physical environment, discarding the two-dimensional spatial information.

Results. Clustering analysis showed significant chemical differentiation across the cores. The clustering successfully reproduces the known molecular segregation of $\text{c-C}_3\text{H}_2$ and CH_3OH in all three cores. Furthermore, it identifies a segregation between $\text{c-C}_3\text{H}_2$ and CH_3CCH , which is not apparent from the emission maps. Key features driving the clustering are integrated intensity, velocity offset, H_2 column density, and H_2 column density gradient. Different environmental conditions are reflected in the variations in the feature relevance across the cores.

Conclusions. This study shows that density-based clustering provides valuable insights into chemical and physical structures of starless cores. It demonstrates that already small datasets covering only two or three molecules can yield meaningful results. In fact, this new approach revealed similarities in the clustering patterns of CH_3OH and CH_3CCH rel-

ative to $\text{c-C}_3\text{H}_2$, suggesting that $\text{c-C}_3\text{H}_2$ traces more outer layers or lower-density regions than to the other two molecules. This allowed for insight into the CH_3CCH peak in L1544, which appears to trace a landing point of chemically fresh gas that is accreted to the core, highlighting the impact of accretion processes on molecular distributions.

Keywords: astrochemistry – ISM: clouds – ISM: molecules – ISM: abundances – stars: formation

3.1 Introduction

To understand star and planetary-system formation, it is crucial to understand the physics and chemistry in star-forming regions. Particularly important are starless dense cores, as they represent the earliest stages of star formation and set its initial conditions. In starless cores, the chemistry and physics can be studied without the complications caused by protostellar feedback. Molecular line emission is a powerful tool for the study of the structure of these dense cores. The distribution and abundance of different molecular species provide information on their chemical composition and physical properties (e.g. Crapsi et al. 2007; Redaelli et al. 2021; Lin et al. 2022a). In addition, the line emission can help recover the chemical evolution of molecules. This has been shown, for example, with the inheritance of water (Cleeves et al. 2014) and methanol (Drozdovskaya et al. 2021, 2022) in the Solar System, which come from the prestellar phase. Prestellar cores are a subset of starless cores that are gravitationally bound and on the verge of star formation (e.g. Andre et al. 2000; Keto & Caselli 2008). This makes prestellar cores dynamically evolved, with higher central densities and more pronounced temperature and velocity gradients compared to unbound starless cores (e.g. see Crapsi et al. 2007).

A well-studied example is the prestellar core L1544 in the Taurus Molecular Cloud. Its molecular emission maps have led to the understanding of its evolutionary status, which is close to gravitational collapse (Williams et al. 1999; Ohashi et al. 1999), as well as its physical structure (volume density, velocity, and the dust and gas temperature profiles; e.g. Crapsi et al. 2007; Keto & Caselli 2008; Keto et al. 2015; Chacón-Tanarro et al. 2019). Spezzano et al. (2016) observed a striking chemical differentiation between the carbon-bearing molecules $\text{c-C}_3\text{H}_2$ and CH_3OH in L1544, driven by differences in the external illumination onto the core. This chemical segregation has also been observed in other starless cores, linking it to their environments (Spezzano et al. 2020). In Spezzano et al. (2017), the authors identified four molecular families in L1544, classified by the location of their emission peaks (carbon-chain peak, CH_3OH peak, dust emission peak, and HNCO peak). Using principal component analysis, they found correlations between the different families and with the physical properties of the core. A similar analysis of the starless core L1521E by Nagy et al. (2019) also reported chemical differentiation between the $\text{c-C}_3\text{H}_2$, the CH_3OH , and the dust emission peak.

In the big data era of astronomy, statistical methods are essential to analysing and interpreting the vast amounts of observational and simulated data. The rapid advancements

in machine learning provide novel approaches to study the molecular complexity during the early stages of star formation. Unsupervised learning algorithms (clustering techniques in particular) are increasingly applied to identify hidden patterns, structures, and relationships in multidimensional datasets (e.g. see review by Fotopoulou 2024). By grouping together data points based on similarities in various physical and chemical parameters, these methods help visualise subtle trends that might not be detected otherwise.

In astrochemistry, clustering methods have primarily been applied to large-scale surveys, for instance, to identify molecular clouds in an unbiased and systematic way (e.g. Colombo et al. 2015; Bron et al. 2018; Yan et al. 2022). By isolating distinct structures in position-position-velocity space, these methods segment clouds into regions with similar physical or chemical properties. Furthermore, Valdivia-Mena et al. (2023) have connected filament scales (<0.1 pc) with envelope scales (>100 au) by identifying (velocity-) coherent structures of inflowing material, known as streamers, through the clustering of molecular emission. Meanwhile, Okoda et al. (2020, 2021) applied principal component analysis to molecular line emission, characterising velocity structures and distinct features surrounding a protostar. Additionally, this method has also been used to disentangle overlapping kinematic components and to understand spectral variations (e.g. Yun & Lee 2023), providing deeper insight into the dynamics of star-forming regions.

In this work, we apply the unsupervised clustering algorithms Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and Hierarchical DBSCAN (HDBSCAN) to investigate the chemical segregation and differences between the molecular emission of $c\text{-C}_3\text{H}_2$, CH_3OH , and CH_3CCH towards the starless cores B68 and L1521E and the prestellar core L1544. The molecules were chosen as representatives of three of the molecular families found in L1544 (see Spezzano et al. 2017). Our goal is to study the chemical segregation previously observed for $c\text{-C}_3\text{H}_2$ and CH_3OH from a different perspective using clustering techniques. Additionally, we aim to investigate the less understood differentiation between the two carbon chains, $c\text{-C}_3\text{H}_2$ and CH_3CCH . In our analysis, we take a novel approach by dropping the two-dimensional spatial information and instead characterising each pixel of the emission maps based on its physical parameters. By concentrating on the core scale and incorporating both starless and prestellar cores, we study the influence of different evolutionary stages on the molecular emission and distribution as well as the effect of different environments.

In Sect. 3.2, we describe the data and the sources analysed in this work. In Sect. 3.3, we explain the preprocessing we applied to our dataset and what methods we used for the unsupervised clustering with DBSCAN and HDBSCAN. The results of the dataset analysis and the clustering are presented in Sect. 3.4. We discuss the results and their implications in Sect. 3.5 and present our conclusions in Sect. 3.6.

3.2 Observations and data reduction

3.2.1 Data

The data presented in this work were taken with the IRAM 30 m single-dish radio telescope on Pico Veleta in the Sierra Nevada, Spain. The observations were carried out between October 2013 and April 2018 (PIs: Silvia Spezzano, Zofia Nagy). The data have also been used in Spezzano et al. (2017), Nagy et al. (2019), and Spezzano et al. (2020). The on-the-fly (OTF) maps were observed in position switching mode, using the EMIR E090 receiver and the Fourier transform spectrometer (FTS) backend with a spectral resolution of 50 kHz. The map sizes, sources and coordinates are listed in Table 3.1. The observed transitions are summarised in Table 3.2.

The data processing was done using the GILDAS software (Pety 2005) and the python packages `pandas` (Wes McKinney 2010) and `spectral_cube` (Ginsburg et al. 2019). All emission maps were gridded to a pixel size of 8'' with the CLASS software in the GILDAS packages; this corresponds to one-third to one-quarter of the actual beam size, depending on the frequency. To create a uniform dataset, we additionally resampled the data to a spectral resolution of 0.18 km s^{-1} , corresponding to the resolution of the lowest frequency observation (82 GHz). The antenna temperature T_A^* was converted to the main beam temperature T_{mb} using the relation $T_{\text{mb}} = F_{\text{eff}}/B_{\text{eff}} \cdot T_A^*$. The corresponding values for the 30 m forward (F_{eff}) and main-beam efficiencies (B_{eff}) are given in Table 3.2.

Table 3.1: Source sample.

	Right ascension (J2000)	Declination (J2000)	V_{sys}^b (km s^{-1})	Distance ^c (pc)	H ₂ column density ^d (10^{22} cm^{-2})	map size ('' $\times$ '')
B68 (s) ^a	17:22:38.9	-23:49:46.0	3.4	150	1.6	3.3×2.5
L1521E (s)	04:29:15.7	+26:14:05.0	6.9	140	2.3	2.5×2.5
L1544 (p)	05:04:17.2	+25:10:42.8	7.2	170	2.8	2.5×2.5

Notes. ^(a)Starless (s) or prestellar (p), following the definition given in Crapsi et al. (2005).

^(b)B68: Spezzano et al. (2020); L1521E: this work; L1544: Caselli et al. (2002a). ^(c)B68: Alves & Franco (2007). L1521E: Galli et al. (2018). L1544: Galli et al. (2019). ^(d)Values derived from *Herschel*/SPIRE observations towards the dust peak (Spezzano et al. 2016, 2020).

3.2.2 Sources

The source sample is listed in Table 3.1. The pre-stellar core L1544 (following the definition of Crapsi et al. 2005) and the two starless cores (L1521E, B68) are located in different star-forming regions (Taurus, Ophiuchus) and therefore cover different evolutionary stages and environmental conditions.

L1521E and L1544 are located in the Taurus molecular cloud. Both cores are located at the edge of their filament, which exposes their southern sides to the local interstellar radiation field (ISRF). The higher illumination leads to an increase of C atoms in the gas

Table 3.2: Spectroscopic parameters of the observed lines.

Molecule	Transition	Frequency ^a (MHz)	$F_{\text{eff}}/B_{\text{eff}}$	E_{up}^a (K)	g_{up}^a	A^a (s ⁻¹)	Reference
c-C ₃ H ₂	2 _{0,2} – 1 _{1,1}	82093.544(1)	0.95/0.81	6.4	5	1.89×10^{-5}	1
c-C ₃ H ₂	3 _{2,2} – 3 _{1,3}	84727.688(2)	0.95/0.81	16.1	7	1.04×10^{-5}	1
CH ₃ OH	2 _{1,2} – 1 _{1,1} (E_2)	96739.358(2)	0.95/0.80	12.5	20	2.56×10^{-6}	2
CH ₃ OH	2 _{0,2} – 1 _{0,1} (A^+)	96741.371(2)	0.95/0.80	7.0	20	3.41×10^{-6}	2
CH ₃ CCH	5 ₁ – 4 ₁	85455.6667(1)	0.95/0.81	19.5	22	1.95×10^{-6}	3
CH ₃ CCH	5 ₀ – 4 ₀	85457.3003(1)	0.95/0.81	12.3	22	2.03×10^{-6}	3
CH ₃ CCH	6 ₁ – 5 ₁	102546.0242(1)	0.95/0.79	24.5	26	3.46×10^{-6}	3
CH ₃ CCH	6 ₀ – 5 ₀	102547.9844(1)	0.95/0.79	17.2	26	3.56×10^{-6}	3

Notes. ^(a)Extracted from the Cologne Database for Molecular Spectroscopy (Müller et al. 2001). **References.** (1) Thaddeus et al. (1985); (2) Xu & Lovas (1997); (3) Bauer & Burie (1969).

phase and subsequently enhanced abundances of carbon chains such as c-C₃H₂, as discussed in Spezzano et al. (2020). L1521E was classified as a very young core because of its high abundances of carbon-chain molecules and low level of CO depletion (Hirota et al. 2002; Tafalla & Santiago 2004; Nagy et al. 2019). The well-studied prestellar core L1544, on the other hand, is more evolved and shows signs of contraction, suggesting that it is on the verge of star formation (Williams et al. 1999; Ohashi et al. 1999; Lee et al. 2001; Caselli et al. 2002a, 2012). Within the central 1000 au, L1544 exhibits an almost total (99.99%) freeze-out of all species heavier than Helium (Caselli et al. 2022).

The isolated starless core B68 is located in the south of the Ophiuchus molecular cloud. It shows kinematic features of oscillation, indicating a stage prior to contraction (Lada et al. 2003; Keto & Caselli 2008). Due to it being a Bok globule (and therefore isolated), it can be assumed to be exposed to uniform external illumination.

3.3 Methodology

3.3.1 Preprocessing

The integrated intensity maps were generated by calculating the zeroth moment of the emission data cubes. We used the resulting map when the emission is extended over at least one telescope beam, as our focus in this work is primarily on the spatial distribution of the molecules. To facilitate comparisons between the molecules, all maps were convolved to an angular resolution of 32'', corresponding to the half-power beam width (HPBW) of the largest observed beam. The resulting integrated intensity maps of c-C₃H₂, CH₃OH and CH₃CCH are shown in Fig. B.1.

To compare the distribution of molecular emission across three different cores with varying map sizes and distances, we treated the individual pixels of each map as input

samples, and discarded their two-dimensional spatial information. Instead, we described the environment of each pixel using the projected distance to the dust peak, the H_2 column density, and the H_2 column density gradient at that position. Additionally, we used the Gaussian fit parameters – velocity and linewidth – obtained from fitting each pixel’s spectrum with a one-dimensional Gaussian profile. This approach is valid as all cores in our sample display only one velocity component along the line of sight. For the analysis, we selected only the fits with a signal-to-noise ratio greater than or equal to three and an error below 70%.

We used six input features in the clustering dataset. The following features characterise one input sample or emission pixel:

Integrated intensity: The intensity was integrated over a range of $\pm 0.7 \text{ km s}^{-1}$ around the centroid velocity. While the expected linewidth for optically thin emission in our cores is around 0.5 km s^{-1} , we used a wider interval here to account for line shifts due the velocity gradients across the cores. For the clustering, we selected only pixels with a signal-to-noise ratio of at least three. Each integrated intensity map was then normalised using the `MinMaxScaler` from the `scikit-learn preprocessing` package, scaling the values from zero to one before being added to the clustering dataset. This standardisation removes information about the absolute brightness of the molecular emission and the intensity ratios between different molecules, focusing instead on the distribution of emission across the core. In the analysis, this feature is referred to as *intensity*.

Velocity offset with respect to source’s V_{LSR} : Here, we applied a selection criterion where the uncertainty in the fitted velocity position, V_{LSR} , has to be smaller than 0.08 km s^{-1} (which is channel width/2.355). To calculate the relative velocity offset of the emission line, we subtracted the systemic velocity from the velocity position: $V_{\text{offset}} = V_{\text{LSR}} - V_{\text{sys}}$. The systemic velocities of the cores are listed in Table 3.1. In the analysis, this feature is referred to as V_{offset} .

Linewidth: The minimum linewidth was set to be the spectral resolution of the spectra (0.18 km s^{-1}). In the analysis, this feature is referred to as *linewidth*.

H_2 column density: We used the H_2 column density maps derived from *Herschel* SPIRE maps (Spezzano et al. 2016, 2020), which are shown in Fig. B.2. Similar to the integrated intensity maps, each column density map was normalised individually before being added to the clustering dataset. In the analysis, this feature is referred to as N_{H_2} .

Distance to the dust continuum peak: This feature was calculated as the distance between the equatorial coordinates of each pixel and the dust emission peak of the corresponding core. To be consistent across all cores, the dust emission peak was approximated by the emission peak in the H_2 column density map. In the analysis, this feature is referred to as *dist2dust*.

H₂ column density gradient: To calculate this feature, we convolved the H₂ column density maps with a Gaussian derivative kernel using a standard deviation of two telescope beams, as described in Soler et al. (2013). Details of the derivation and the resulting gradients can be found in Sec. 3.4.1 and Appendix B.2, where the gradient maps are presented in Fig. B.2. As with the integrated intensity and H₂ column density maps, each gradient map was normalised individually before being added to the clustering dataset. In the analysis, this feature is referred to as ∇N_{H_2} .

The features V_{offset} , *linewidth*, and *dist2dust* are normalised only when chosen for a specific sub-dataset (see Tables 3.3 and 3.4). This ensures that the feature values of all selected molecules are standardised together rather than individually, as is done for the emission.

3.3.2 Clustering

We used two density-based clustering algorithms in this analysis: DBSCAN (Ester et al. 1996) and HDBSCAN (Campello et al. 2013). For DBSCAN, we used the `scikit-learn` implementation (Pedregosa et al. 2011) and for HDBSCAN the `hdbscan` Python package (McInnes et al. 2017). Density-based clustering separates high-density areas from low-density areas by grouping points that are closer than a given distance threshold, which is defined by the hyperparameter `EPSILON`. The minimum number of points required to form a cluster is set by the hyperparameter `MIN_SAMPLES`; groupings smaller than this are considered as noise. Unlike K-MEANS clustering, which defines clusters as spherical, density-based clustering allows for clusters of arbitrary shape. DBSCAN classifies points into three types: core points, which have at least `MIN_SAMPLES` of neighbouring points within the `EPSILON` distance; border points, which have less than `MIN_SAMPLES` of neighbouring points, but are within `EPSILON` distance to at least one core point; noise points, which have less than `MIN_SAMPLES` of neighbouring points and are not within `EPSILON` distance to any core point, so they are not part of any cluster. HDBSCAN extends DBSCAN by converting it into a hierarchical clustering algorithm, allowing for clusters with varying densities. Clusters are defined by the hyperparameter `MIN_CLUSTER_SIZE`, which sets the minimum number of points required for a group to be considered a cluster, and the hyperparameter `MIN_SAMPLES`, which sets the minimum number of neighbouring points needed for a point to be considered a core point.

We used three features at a time as inputs to both DBSCAN and HDBSCAN, which helped to simplify the output interpretation. The integrated intensity is kept as a fixed input feature to retain the information about the molecular emission, while the other two features are varied. The ten resulting feature combinations are listed in Table 3.3. To optimise the clustering, we perform a systematic grid search for the respective hyperparameters: `EPSILON` (ranging from 0.05 to 0.155, in steps of 0.005) and `MINSAMPLES` (ranging from 10 to 100, in steps of 5) for DBSCAN; and `MINCLUSTERSIZE` (ranging from 5 to 20) and `MINSAMPLES` (ranging from 1 to `MINCLUSTERSIZE`) for HDBSCAN.

To determine the best tuning, we use the density-based clustering validation (DBCV)

Table 3.3: Feature combinations used in the clustering.

Combination	Feature 1	Feature 2	Feature 3
1	<i>intensity</i>	V_{offset}	<i>linewidth</i>
2	<i>intensity</i>	V_{offset}	<i>dist2dust</i>
3	<i>intensity</i>	V_{offset}	N_{H_2}
4	<i>intensity</i>	V_{offset}	∇N_{H_2}
5	<i>intensity</i>	<i>linewidth</i>	<i>dist2dust</i>
6	<i>intensity</i>	<i>linewidth</i>	N_{H_2}
7	<i>intensity</i>	<i>linewidth</i>	∇N_{H_2}
8	<i>intensity</i>	<i>dist2dust</i>	∇N_{H_2}
9	<i>intensity</i>	<i>dist2dust</i>	N_{H_2}
10	<i>intensity</i>	N_{H_2}	∇N_{H_2}

score (Moulavi et al. 2014). This score is particularly well-suited to validating density-based clustering methods because it accounts for outliers and noise points (unlike cross-validation, for example). To calculate the score, a kernel density function first estimates the local density of data points around each object. Then, the cluster quality is evaluated by comparing the minimum density within clusters (which represents cluster cohesion) to the maximum density between clusters (which represents cluster separation). We use the HDBSCAN function `validity_index`, which is a fast approximation of the original DBCV score. Although this function is provided by the `hdbscan` package, it only requires the data points and the corresponding cluster labels as input, and therefore we also applied it to the DBSCAN results. For HDBSCAN, we additionally used the intrinsic attribute `relative_validity`, which is another fast approximation of the DBCV score. Since this intrinsic score gives slightly different results from the `validity_index` function, we consider both scores in our evaluation. Finally, we select the best result based on the number of noise points and the relative sizes of the clusters. However, since both the `validity_index` function and the `relative_validity` attribute are only approximations of the DBCV score, we treat them as relative measures and use them only to compare results across different hyperparameters choices, not across different datasets.

In addition to the DBCV score, we applied two postprocessing criteria to determine the optimal clustering results for each method, feature combination, and dataset: (1) the number of clusters found by the algorithm has to be between two and five (inclusive), and (2) the found clusters must collectively cover at least 50% of the data points. Tests with more than five clusters showed that in this case, larger clusters are merely subdivided into smaller ones covering the same total amount of data and provide no additional insights. Conversely, when only one cluster is found, it typically contains over 90% of the input data points, offering little value to understanding the data distribution. Most datasets show two to four prominent trends that are effectively captured by the clustering. When the resulting clusters cover less than 50% of the data, they fail to represent the overall trends

and instead highlight minor sub-clusters while overlooking the majority of the data.

3.3.3 Case studies

In addition to $c\text{-C}_3\text{H}_2$, CH_3OH , and CH_3CCH , the dataset used for this study covers molecules such as CCS , HC_3N , HC^{18}O^+ , C_4H , HNCO , and CS , with between 8 to 20 detected molecules per core. However, to enhance the interpretability and extract meaningful patterns, we focused our analysis on the three key molecules $c\text{-C}_3\text{H}_2$, CH_3OH , and CH_3CCH . This targeted approach allowed for a clear understanding of the relationships between these molecules, and helped to maintain clarity in the complex clustering output. During the clustering process, the features of the molecular maps were combined without giving the algorithm any prior information about which data point corresponds to each molecule.

We used the following four datasets as input to the clustering:

Case 1: $c\text{-C}_3\text{H}_2$ vs CH_3OH

The two molecules show a well-known and well-studied segregation in the sources of our dataset (Spezzano et al. 2016, 2020), which is driven by environmental effects. Through clustering, we analyse how this chemical differentiation is represented in the six features and how it influences the clusters identified by the algorithm. This approach allowed us to assess the effectiveness of the clustering technique. The transitions used in each core are listed in Table 3.4, along with the initial ratio between the data points of each molecule in this sub-dataset.

Case 2: $c\text{-C}_3\text{H}_2$ vs CH_3CCH

These two molecules display a chemical differentiation in the prestellar core L1544 that is not yet understood (Spezzano et al. 2017). While CH_3CCH is a carbon chain such as $c\text{-C}_3\text{H}_2$ and is therefore expected to peak in the carbon-chain rich south-east of the core, similar to $c\text{-C}_3\text{H}_2$, it instead peaks in the north-west of the core. To ensure a balanced number of data points between the two molecules, we use two transitions of CH_3CCH for B68, and three transitions for L1521E, as shown in Table 3.4.

Case 3: CH_3OH vs CH_3CCH

The two molecules show a chemical differentiation in the prestellar core L1544 (Spezzano et al. 2017). Both peak in the northern part of the core, CH_3OH in the north-east, and CH_3CCH in the north-west. We use this combination to rule out biases that might arise from a clustering with $c\text{-C}_3\text{H}_2$.

Case 4: $c\text{-C}_3\text{H}_2$ vs CH_3OH vs CH_3CCH

We use the combination of all three molecules to validate the results from the other two case studies. This approach also helps to eliminate biases in the algorithm that might arise

Table 3.4: Molecular transitions used in Case 1, Case 2, and Case 3. For Case 4, the transitions from Case 1 and Case 2 are combined for each core. Additionally, the ratio of data points for c-C₃H₂ to CH₃OH (Case 1), c-C₃H₂ to CH₃CCH (Case 2), and CH₃OH to CH₃CCH (Case 3) in each dataset is given as percentages.

Case 1			
	c-C ₃ H ₂	CH ₃ OH	ratio
B68	2 ₀₂ – 1 ₁₁	2 ₀₂ – 1 ₀₁	50/50
L1521E	2 ₀₂ – 1 ₁₁	2 ₀₂ – 1 ₀₁	45/55
L1544	3 _{2,2} – 3 _{1,3}	2 ₀₂ – 1 ₀₁	42/58
Case 2			
	c-C ₃ H ₂	CH ₃ CCH	ratio
B68	2 ₀₂ – 1 ₁₁	5 ₀ – 4 ₀ , 5 ₁ – 4 ₁	57/43
L1521E	2 ₀₂ – 1 ₁₁	5 ₀ – 4 ₀ , 5 ₁ – 4 ₁ , 6 ₀ – 5 ₀	45/55
L1544	3 _{2,2} – 3 _{1,3}	6 ₁ – 5 ₁	46/54
Case 3			
	CH ₃ OH	CH ₃ CCH	ratio
B68	2 ₀₂ – 1 ₀₁	5 ₀ – 4 ₀ , 5 ₁ – 4 ₁	57/43
L1521E	2 ₀₂ – 1 ₀₁	5 ₀ – 4 ₀ , 5 ₁ – 4 ₁ , 6 ₀ – 5 ₀	50/50
L1544	2 ₀₂ – 1 ₀₁	6 ₁ – 5 ₁	54/46

from comparing only two molecules. The dataset for each core contains the combined data of Case 1 and Case 2 (see Table 3.4).

3.4 Results

3.4.1 H₂ column density gradient

To derive the H₂ column density gradients for our three cores, we apply the method presented in Soler et al. (2013). Therefore, the gradient is calculated by convolving the H₂ column density maps with a Gaussian derivative kernel in the x and y direction. To derive the total gradient, we combine the two directions. We use the method `gaussian_filter` from the Python package `scipy.ndimage`. To depict the filament environments of the cores, we choose a Gaussian kernel with standard deviation equivalent to two telescope beams (=2x 32" or 2x 4 pixels).

The derived H₂ column density gradient maps are presented in Fig. B.2 in the Appendix, alongside the corresponding H₂ column density maps. The observed gradients agree with the different levels of exposure to the ISRF of the different cores (see Sec. 3.2.2). The isolated starless core B68 has a mostly uniform N(H₂) gradient, forming a ring-like structure

at the edges and a $\nabla N_{\text{H}_2} \approx 0$ in the centre, representing a uniform external illumination. For L1521E, the larger illumination towards the south is represented by an increased $N(\text{H}_2)$ gradient along the south and lower values in the protected centre. Similar to L1521E, the $N(\text{H}_2)$ gradient of the prestellar core L1544 depicts how the south of the core is more exposed to the ISRF, resulting in larger values of ∇N_{H_2} . In contrast, the $N(\text{H}_2)$ gradient is much lower in the more protected centre and north of the core, where both CH_3OH and CH_3CCH peak.

3.4.2 Comparison of dataset features

Figure 3.1 shows a comparison of selected features of the dataset, illustrating the data points for $\text{c-C}_3\text{H}_2$ (green circles), CH_3OH (blue crosses), and CH_3CCH (red diamonds) observed towards B68 (left), L1521E (centre), and L1544 (right). The plots reveal distinct patterns and behaviours that vary depending on the core, molecule, and feature.

The distributions of *intensity* over N_{H_2} and *intensity* over V_{offset} , shown in the top two rows of Fig. 3.1, reflect the distribution of the molecular emission across the cores. In B68, all molecules exhibit similar distributions, with peak *intensity* at the highest H_2 column density. In L1521E, $\text{c-C}_3\text{H}_2$ and CH_3CCH peak in the south-eastern part of the core at lower N_{H_2} compared to CH_3OH , which peaks at the dust peak. In contrast, in L1544, all molecules peak in different locations across the core with varying projected distances to the dust peak. Velocity-wise, CH_3OH in L1521E shows a slightly different behaviour compared to the carbon chains: at high *intensity* it extends to higher V_{offset} , while at lower *intensity* it spreads to lower V_{offset} . Additionally, in L1544, CH_3CCH spans a broader velocity range than $\text{c-C}_3\text{H}_2$ and CH_3OH , suggesting that it traces a different layer of the core.

The $\text{c-C}_3\text{H}_2$ emission shows a wide range of linewidths, which are in general broader and reach higher values compared to the other molecules (see third row in Fig. 3.1), likely tracing more turbulent material (compare e.g. Lin et al. 2022a for L1544). In L1521E, the linewidths of $\text{c-C}_3\text{H}_2$ have a more compact distribution around a value of 0.4 km s^{-1} . For CH_3OH , the linewidths are typically smaller in B68 and L1544, between $0.25\text{--}0.30 \text{ km s}^{-1}$ and $0.30\text{--}0.35 \text{ km s}^{-1}$, respectively, while in L1521E, they are slightly higher, around 0.4 km s^{-1} , with some values reaching up to 0.65 km s^{-1} . The linewidths of CH_3CCH are more compactly distributed at lower values in the two starless cores (averaging around 0.3 km s^{-1}), but extend up to 0.5 km s^{-1} in the prestellar core L1544, possibly indicating the presence of more turbulent material.

The ∇N_{H_2} -*intensity* distribution also reflects the morphology of the molecular emission across the cores (see bottom row in Fig. 3.1). In L1521E, all molecules display a similar behaviour, with the highest intensities occurring at high, though not maximum, ∇N_{H_2} values. In contrast, in L1544, the $\text{c-C}_3\text{H}_2$ *intensity* peaks at a much higher ∇N_{H_2} than CH_3OH and CH_3CCH . This illustrates the active photochemistry in the south of the core caused by the external illumination of the core, leading to an increased abundance of carbon-chains such as $\text{c-C}_3\text{H}_2$. This is supported by a local peak in CH_3CCH emission in the south and a sharp decline in CH_3OH *intensity* at a higher ∇N_{H_2} . In B68, which has a more spherical shape and is uniformly illuminated, the molecular emissions peak at lower

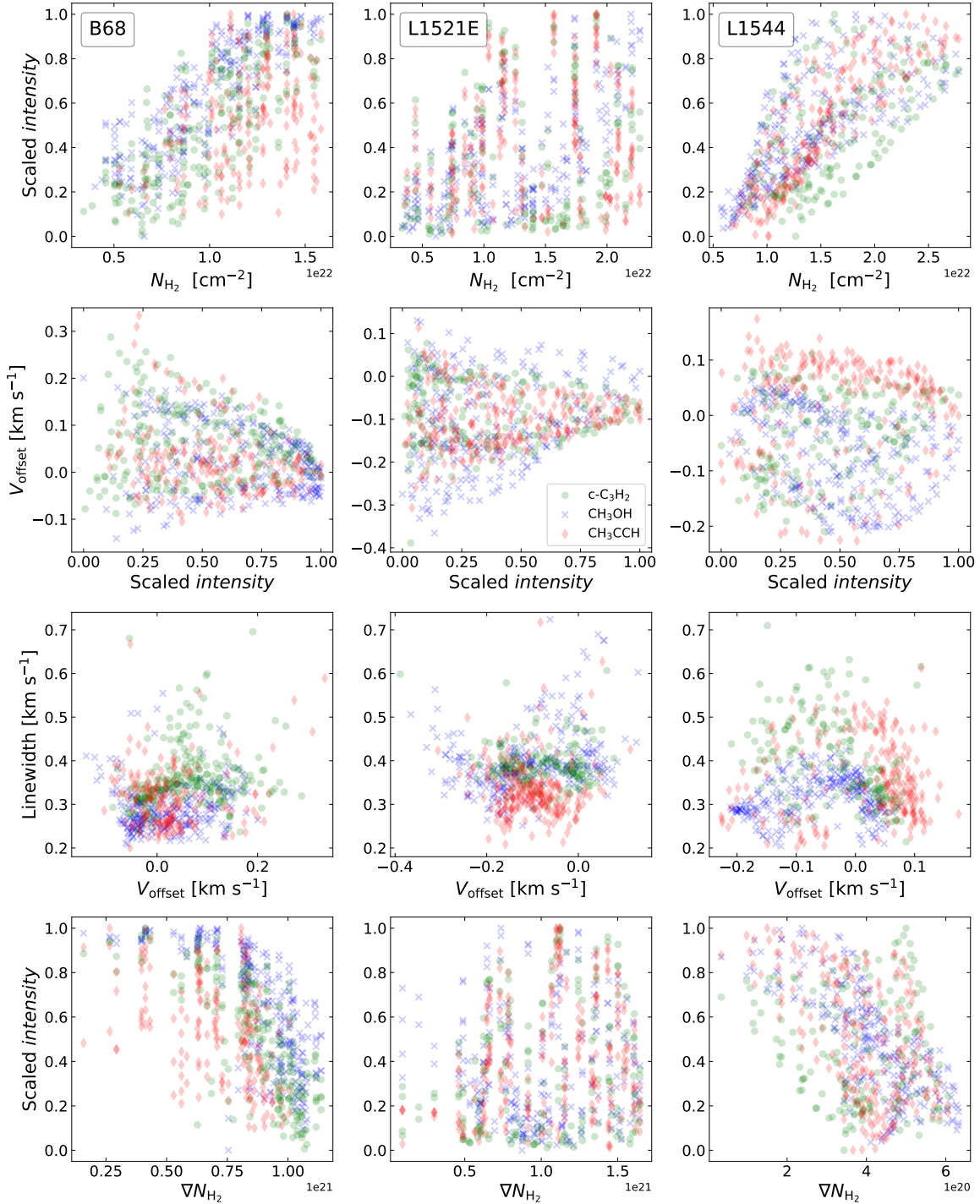


Figure 3.1: Comparison of different features for B68 (*left*), L1521E (*middle*), and L1544 (*right*), with c-C₃H₂ given in green (circle), CH₃OH in blue (cross), and CH₃CCH in red (diamond).

∇N_{H_2} values in the protected centre of the core. The emission then shows a rather steep decline at a higher ∇N_{H_2} in the outer parts of the core. Additionally, CH_3CCH exhibits lower *intensity* values even at a lower ∇N_{H_2} , because the emission map is less spatially extended across this core compared to $\text{c-C}_3\text{H}_2$ and CH_3OH (see Fig. B.1).

In summary, the plots point out the wealth of chemical and physical differences between the cores. In L1544, all molecules are clearly separated and behave differently, which indicates a more profound chemical segregation in this core. In contrast, in B68 and L1521E, the molecules show a more similar behaviour and are less segregated, which could be linked to their earlier evolutionary stages compared to L1544. Consequently, an unbiased approach such as clustering can provide valuable insights into the varying chemical environments across these cores.

3.4.3 Clustering

In Figs. 3.2 - 3.3, and B.3 - B.6, we show the clustering results for B68, and for L1521E and L1544, respectively. They present the results for the feature combinations 2, 3, 4, 9, and 10 for Case 1 and Case 2 (for details see Table 3.3). The remaining results for Cases 1 and 2, together with the results for Cases 3 and 4 are published on Zenodo¹ as Figs. C.3 - C.14. The molecular ratio (i.e. the ratio of data points belonging to each of the molecules in a specific cluster) and the number of data points assigned to a cluster are given in Tables B.1 - B.3. In the following, we describe the results for each case individually.

Both the DBSCAN and the HDBSCAN results are included in the analysis. However, to improve readability, only one result is represented for each combination. This was decided manually for each combination based on the amount of noise points and the number of clusters. This ensures that all scientific results discussed in this work are also presented visually.

To evaluate the possible chemical segregation within the clusters, we focused on imbalanced clusters, where the molecular ratio deviates from the initial ratio by at least 10% (e.g. 37/63 instead of 47/53). These imbalanced clusters are particularly insightful because they highlight regions where specific molecular abundances diverge from the average distribution, and they potentially indicate distinct physical or chemical processes at work. The imbalanced clusters are marked in boldface in Tables B.1 - B.3. Our primary focus is to extract meaningful chemical and scientific insights from the clustering patterns. In fact, many clusters exhibit an excess of, or are dominated by, one molecule, indicating chemical differentiation across all three cases and all three cores in this study. However, clusters that exclusively contain data points from a single molecule are rare and typically contain only a small number of points ($N \leq 20$). Overall, we prioritised clusters that cover at least one telescope beam, which corresponds to a size of roughly 20 data points or pixels.

¹zenodo.org/records/15519030 – Giers et al. 2025: Appendix C.

Case 1: c-C₃H₂ vs CH₃OH

In some results, the clusters vary greatly in size. The largest cluster typically maintains a balanced molecular ratio, similar to the input ratio (see Table 3.4), while the smaller clusters tend to show more variation (see Tables B.1 - B.3). For L1521E and L1544, the input data has a slightly higher proportion of CH₃OH compared to c-C₃H₂ (see Table 3.4), causing the largest cluster to often show a slight excess in CH₃OH. In the following section, we discuss the clustering results for each core individually.

B68: Imbalanced clusters with an excess of c-C₃H₂ or CH₃OH are spatially separated into different regions of the core. CH₃OH-dominated clusters are concentrated around the dust peak at the centre of the core, characterised by features such as high *intensity* and high N_{H_2} (e.g. red clusters in combs. 2, 4, and 7, or yellow cluster in comb. 3; see Figs. 3.2 and C.3, and Table B.1). In contrast, c-C₃H₂-dominated clusters are confined to the (south-) west region and associated with data points of moderate *intensity* (see yellow and purple clusters in comb. 2 in Fig. 3.2). In combination 9 and 10, N_{H_2} and ∇N_{H_2} are structured in ring-like clusters around the dust peak (see Fig. 3.2), all dominated by CH₃OH, and with no significant contribution from c-C₃H₂. Interestingly, in B68, the largest cluster of each combination (shown in blue) is slightly imbalanced towards CH₃OH (on average 3%, see Table B.1), even though the initial molecular ratio is 50/50. The smaller clusters, however, display greater variation.

L1521E: In this core, many features are clustered into separate structures dominated by either CH₃OH or c-C₃H₂. This molecular segregation is evident across multiple combinations: The *intensity-dist2dust* distribution is clustered into two separate diagonals (see blue and red clusters in combs. 2 and 5 in Figs. B.3 and C.4). The lower *intensity* diagonal (blue) corresponds to regions in the north-west of the core and is imbalanced towards c-C₃H₂ (see Table B.2). Conversely, the upper *intensity* diagonal (red) is associated with the south-east of the core and is dominated by CH₃OH. The same cluster distribution, with similar molecular imbalance, appears in the *intensity-N_{H2}* and the *intensity-V_{offset}* distributions: clusters dominated by c-C₃H₂ (in the north-west) show high N_{H_2} and high V_{offset} (see blue cluster in comb. 2, red cluster in comb. 3), while CH₃OH-dominated clusters (south-east) show low N_{H_2} and lower V_{offset} (see red clusters in comb. 2 and 9, and blue cluster in comb. 3). Additionally, combination 2 (Fig. B.3) shows that the CH₃OH-dominated clusters appear as a narrow diagonal structure in *dist2dust* over V_{offset} , while the c-C₃H₂ clusters present a separate, more diffuse distribution. To summarise, c-C₃H₂ and CH₃OH show a separation in N_{H_2} and V_{offset} in L1521E, and cluster into separate diagonals in the *intensity-dist2dust* distribution. Beyond the north-west/south-east separation, the $V_{\text{offset}}-\nabla N_{\text{H}_2}$ distribution is split into c-C₃H₂ at high ∇N_{H_2} and mid V_{offset} in the south of the core, and CH₃OH at mid ∇N_{H_2} and lower V_{offset} , in the east (c-C₃H₂: blue cluster in comb. 4; CH₃OH: red cluster in comb. 4; see Figs. B.3). Additionally, CH₃OH shows a cluster at the core centre, characterised by high *intensity* and high N_{H_2} (see yellow cluster in comb. 9 in Fig. B.3).

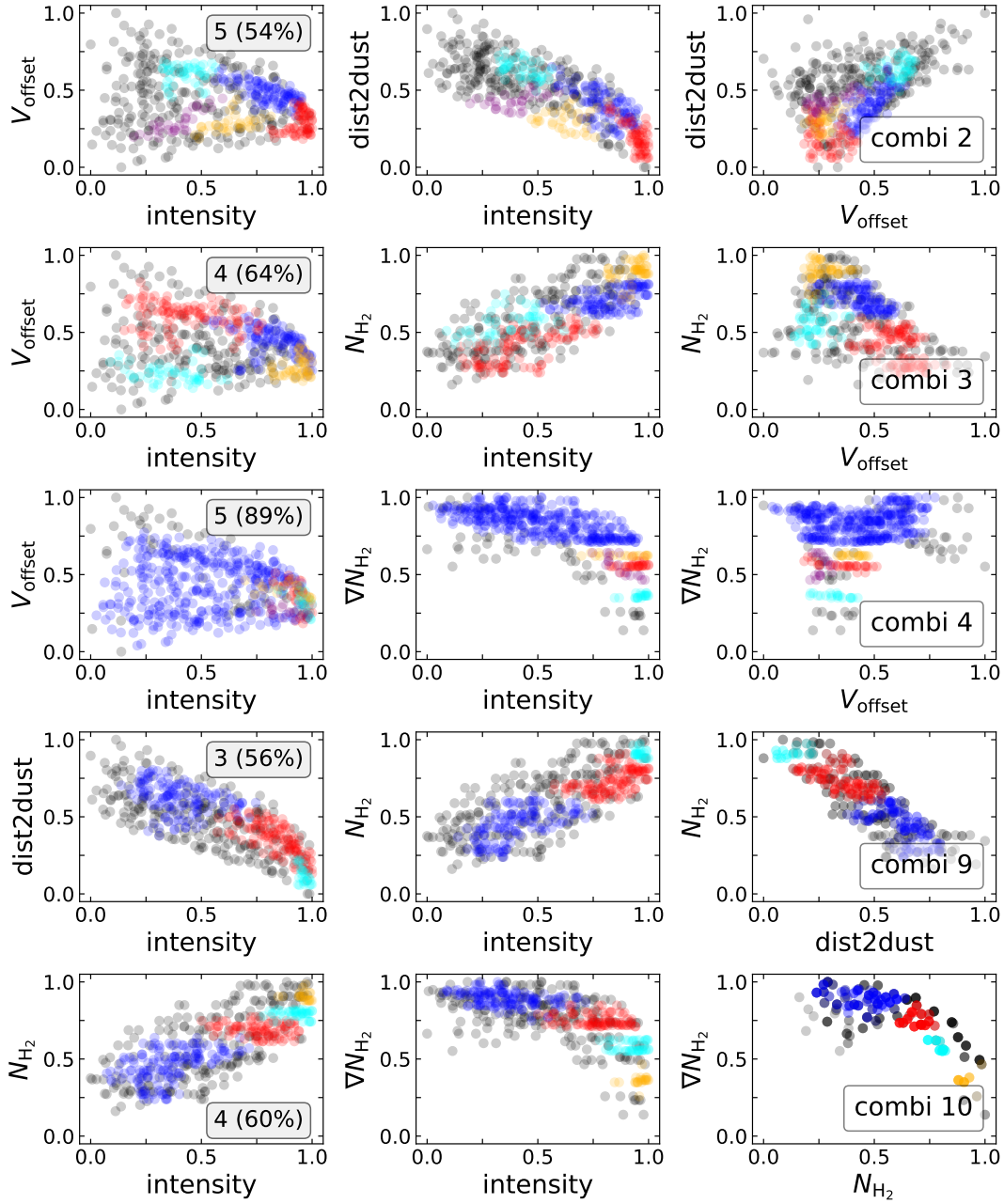
B68: $\text{c-C}_3\text{H}_2$ vs CH_3OH (Case 1)

Figure 3.2: Clustering results for the starless core B68 for the dataset of Case 1 for feature combinations 2, 3, 4, 9, and 10. Each row represents a different combination of features (see Table 3.3). The annotations provide information on how many clusters were found by the algorithm (two to five) and what percentage of data points are assigned to the clusters. Noise points (=points not assigned to any cluster) are plotted in black. The colours of the clusters are ordered by cluster size: the biggest cluster is given in blue, followed by red, cyan, yellow, and purple.

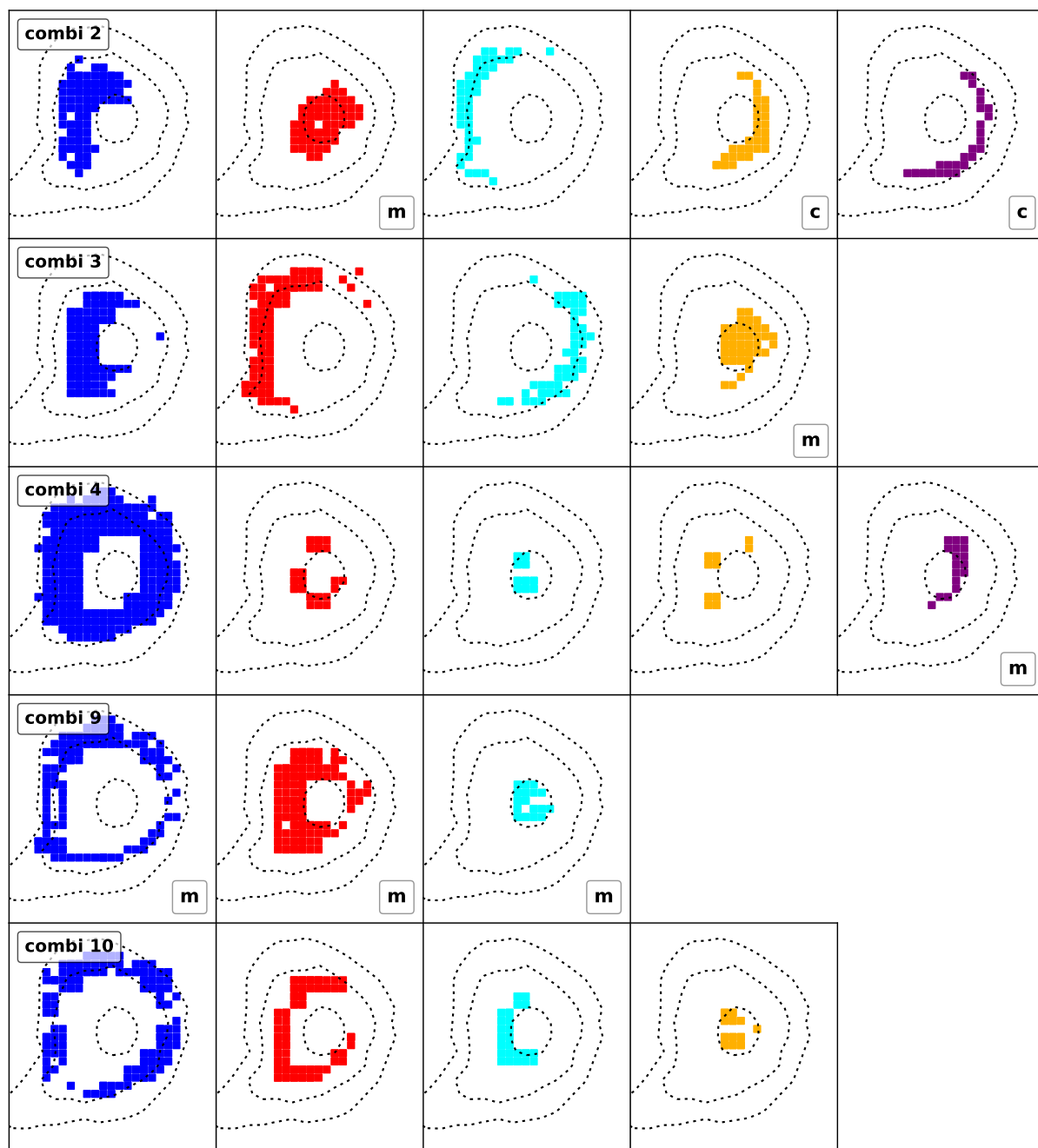


Figure 3.2: continued. Corresponding spatial distribution of each cluster across the starless core B68. The annotations indicate if a cluster contains more than 60% of one molecule (c: $\text{c-C}_3\text{H}_2$; m: CH_3OH ; p: CH_3CCH). The dashed line contours represent 30%, 50%, 90% of the H_2 column density peak derived from *Herschel* maps Spezzano et al. (2020).

L1544: Similar to L1521E, many features in L1544 are clustered into separate structures (V_{offset} over *intensity*, *dist2dust* over *intensity*, N_{H_2} over V_{offset} , *dist2dust* over V_{offset} , *linewidth* over V_{offset} , ∇N_{H_2} over V_{offset} , N_{H_2} over *dist2dust*), dividing the core into north/south, and on-centre/off-centre regions. However, unlike in L1521E, these structures do not consistently correspond to a specific molecular segregation. A separation of the two molecules is visible in V_{offset} and ∇N_{H_2} in combination 4, where a c-C₃H₂-dominated cluster (red) is concentrated in the northern part of the core and a CH₃OH-dominated cluster (blue) is found in the south, both with lower *intensity* (see Fig. B.5). Additionally, c-C₃H₂ clusters appear at the core centre with high *intensity*, high *linewidth*, high N_{H_2} and low ∇N_{H_2} (see cyan cluster in comb. 3 in Fig. B.5, and red clusters in combs. 7 and 10 in Figs. B.5 and C.5). CH₃OH, on the other hand, forms a cluster on the CH₃CCH peak in the north-west, characterised by low V_{offset} and higher *intensity* (see red cluster in comb. 2 in Fig. B.5). It also appears across the northern part of the core with lower *intensity* and low V_{offset} (red cluster in comb. 3, see Fig. B.5). In combination 10, c-C₃H₂ (red) is clustered on the dust peak, while CH₃OH (blue) is off-peak, with separation visible in N_{H_2} , and ∇N_{H_2} (see Fig. B.5). A similar split appears in N_{H_2} in combination 6 (see Fig. C.5), though both corresponding clusters (blue and red) are imbalanced towards CH₃OH without a molecular separation.

Summary of Case 1: All cores display cluster structures with a clear separation between c-C₃H₂ and CH₃OH. However, the molecules are not necessarily clustered on their respective peaks. The molecular separation is primarily visible in the features *intensity*, V_{offset} , and N_{H_2} , and for L1521E and L1544 also in ∇N_{H_2} . In general, the clustering reveals recurring structures in several feature combinations, where one molecule dominates over the other. Additionally, B68 and L1544 show structures in N_{H_2} and ∇N_{H_2} that divide the core into on-centre and around-centre. These divisions, however, are not always linked to a molecular separation.

Case 2: c-C₃H₂ vs CH₃CCH

Similar to Case 1, c-C₃H₂ is slightly underrepresented in the datasets for L1521E and L1544, resulting in clusters that are more imbalanced towards CH₃CCH. For B68, the opposite is true, resulting in a slight excess of c-C₃H₂ in many clusters. In the following, we discuss the clustering results for each core individually:

B68: The clusters dominated by either c-C₃H₂ or CH₃CCH are spatially separated, showing behaviour similar to the segregation of c-C₃H₂ and CH₃OH in Case 1. High *intensity* CH₃CCH is clustered at the core centre, where the molecule peaks (see red and cyan clusters in comb. 4, red cluster in comb. 8, and cyan and yellow clusters in comb. 10 in Figs. 3.3 and C.6). In contrast, c-C₃H₂ is clustered off its emission peak, along the east side of the core, with lower *intensity* (see red clusters in combs. 1,3, and 9 in Fig. 3.3 and C.6). Beyond *intensity*, the molecular separation is also evident in V_{offset} and N_{H_2} : c-C₃H₂ is associated with high V_{offset} and low N_{H_2} , while CH₃CCH is associated with low V_{offset}

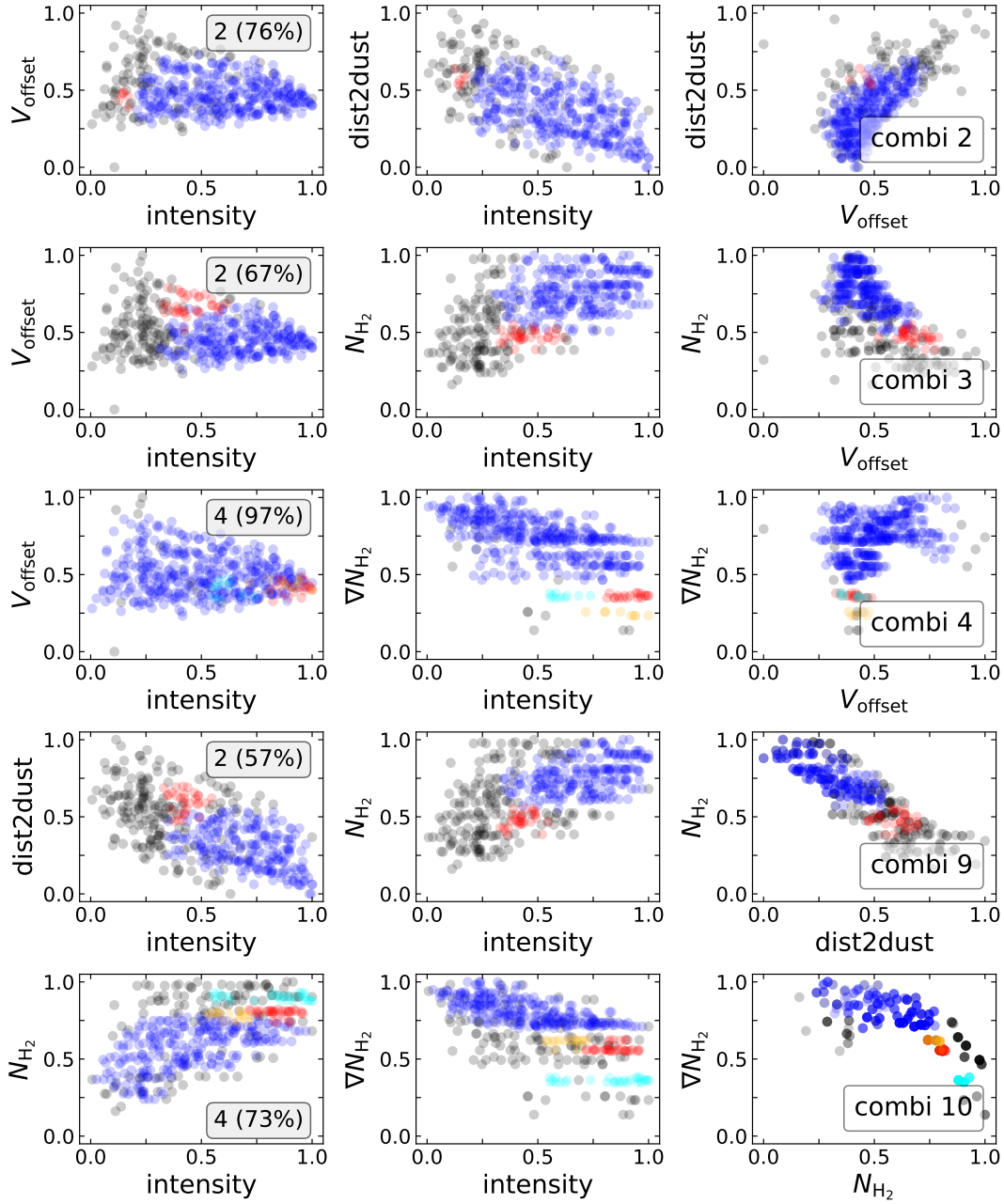
B68: $c\text{-C}_3\text{H}_2$ vs CH_3CCH (Case 2)

Figure 3.3: Clustering results for the starless core B68 for the dataset of Case 2 for feature combinations 2, 3, 4, 9, and 10. Each row represents a different combination of features (see Table 3.3). The annotations provide information on how many clusters were found by the algorithm (two to five) and what percentage of data points are assigned to the clusters. Noise points (=points not assigned to any cluster) are plotted in black. The colours of the clusters are ordered by cluster size: the biggest cluster is given in blue, followed by red, cyan, yellow, and purple.

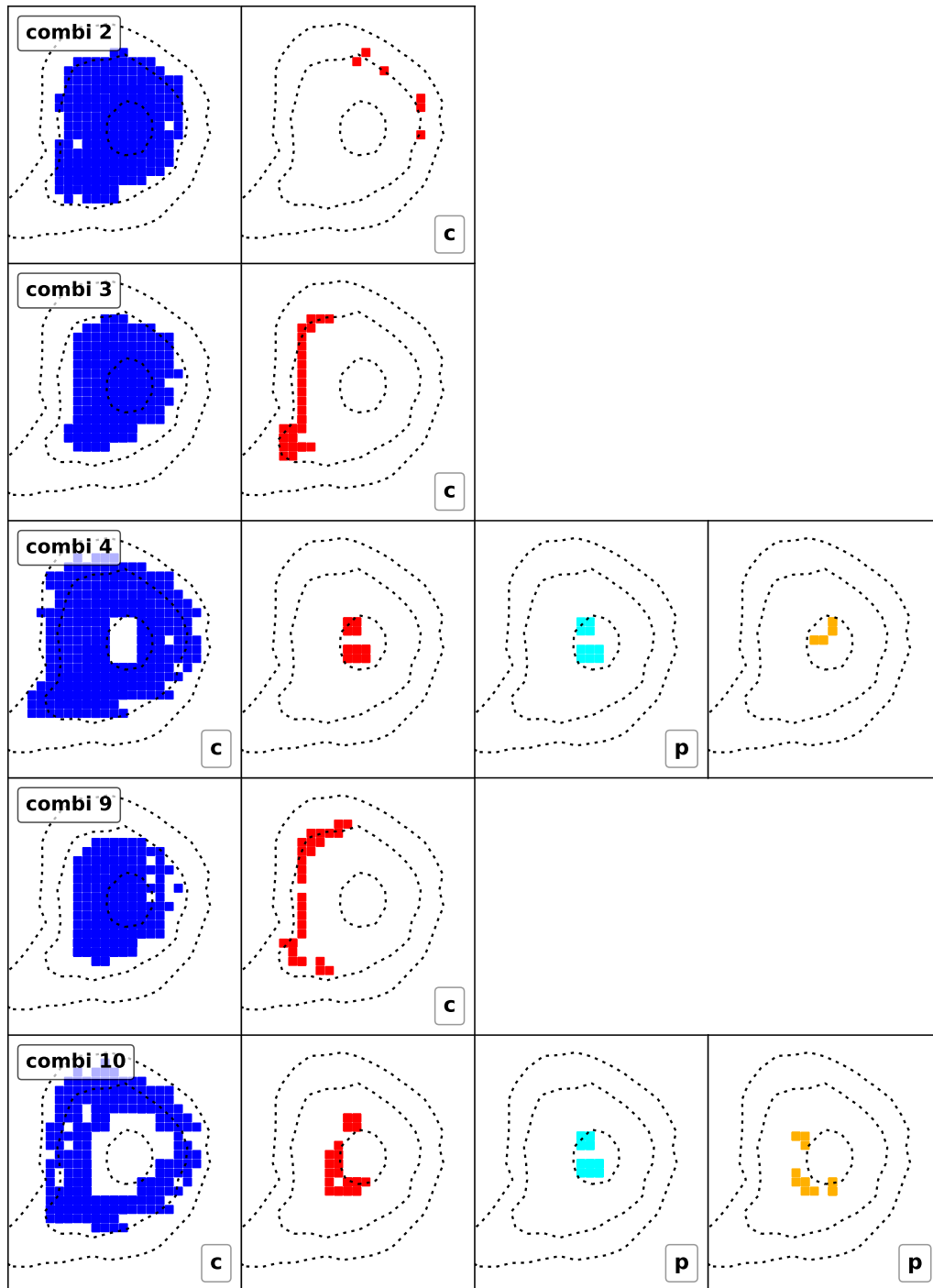


Figure 3.3: continued. Corresponding spatial distribution of each cluster across the starless core B68. The annotations indicate if a cluster contains more than 60% of one molecule (c: c-C₃H₂; m: CH₃OH; p: CH₃CCH). The dashed line contours represent 30%, 50%, 90% of the H₂ column density peak derived from *Herschel* maps Spezzano et al. (2020).

and high N_{H_2} . Interesting to note is that this c-C₃H₂ cluster is shaped along high values of ∇N_{H_2} (see Fig. B.2), which is not included in the mentioned feature combinations (1, 3, 9). In combination 10, a c-C₃H₂-dominated cluster forms a broad ring (blue), corresponding to high values of ∇N_{H_2} (see also Fig. B.2), surrounding CH₃CCH-dominated clusters in the core centre (cyan and yellow). This separation is visible in ∇N_{H_2} and N_{H_2} (see Fig. 3.2).

L1521E: Similar to Case 1, clusters dominated by c-C₃H₂ and CH₃CCH are spatially separated, forming separate structures in various features (e.g. V_{offset} over *intensity*, *dist2dust* over *intensity*, *dist2dust* over V_{offset} , N_{H_2} over *intensity*, N_{H_2} over *linewidth*). As in Case 1, c-C₃H₂ is clustered in the north and north-west, characterised by high N_{H_2} , high V_{offset} , and appearing as lower diagonal in the *intensity-dist2dust* distribution (see red clusters in combs. 2, 3, and 6, and blue clusters in combs. 5 and 9 in Figs. B.4 and C.7). In contrast, CH₃CCH-dominated clusters are found at low N_{H_2} and low V_{offset} , similar to CH₃OH in Case 1, building the upper diagonal in the *intensity-dist2dust* and located in the south and south-east of the core (see blue clusters in combs. 2, 3, and 9, and red cluster in comb. 5 in Figs. B.4 and C.7). CH₃CCH-dominated clusters are concentrated around the carbon peak in the south-east and are associated with higher N_{H_2} (see cyan and yellow clusters in comb. 4, cyan cluster in comb. 6, and cyan and purple clusters in comb. 10 in Figs. B.4 and C.7). Additionally, a low *intensity* c-C₃H₂ cluster is located along the sharp edge of the core in the south-west (see cyan cluster in comb. 9 in Fig. B.4). Similar to B68, the shape of this cluster follows values of high ∇N_{H_2} , even though this feature is not included in combination 9.

L1544: For c-C₃H₂ and CH₃CCH, the clustered features reveal structures similar to those found in Case 1 (e.g. in V_{offset} over *intensity*, *linewidth* over V_{offset} , N_{H_2} over V_{offset} , ∇N_{H_2} over V_{offset} , N_{H_2} over *linewidth*), which again divide the core into north/south and on-centre/off-centre regions. As in Case 1, c-C₃H₂ is predominantly associated with the northern part of the core, while CH₃CCH is dominant in the south. This separation is visible in V_{offset} and ∇N_{H_2} (c-C₃H₂: red clusters in combs. 1, 2, and 4; CH₃CCH: blue clusters in combs. 1 and 4; see Figs. B.6 and C.8). In addition, both molecules cluster in the core centre, exhibiting high *intensity*, higher V_{offset} and higher *linewidth* (c-C₃H₂: cyan cluster in comb. 1; CH₃CCH: yellow cluster in comb. 1, and red cluster in comb. 6; see Figs. C.8). In combination 3, a CH₃CCH-dominated cluster also covers the c-C₃H₂ peak in the south-east of the core, characterised by high *intensity*, high N_{H_2} , and high V_{offset} (see red cluster in Fig. B.6).

Summary of Case 2: Clusters dominated by c-C₃H₂ or CH₃CCH show spatial separation across all cores, similar to the segregation observed between c-C₃H₂ and CH₃OH in Case 1, revealing distinct structures in various features combinations. The molecular segregation is evident in the same features as in Case 1, *intensity*, V_{offset} , and N_{H_2} , with B68 and L1544 also showing separation in ∇N_{H_2} . Overall, the clustering shows the same strong divisions of the cores into north/south, east/west, on-centre/off-centre regions as

seen in Case 1, highlighting structural and chemical themes across the cores.

Case 3: CH₃OH vs CH₃CCH

For B68, CH₃CCH is slightly underrepresented in this dataset, resulting in clusters that are more imbalanced towards CH₃OH. In the following, we discuss the clustering results for each core individually.

B68 (Fig. C.9): CH₃OH-dominated clusters and CH₃CCH-dominated clusters are not clearly spatially separated. Instead, both are found in the central area of the core and on the core centre (CH₃OH: blue cluster in combs. 1, 2, 6, 9, cyan cluster in comb. 6; CH₃CCH: red cluster in comb. 1 and cyan cluster in combs. 8 and 10), reproducing the clustering behaviour of Case 1 for CH₃OH and Case 2 for CH₃CCH. A direct, feature-wise separation of the two molecules occurs only in combination 1 in *intensity*, with CH₃OH at higher and CH₃CCH at lower values. Apart from that, CH₃OH is clustered along the east of the core with high V_{offset} , and along the west of the core with low V_{offset} (east: cyan cluster in comb. 1, red cluster in combs. 2 and 3; west: yellow cluster in comb. 4). The shapes of those clusters follow high values of ∇N_{H_2} , even though this feature is only included in combination 4, resembling patterns of c-C₃H₂ in Case 2 (east) and Case 1 (west). In combinations 8 and 10, the clustering creates ring-like structures around the dust peak, visible in N_{H_2} , *dist2dust*, and ∇N_{H_2} , with CH₃CCH concentrated at the centre and CH₃OH forming the outer rings - reflecting structures found in Case 1 (CH₃OH) and Case 2 (CH₃CCH).

L1521E (Fig. C.10): As in Cases 1 and 2, we see a spatial separation of the two input molecules, CH₃OH and CH₃CCH, but here it is less pronounced. CH₃CCH is primarily clustered in the (south-) east of the core, similar to Case 2 (e.g. see blue and red clusters in comb. 1, cyan cluster in combs. 4, 5 and 10, yellow cluster in combs. 7 and 8). In contrast, CH₃OH is clustered along the north of the core (see red cluster in combs. 3 and 8, blue cluster in comb. 9), showing patterns similar to c-C₃H₂ in Case 1 and Case 2, but contrary to its own clustering behaviour in Case 1. In terms of features, we see an indirect separation in N_{H_2} values: CH₃CCH is associated with mid N_{H_2} , and CH₃OH with higher N_{H_2} (CH₃CCH: cyan cluster in comb. 10; CH₃OH: blue cluster in comb. 9). Combination 5 shows a direct spatial separation of the two molecules into east (CH₃CCH) and west (CH₃OH), visible as upper and lower diagonal in the *intensity/dist2dust*-distribution. Apart from that, CH₃CCH is clustered at the sharp edge of the core in the south-west, with high ∇N_{H_2} and mid V_{offset} (see red cluster in comb. 4, blue cluster in comb. 7), similar to c-C₃H₂ in Case 2. Both CH₃OH and CH₃CCH are additionally clustered at the core centre, at high N_{H_2} , with CH₃OH at high *intensity* (see red cluster in comb. 6), and CH₃CCH at lower *intensity* (see cyan cluster in comb. 3, purple cluster in comb. 6 and yellow cluster in comb. 9). This pattern was not observed for either CH₃OH or CH₃CCH in Case 1 or Case 2.

L1544 (Fig. C.11): For CH_3OH and CH_3CCH , the clustered features reveal structures similar to those found in Case 1 and Case 2 (e.g. in V_{offset} over *intensity*, *linewidth* over V_{offset} , N_{H_2} over V_{offset} , ∇N_{H_2} over V_{offset} , N_{H_2} over *linewidth*). As before, this leads to a division of the core into north/south and on-centre/off-centre regions. CH_3OH exhibits a north-south separation, visible in the features V_{offset} , *intensity*, and ∇N_{H_2} (see blue and red clusters in combs. 1-4), similar to the pattern seen in Case 1. In contrast, CH_3CCH is predominantly clustered in the south and on the $\text{c-C}_3\text{H}_2$ peak (see yellow cluster in combs. 3, 4, and 7), as well as at the core centre (see yellow cluster in comb. 2, cyan cluster in comb. 3, red cluster in combs. 5, 6 and 9), which were both observed in Case 2. Additionally, CH_3CCH is clustered at its molecule peak in the north-west, with high *intensity*, high *linewidth*, and low ∇N_{H_2} , which was not seen in Case 2 (see cyan cluster in comb. 7). Direct molecular segregation occurs only in comb. 3, where CH_3CCH is clustered at the core centre (cyan cluster) with high *intensity* and high N_{H_2} , while CH_3OH is clustered around the centre at lower *intensity* and lower N_{H_2} (blue and red clusters). Ring-like cluster structures are visible in N_{H_2} and ∇N_{H_2} in combinations 6, 8, 9, and 10; however, they do not display any molecular segregation but instead a balanced ratio between CH_3CCH and CH_3OH .

Summary of Case 3: All cores display cluster structures with feature-wise or spatial separation between CH_3OH and CH_3CCH . However, the molecular segregation is less distinct than in Case 1 and Case 2, and it is mainly visible in the features *intensity*, V_{offset} , N_{H_2} , and ∇N_{H_2} . In B68 and L1544, the clustering predominantly reproduces the structures and clusters found in Case 1 and 2. All three cores show minor differences of cluster behaviour compared to Case 1 and Case 2, where CH_3OH or CH_3CCH behave similar to $\text{c-C}_3\text{H}_2$. This is particularly evident in L1521E, where CH_3OH and CH_3CCH show a more distinct separation, similar to Case 1 and 2.

Case 4: $\text{c-C}_3\text{H}_2$ vs CH_3OH vs CH_3CCH

With a dataset containing three molecules, it is more difficult and less clear to identify molecular segregation, as the ratios between the molecules mostly do not show big variations from the initial ratio of the dataset (the initial ratios for $\text{c-C}_3\text{H}_2/\text{CH}_3\text{OH}/\text{CH}_3\text{CCH}$ are 37/36/27 for B68, 29/35/36 for L1521E, and 28/39/33 for L1544). In the following, we discuss the clustering results for each core individually.

In B68 (see Fig. C.12), $\text{c-C}_3\text{H}_2$ is clustered in a shell along the east, reproducing the structure of Case 1. CH_3CCH is clustered in the core centre, reproducing the behaviour in Case 2 and Case 3. However, CH_3OH is not clustered in the core centre as in Case 1 but instead around the centre in shells along the east and along the west, similar to what was found in Case 3. Additionally, $\text{c-C}_3\text{H}_2$ and CH_3OH show concentric clusters around the centre, visible in N_{H_2} , similar to before.

In L1521E (see Fig. C.13), $\text{c-C}_3\text{H}_2$ is clustered in the northern part of the core, reproducing the clusters in both Case 1 and 2. It also shows the small cluster along the sharp edge in the south-west of the core, seen in Case 2. CH_3CCH is clustered in the south of the

core and the core centre, reproducing the cluster structures of Case 2 and Case 3. CH_3OH , on the other hand, is not clustered in the south as seen in Case 1 but instead in the core centre, as in Case 3 (where its emission peaks) and the north-west of the core.

In L1544 (see Fig. C.14), the division of the core into north/south is visible in V_{offset} for CH_3OH , similar to Cases 1 and 3. For $\text{c-C}_3\text{H}_2$, the association with the northern part seen in Case 1 and Case 2 cannot be reproduced. Instead, it is clustered only on the CH_3OH peak in the north-east of the core. The cluster in the core centre can be reproduced. Additionally, both $\text{c-C}_3\text{H}_2$ and CH_3OH are clustered on their respective molecular peaks, which was not seen in Case 1 or Case 2. For CH_3CCH , both the cluster in the core centre and on the $\text{c-C}_3\text{H}_2$ peak are reproduced. However, Case 4 does not recreate CH_3CCH as being associated with the southern part of the core in the north-south division as in Cases 2 and 3. Also, the molecular separation in ∇N_{H_2} and V_{offset} that was seen in combination 4 in Case 1 and 2 is not reproduced with this combined dataset.

To summarise, in B68 and L1521E, the clustering behaviour of $\text{c-C}_3\text{H}_2$ and CH_3CCH seen in Case 2 can be reproduced, but CH_3OH behaves now differently than in Case 1. In L1544, however, the behaviour of CH_3OH and part of CH_3CCH seen in Case 1 and Case 2 can be reproduced, but not the behaviour of $\text{c-C}_3\text{H}_2$. We discuss this further in Sect. 3.5.1.

3.4.4 CH_3CCH abundances

This section focuses on comparing the CH_3CCH abundances towards the dust peaks of the different cores. In addition to B68, L1521E, and L1544, we include data from the prestellar cores HMM-1, L429, L694-2, and OphD, which were observed but not analysed in the IRAM project of Spezzano et al. (2020). To calculate the abundances at the dust peaks, we divide the CH_3CCH column density by the H_2 column density. The H_2 column density at the dust peak is extracted from the respective $N(\text{H}_2)$ map (derived from *Herschel* SPIRE observations, Spezzano et al. 2020) using a circular aperture with a diameter of $16''$ (matching the *Herschel* map pixel size). We assume a 20% uncertainty for the resulting values. To derive $N(\text{CH}_3\text{CCH})$, we convolve the CH_3CCH spectral cubes with the $40''$ beam of the *Herschel* telescope and extract the spectrum at the dust peak using the same $16''$ aperture. The column density is calculated using a one-dimensional Gaussian fit, and following Mangum & Shirley (2015), with the respective spectroscopic parameters listed in Table 3.2. Since we do not have sufficient data to determine a precise excitation temperature via a rotational diagram for all cores, we adopt a standard excitation temperature of 8 K for consistency. A lower (higher) excitation temperature only affects the abundances of CH_3CCH (and $\text{c-C}_3\text{H}_2$ in L1544) by shifting them to higher (lower) values, but the overall trend stays the same. For each core, we select the most optically thin transition with the smallest (propagated) error in column density. Specifically, we use the CH_3CCH ($5_1 - 4_1$) transition for most cores, except for L694-2 and L1544, where the ($5_0 - 4_0$) and ($6_1 - 5_1$) transitions are used, respectively.

For comparison, we also calculate the abundances of $\text{c-C}_3\text{H}_2$, using the ($2_{02} - 1_{11}$) transition for all cores except L1544, where the ($3_{2,2} - 3_{1,3}$) transition is applied. Figure B.7 displays the extracted spectra along with their Gaussian fits. Figure 3.4 presents the

resulting CH_3CCH (blue circles) and $\text{c-C}_3\text{H}_2$ (orange squares) abundances at the dust peaks of the different cores for an assumed excitation temperature of 8 K.

The CH_3CCH abundances of the starless cores (see left part of Fig. 3.4) are about one order of magnitude higher than the values of the prestellar cores (see right part of Fig. 3.4), suggesting an evolutionary trend of CH_3CCH from the starless to the prestellar phase. Notably, the CH_3CCH abundances in L1544 are significantly higher compared to the other prestellar cores and even compared to the starless cores. This suggests that the observed variations in CH_3CCH are probably influenced not only by the evolutionary stage but also by environmental factors. This is to be discussed further in Sec. 3.5.2. However, to further study the interplay between environmental and evolutionary or dynamical effects on the CH_3CCH abundance spread, additional work is necessary that goes beyond the scope of this paper.

The $\text{c-C}_3\text{H}_2$ abundances show much less variation than those of CH_3CCH and do not exhibit a significant difference between the starless and the prestellar stages. The abundances are spread within one order of magnitude, suggesting that $\text{c-C}_3\text{H}_2$ and CH_3CCH trace different layers within the core. This indicates that $\text{c-C}_3\text{H}_2$ is largely unaffected by the evolution from the starless to the prestellar stage.

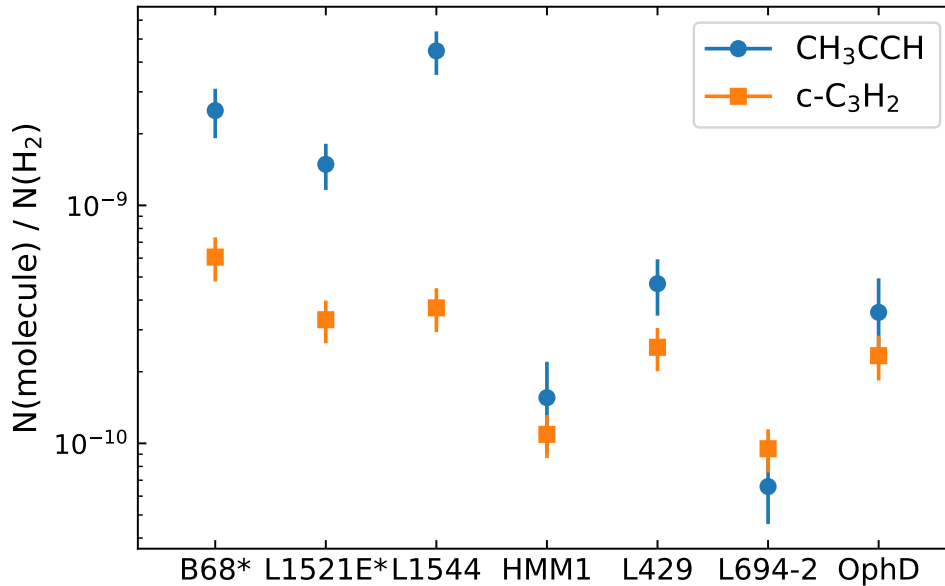


Figure 3.4: Comparison of the abundances of CH_3CCH and $\text{c-C}_3\text{H}_2$ at the dust peaks of different starless and prestellar cores. The molecular column density was calculated assuming $T_{\text{ex}} = 8$ K (see text for details). The starless cores are marked with an asterisk.

3.5 Discussion

3.5.1 Density-based clustering

Segregation between c-C₃H₂ and CH₃OH. Density-based clustering is able to find molecular differentiation in our dataset. In particular, the clustering with the dataset containing c-C₃H₂ and CH₃OH (Case 1) successfully reproduces the known molecular segregation between these two molecules in B68, L1521E, and L1544. This segregation is attributed to uneven illumination across the cores, as discussed in Spezzano et al. (2016, 2020). The differentiation mainly appears in the features *intensity*, V_{offset} , N_{H_2} , and ∇N_{H_2} . In addition, the following pairs of features frequently show segregation: $\text{intensity}/V_{\text{offset}}$, $\text{intensity}/\text{dist2dust}$, $\text{intensity}/N_{\text{H}_2}$, $\text{intensity}/\nabla N_{\text{H}_2}$, $V_{\text{offset}}/N_{\text{H}_2}$, $V_{\text{offset}}/\nabla N_{\text{H}_2}$.

Segregation between c-C₃H₂ and CH₃CCH. The clustering analysis of the dataset containing c-C₃H₂ and CH₃CCH (Case 2) reveals molecular segregation between these two carbon chains in all three cores. Like in Case 1, the segregation appears in the features *intensity*, V_{offset} , N_{H_2} , ∇N_{H_2} , and similar pairs of features. However, in B68 and L1521E, a differentiation between these two molecules is not apparent in their emission maps (see Fig. B.1) and was therefore previously unrecognised. The segregation between c-C₃H₂ and CH₃CCH suggests that these molecules trace different layers in the cores, representing different physical conditions. A similar result was discussed in Lin et al. (2022a), where c-C₃H₂ was found to trace lower density regions than for example CH₃OH.

In B68 and L1521E, the emission of CH₃CCH is less spatially extended compared to c-C₃H₂ (see Fig. B.1). In B68, the CH₃CCH emission is concentrated on the core centre, while in L1521E, it is primarily located in the eastern part of the core. Due to the lower number of available data points for CH₃CCH compared to c-C₃H₂, the clustering dataset in Case 2 includes two CH₃CCH transitions for B68 and three transitions for L1521E (as detailed in Table 3.4). This results in a higher density of data points in the core centre of B68 and the eastern part of L1521E, increasing the likelihood of forming clusters at these locations with density-based algorithms such as DBSCAN and HDBSCAN. The incomplete coverage of CH₃CCH across the cores becomes more apparent in Case 3 (CH₃OH, and CH₃CCH) and Case 4 (c-C₃H₂, CH₃OH, and CH₃CCH). In Case 4, CH₃CCH forms clusters similar to those in Case 2 (c-C₃H₂ and CH₃CCH), but CH₃OH shows different behaviour compared to Case 1 (c-C₃H₂ and CH₃OH). In contrast, in L1544 – where the emission maps of c-C₃H₂, CH₃OH, and CH₃CCH extend across the entire core – the clustering results of Case 4 differ significantly from Case 1 and 2. In Case 3, on the other hand, most clusters found in Case 1 and 2 are recreated. The only exception is CH₃OH in L1521E, that mimics the clustering behaviour of c-C₃H₂ instead.

Similarities between CH₃OH and CH₃CCH. Our analysis reveals that in B68 and L1521E, the clusters dominated by CH₃OH in Case 1 behave very similarly to those dominated by CH₃CCH in Case 2, despite the fact that the CH₃OH emission is as spatially

extended as $\text{c-C}_3\text{H}_2$ in both cores and CH_3CCH is not. The CH_3CCH - and CH_3OH -dominated clusters are associated with the same features – N_{H_2} and V_{offset} for L1521E, *intensity* for B68 – and are located in the same regions within the cores (south-east for L1521E, and the core centre for B68). Additionally, these clusters are spatially distinct from the $\text{c-C}_3\text{H}_2$ clusters. As shown in Fig 3.1, $\text{c-C}_3\text{H}_2$ exhibits broader linewidths than CH_3OH and CH_3CCH across B68, likely indicating that it traces a different, more turbulent layer. These clustering results may therefore reflect differences in the physical layers traced by the molecules.

In L1544, the similarity in cluster behaviour between CH_3OH and CH_3CCH relative to $\text{c-C}_3\text{H}_2$ is observed in only one feature pairing: ∇N_{H_2} and V_{offset} (see combination 4). This could be linked to both CH_3OH and CH_3CCH tracing gas influenced by slow accretion flows. For CH_3OH , such an association has been demonstrated and discussed, for instance, by Lin et al. (2022a). In Sec. 3.5.2, we further explore how the CH_3CCH peak in L1544 may also be impacted by inflowing gas.

Relevance of different features in the clustering analysis. In our clustering analysis, V_{offset} appears to be a dominant feature that drives the division of the cores into the different clusters, and in some cases, molecular separation. In L1521E, the core is divided into north-west/south-east, while L1544 shows a north/south split, both reflecting the velocity structure of the cores. A similar split is indicated in B68, with a separation between the core centre and a shell along the eastern side, although this division is less pronounced. The strong dependence of velocity structure with chemical prominence indicates that static chemical models might not be sufficient to predict observed features in full. While the overall physical structure of these cores is generally well-described by quasi-static models, understanding the anisotropic chemical structures requires a more dynamic approach (see also Lin et al. 2022a).

Both B68 and L1544 show clusters with concentric ring structures around their core centres, following the patterns of N_{H_2} and ∇N_{H_2} . This ring-like pattern is a result of the rather spherical shape of these cores. In contrast, L1521E, which is more elongated and irregularly shaped, does not show this pattern. The clustering analysis of both L1521E (Case1) and L1544 (Case1/Case2) reveals molecular separation in the $V_{\text{offset}}-\nabla N_{\text{H}_2}$ distribution, which does not appear for B68. This difference may be due to environmental factors, as B68, a Bok globule, is exposed to relatively uniform external illumination. In L1521E, the clusters with the highest ∇N_{H_2} values in this feature pairing (see combination 4) are dominated by $\text{c-C}_3\text{H}_2$. These data points are located at the filamentary edge in the southern part of the core, where $\text{c-C}_3\text{H}_2$ peaks and external illumination is strongest. In L1544, the data points with highest ∇N_{H_2} values also come from the southern part of the core, near the filamentary edge with high external illumination. However, in this case they are dominated by CH_3OH and CH_3CCH instead of $\text{c-C}_3\text{H}_2$. This suggests that the clustering results reflect the distinct environmental conditions within each core. Overall, the clusters with prominent N_{H_2} and ∇N_{H_2} features appear to represent the chemical patterns across the core structures, with differences in the clustering likely tied to varying

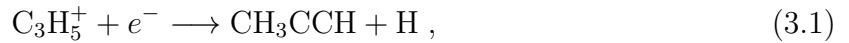
environmental conditions.

The features *dist2dust* and *linewidth* appear to be less significant in our analysis, as they rarely exhibit distinct structures or molecular separations by themselves. However, when combined with other features, such as *dist2dust/intensity* or *linewidth/V_{offset}*, they provide additional insights. The one-dimensional projected distance to the dust peak seems less relevant in the clustering analysis compared to N_{H_2} and ∇N_{H_2} , as these two features better characterise the immediate environment of a data point or pixel.

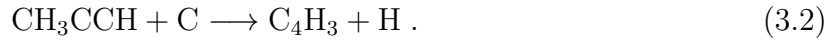
3.5.2 Evolution traced by CH₃CCH

In the starless cores B68 and L1521E, the distribution of the CH₃CCH emission overlaps with that of c-C₃H₂ (see Fig. B.1). In L1544, however, the peak of the CH₃CCH emission is not located in the south-east of the core, where the other carbon chains are found, but rather in the north-west. It is known that in the north-east of L1544, around the CH₃OH peak, two filaments converge (e.g. see André et al. 2010; Spezzano et al. 2016). This may result in slow accretion flows (Punanova et al. 2018; Lin et al. 2022a), which could deliver fresh material to the core and help replenish CH₃CCH.

To investigate the formation and destruction routes of CH₃CCH, we conducted chemical simulations using the gas-grain chemical model **pyRate** (Sipilä et al. 2015), applied to the standard physical model of L1544 (Keto et al. 2015). For the gas-phase chemical network we adopted the 2014 public release of the KIDA chemical network (kida.uva.2014; Wakelam et al. 2015), while the grain-surface network is an updated version of the one presented in Semenov et al. (2010). The simulation yields radial abundance profiles as a function of time; we checked the results at an evolutionary time of 10⁵ yrs. The model shows that at intermediate densities ($n \sim 10^4 \text{ cm}^{-3}$), CH₃CCH is mainly formed by the dissociative recombination of C₃H₅⁺:



while it mainly gets destroyed by the reaction with free carbon:



Following this, in regions with high irradiation and therefore active gas-phase chemistry, such as the carbon-chain peak in L1544, CH₃CCH is quickly destroyed due to the high amounts of atomic carbon present in the gas phase. In contrast, the north-western part of L1544 is more shielded from irradiation, allowing CH₃CCH to form from fresh material brought in by the filamentary flow from the north-west. Here, the low abundance of free carbon in the gas phase protects the CH₃CCH from destruction. This is further supported by the velocity and linewidth maps of CH₃CCH, as shown in Fig. 3.5, along with the results for c-C₃H₂ and CH₃OH. The linewidth map of CH₃CCH (bottom right) shows increased linewidths near the emission peak, while the velocity map (top right) reveals a sharp velocity gradient in the same area. This suggests that the CH₃CCH emission peak in L1544 is the landing point or accumulation point of the incoming fresh gas. Similar signs of this active chemistry are observed in c-C₃H₂. The integrated intensity map (Fig. B.1) shows

a local maximum in this region, while the linewidth map (bottom left in Fig. 3.5) exhibits the highest linewidths not at the $c\text{-C}_3\text{H}_2$ peak in the south-east but at the CH_3CCH peak in the north-west. The velocity map of $c\text{-C}_3\text{H}_2$ (top left) also shows a gradient around this area.

The fact that CH_3CCH is an early-type molecule explains why, in our dataset, extended CH_3CCH emission is only observed in the starless cores (B68, L1521E) and the prestellar core L1544 due to the possible material accretion. In contrast, in the other prestellar cores (HMM1, OphD, L694-2, L429), we detect some emission at the respective dust peaks, but no significant emission beyond that. This evolutionary trend is further supported by the CH_3CCH abundances at the dust peaks (shown in Fig. 3.4), where the abundances in the starless cores are approximately one order of magnitude higher than those in the prestellar cores, except for L1544.

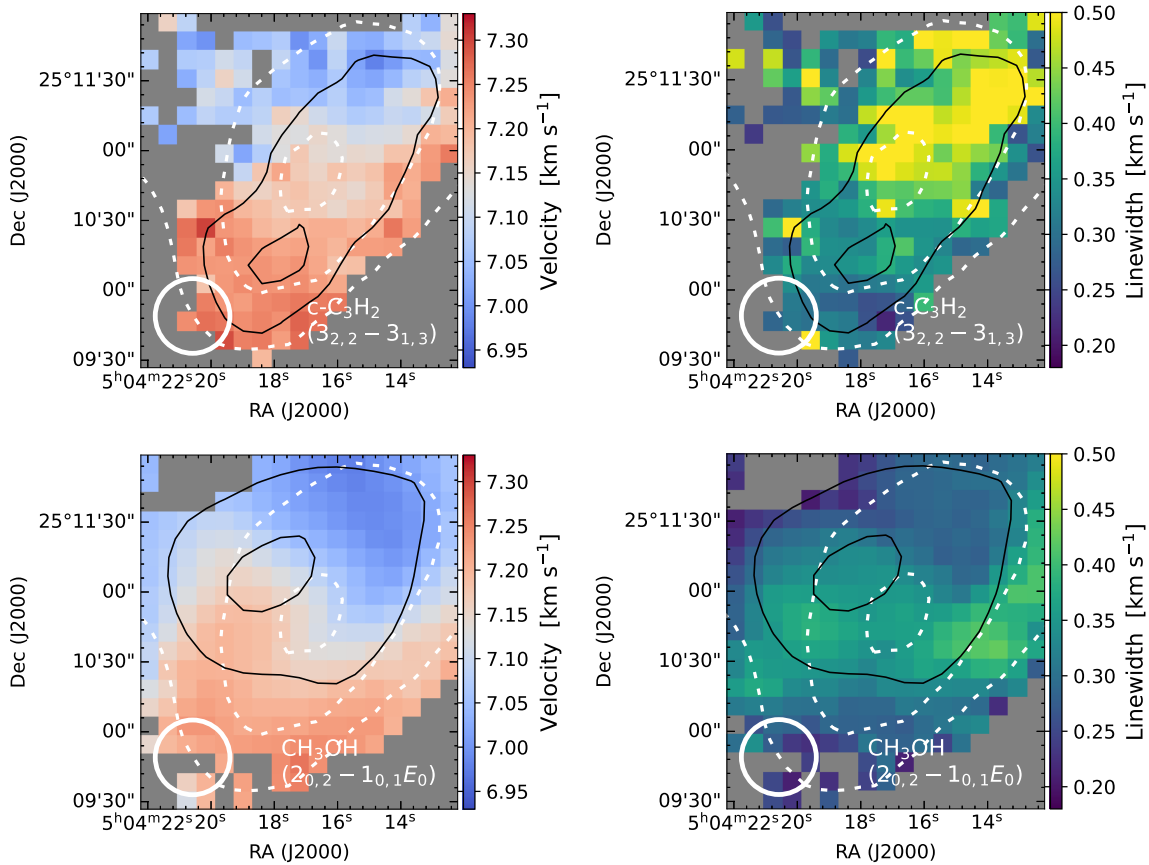


Figure 3.5: Centroid velocities (*left*) and linewidths (*right*) of $c\text{-C}_3\text{H}_2$ (*top*), CH_3OH (*middle*), and CH_3CCH (*bottom*) towards the prestellar core L1544. Black contours show 50% and 90% of the respective molecular emission peak. White contours show 30%, 50%, and 90% of the H_2 column density peak derived from *Herschel* maps (Spezzano et al. 2016). The white circle in the bottom-left corner indicates the beam size of the IRAM 30 m telescope (32'').

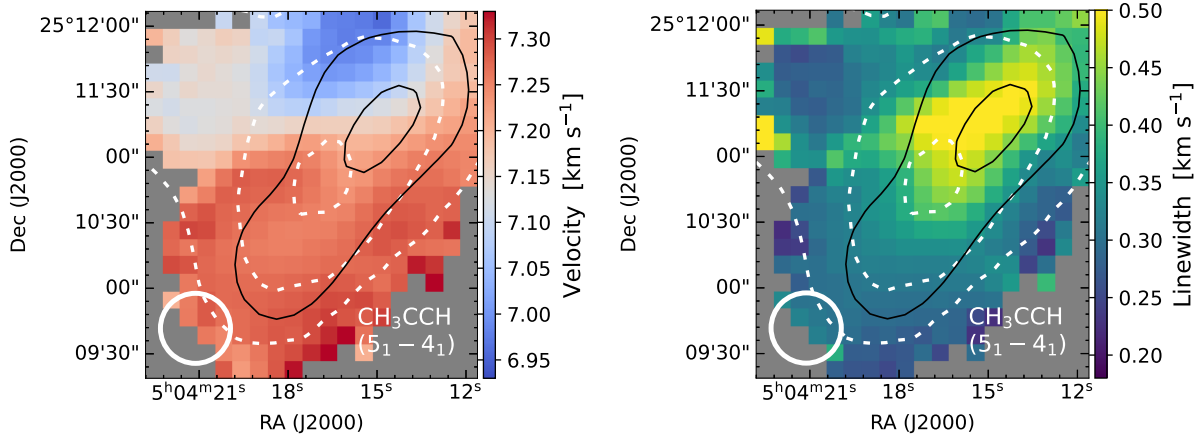


Figure 3.6: continued.

3.6 Conclusions

We have presented an analysis of molecular differentiation using the density-based clustering algorithms DBSCAN and HDBSCAN. The clustering was applied to four different datasets, in order to compare the emission morphologies of $c\text{-C}_3\text{H}_2$, CH_3OH , and CH_3CCH observed towards the starless cores B68 and L1521E and the prestellar core L1544.

Our main results can be summarised as follows:

- The analysis with density-based clustering finds a significant chemical differentiation across the cores in our dataset. It successfully reproduces the known molecular segregation of $c\text{-C}_3\text{H}_2$ and CH_3OH for B68, L1521E, and L1544. Furthermore, the clustering analysis identifies a segregation between $c\text{-C}_3\text{H}_2$ and CH_3CCH in all three cores, which is not apparent from visual inspection of the emission maps.
- The most relevant features in the clustering analysis are integrated intensity, velocity offset, H_2 column density, and H_2 column density gradient. Distinct and recurring cluster structures in the H_2 column density and the gradient highlight structural and chemical patterns across the cores. Differences in the relevance of these two features for the three cores reflect the varying environmental conditions within each core. The strong relation between molecular emission and velocity structure suggests that to understand anisotropic chemical structures, static chemical models are not sufficient, but dynamical models are necessary.
- Increased CH_3CCH abundances towards the starless cores compared to prestellar cores indicate an evolutionary trend. Increased CH_3CCH abundances towards L1544 suggest an additional influence of environmental factors. In fact, in L1544, the CH_3CCH peak in the north-west of the core appears to trace the landing point

of chemically fresh gas that is accreted to the core. Unlike the photochemically active south of the core, this area is shielded from external irradiation, which protects CH_3CCH from being destroyed by free carbon atoms.

- The clustering analysis finds a similar behaviour between CH_3OH and CH_3CCH relative to $\text{c-C}_3\text{H}_2$ in all cores. This indicates that $\text{c-C}_3\text{H}_2$ traces an outer layer of gas and possibly a lower-density shell compared to the other two molecules. In L1544, the similar clustering patterns observed for CH_3OH and CH_3CCH may reflect the influence of accretion processes in shaping the molecular distribution.

Our results demonstrate that a successful density-based clustering approach for studying astrochemical processes does not require a large dataset covering multiple molecules across various cores. In fact, the results are often easier to interpret when only two or three molecules are considered. While this clustering method is more time-consuming than techniques such as principal component analysis, it can process much more detailed information and provide deeper insights into the core's structure. Using the more general approach of describing a data point's location through its H_2 column density and its gradient, rather than relying on spatial coordinates, also enables simple comparisons between cores. In future studies, we aim to expand our analysis of molecular differentiation with density-based clustering to include more cores and molecules, especially those that trace other physical or chemical features. This will also help explore any evolutionary effects that the cores or their environment might have on the molecular distribution.

Data availability

Figures C.3-C.14, presenting the detailed clustering results for Cases 1-4, are published on Zenodo (zenodo.org/records/15519030).

Chapter 4

Deuteration of HC_3N and CH_3CCH in the prestellar core L1544

The content of this chapter was submitted for publication to the *Astronomy & Astrophysics* journal on August 5th, 2025. Credit: Giers, K., Spezzano, S., Lin, Y., Caselli, P., Sipilä, O., 2025, subm.

Abstract

Context. Deuterated molecules are a useful diagnostic tool to probe the evolution and the kinematics in the earliest stages of star formation. Due to the low temperatures and high densities in the centre of prestellar cores, the deuterium fraction is enhanced there by several orders of magnitude with respect to the cosmic D/H abundance ratio.

Aims. We study the distribution of the emission and the deuteration of the two carbon chains HC_3N and CH_3CCH throughout the prototypical prestellar core L1544.

Methods. We analyse emission maps of CH_3CCH , CH_2DCCH , CH_3CCD , HC_3N , HCC^{13}CN , and DC_3N , observed towards L1544 with the IRAM 30 m single-dish radio telescope. We use non-local thermodynamic equilibrium radiative transfer calculations, combined with chemical modelling of the molecular abundances, to constrain physical parameters of the observed species. Following this, we derive the corresponding column density and deuteration maps and analyse the chemical processes influencing the different molecular distributions.

Results. We find maximum levels of deuteration of $\text{N}(\text{DC}_3\text{N})/\text{N}(\text{HC}_3\text{N}) = 0.07 \pm 0.01$, $\text{N}(\text{CH}_2\text{DCCH})/\text{N}(\text{CH}_3\text{CCH}) = 0.15 \pm 0.03$, and $\text{N}(\text{CH}_3\text{CCD})/\text{N}(\text{CH}_3\text{CCH}) = 0.09 \pm 0.02$. The deuteration of HC_3N appears very homogeneous across the core, with widespread D-fraction values above 0.06, tracing intermediate-density gas in the outer layers of the core. CH_3CCD is most efficiently formed in the higher-density regions towards the core centre, while the deuteration fraction of CH_2DCCH traces a local density enhancement in the north-east of the core, coinciding with the CH_3OH emission peak.

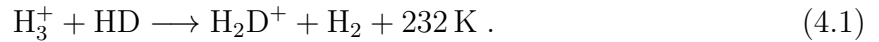
Conclusions. The results suggest that gas-phase reactions dominate the formation and

deuteration of both HC₃N and CH₃CCH in L1544, with spatial variations driven by physical structure, density and external radiation. The significantly larger deuteration fraction of CH₂DCCH compared to CH₃CCD and a tentative gradient with larger values in the north suggest that there are different deuteration mechanisms for the two functional groups, with varying efficiency across the core. Similarities of the CH₂DCCH emission to CH₂DOH might indicate an additional deuteration pathway of CH₃CCH on the surfaces of dust grains, like observed for H₂CO.

Keywords: astrochemistry – ISM: clouds – ISM: molecules – ISM: abundances – stars: formation

4.1 Introduction

The deuterium fractionation of molecules is a powerful tool to study the early stages of star formation. Due to the combination of high densities and low temperatures in the prestellar phase, CO is largely depleted onto the surfaces of dust grains (see e.g. Caselli et al. 1999). As CO is the main destroyer of several ions (e.g. H₃⁺ and its deuterated forms; Dalgarno & Lepp 1984), its depletion leads to an enhancement in their abundances. Simultaneously, the main pathway for the formation of H₂D⁺ opens up:



As this reaction is exothermic, the backwards reaction is suppressed at the low temperatures present in prestellar cores (assuming a low ortho-to-para H₂ ratio; e.g. Pagani et al. 1992). The enhanced abundance of H₂D⁺, as well as of D₂H⁺ and D₃⁺ then leads to an increased deuteration of other molecules, via a deuteron transfer, as predicted by models (e.g. Walmsley et al. 2004).

The evolved prestellar core L1544 is known for its high levels of deuteration (see e.g. Crapsi et al. 2005; Chantzos et al. 2018; Giers et al. 2023). The distribution of the deuterium fractionation across this core has been studied for a variety of molecules: HCO⁺, N₂H⁺, CH₃OH, H₂CO, c-C₃H₂, and H₂CS (Caselli et al. 2002a; Redaelli et al. 2019; Chacón-Tanarro et al. 2019; Giers et al. 2022; Spezzano et al. 2022a). Those observations show that, for most species, the deuterium fractionation in L1544 is most efficient close to the dust peak of the core, in higher-density gas layers. On the other hand, the different molecules show a wide range of deuteration levels: While H₂CO and HCO⁺ show rather low values of around 3%, N₂H⁺ and H₂CS reach high values of up to 30%. On top of its chemical richness, the core also shows a spatial molecular segregation with characteristic emission peaks of molecules, namely c-C₃H₂, CH₃OH, and HNCO (Spezzano et al. 2017). Those seem to be the emission peak of several other species that are chemically related to the respective molecules: C-bearing molecules are grouped in the south on the c-C₃H₂ peak, O-bearing molecules gather in the north-east on the CH₃OH peak, and HNCO in the north-west is joined by CH₃CCH. The spatial differentiation between carbon chains in the south and CH₃OH in the north has been explained by a nonuniform external illumination of the core (Spezzano et al. 2016). More precisely, L1544 is located at the edge of a filament in

the Taurus molecular cloud, which leads to the southern part being exposed to the interstellar radiation field (ISRF), while the northern part is more protected by the molecular cloud. The ISRF drives the photochemistry in the south, enhancing the abundance of free carbon atoms in the gas phase, and subsequently the abundance of carbon chains. In the protected north, however, photochemistry is not a dominant process, and the carbon is mainly locked in CO. The northern part of the core is also known to be the meeting point of two filamentary structures (e.g. see Spezzano et al. 2016; Lin et al. 2022a). Giers et al. (2025) suggest that inflowing, fresh material actually enhances the abundances of chemically young species in this area, causing the HNCO peak in the north-west of the core.

This work focuses on the emission and distribution of the two carbon chains HC_3N and CH_3CCH and their singly deuterated isotopologues. In L1544, the two molecules show a spatial differentiation, where the emission of HC_3N follows other carbon chains and peaks in the south of the core, while CH_3CCH peaks in the north-west on the HNCO peak. The fact that the two carbon-chain molecules exhibit distinct spatial distributions is particularly interesting, given that such a segregation is typically observed between different molecular families rather than within the same class of species. These observations offer new insights into the complex chemical structure of the prototypical core L1544. Furthermore, this dataset gives the opportunity to study the deuteration of larger molecules, providing deeper understanding of the chemical processes that dominate in prestellar cores and shape the observed chemical structures. In addition, the study of the deuteration can help to better understand the formation pathways of the molecules and, eventually, put constraints on chemical models.

Cyanoacetylene, HC_3N , is the smallest molecule from the family of cyanopolyynes, which are possible precursors of prebiotic molecules (see e.g. Shingledecker et al. 2021). After its first detection in space by Turner (1971), HC_3N has since been detected in various interstellar environments (see e.g. Morris et al. 1976; Walmsley et al. 1980; Chapillon et al. 2012; Spezzano et al. 2017). The deuterated counterpart, DC_3N , was first detected by Langer et al. (1980) towards the low-mass dark cloud TMC-1, and by Howe et al. (1994) towards several other dark cores. Methylacetylene, CH_3CCH , is a large carbon chain, and a precursor in the formation of polycyclic aromatic hydrocarbons (PAHs) (Parker & Kaiser 2017). It contains a methyl group with three equivalent H atoms and one H atom in the terminal CCH group, which makes the study of its deuteration very useful to understand the molecule's formation pathways. In the ISM, CH_3CCH was first detected by Buhl & Snyder (1973), followed by detections in various other environments (van Dishoeck et al. 1995; Fontani et al. 2002; Vastel et al. 2014; Gratier et al. 2016; Spezzano et al. 2017; Agúndez et al. 2019; Lin et al. 2022b). The singly deuterated isotopologues CH_2DCCH and CH_3CCD were first detected in space by Gerin et al. (1992) and Markwick et al. (2005), respectively. The detection of the doubly deuterated forms CHD_2CCH and CH_2DCCD was reported by Agúndez et al. (2021).

Both HC_3N and CH_3CCH are believed to primarily form in the gas phase. The main formation pathways for CH_3CCH are thought to be via ion-molecule reactions, neutral-neutral reactions, and dissociative recombination of larger hydrocarbons (Schiff & Bohme

1979; Turner et al. 1999; Calcutt et al. 2019; Giers et al. 2025), while for HC₃N, potential formation routes are via HCN, CN, or HNC (Hily-Blant et al. 2018). There is not much known about the deuteration process of either species, but there is evidence that it happens mainly in the gas phase. According to Rivilla et al. (2020), DC₃N most likely forms via ion-molecule reactions, for example of HC₃N directly reacting with H₂D⁺, whereas the deuteration of CH₃CCH is believed to happen via the dissociative recombination of C₃H₆D⁺ and C₃H₅D⁺ (Agúndez et al. 2021).

In this work, we present the emission maps of CH₂DCCH, CH₃CCD, and DC₃N, and their non-deuterated isotopologues CH₃CCH, HC₃N, and HCC¹³CN, observed towards the prestellar core L1544. We report the observations and detected lines in Sect. 4.2. In Sect. 4.3.2, we perform non-LTE radiative transfer calculations to model spectral lines extracted towards the dust peak of the source to constrain physical parameters of the transitions. Following this, we derive the corresponding column density and deuterium fraction maps of the different species in Sect. 4.3.3. In Sect. 4.4, we discuss the deuterium fractionation of HC₃N and CH₃CCH in L1544 and the implications on the understanding of the chemical processes happening. Section 4.5 concludes the article.

4.2 Data

The data presented in this work were taken with the IRAM 30 m single-dish radio telescope on Pico Veleta in the Sierra Nevada, Spain. The observations were carried out between October 2013 and June 2022 (PIs: Silvia Spezzano, Katharina Giers). The data on HC₃N, HCC¹³CN, CH₃CCH, and CH₂DCCH have also been used in Spezzano et al. (2017) and Giers et al. (2025). The on-the-fly (OTF) maps were observed in position switching mode, using the EMIR E090 receiver and the Fourier transform spectrometer (FTS) backend with a spectral resolution of 50 kHz. The observed 2.5' × 2.5' OTF maps are centred on the source dust emission peak ($\alpha_{2000} = 05^{\text{h}}04^{\text{m}}17^{\text{s}}.21$, $\delta_{2000} = +25^{\circ}10'42''.8$, Ward-Thompson et al. 1999). The observed transitions are summarised in Table 4.1.

The data processing was done using the GILDAS software (Pety 2005) and the python packages `pandas` (Wes McKinney 2010) and `spectral-cube` (Ginsburg et al. 2019). All emission maps were gridded to a pixel size of 8'' with the CLASS software in the GILDAS packages; this corresponds to one-third to one-quarter of the actual beam size, depending on the frequency. To create a uniform dataset, we additionally resampled the data to a spectral resolution of 0.18 km s⁻¹, corresponding to the resolution of the lowest frequency observation (84.4 GHz). The antenna temperature T_A^* was converted to the main beam temperature T_{mb} using the relation $T_{\text{mb}} = F_{\text{eff}}/B_{\text{eff}} \cdot T_A^*$. The corresponding values for the 30 m forward (F_{eff}) and main-beam efficiencies (B_{eff}) are given in Table 4.1. Figure 4.1 shows the integrated intensity maps of the observed transitions, computed by integrating over a velocity range of 0.7 km s⁻¹ around each line. To enable the comparison between the molecules in the later analysis, all maps were convolved to an angular resolution of 31'', which corresponds to the half-power beam width (HPBW) of the largest observed beam.

Table 4.1: Spectroscopic parameters of the observed lines. The transitions used to derive the molecular column density are marked with an asterisk.

Molecule	Transition	Frequency ^a (MHz)	$F_{\text{eff}}/B_{\text{eff}}$	E_{up}^a (K)	g_{up}^a	A^a (s ⁻¹)
HC ₃ N	11 – 10	100076.39(2)	0.95/0.79	28.8	23	7.77×10^{-5}
HCC ¹³ CN*	10 – 9	90601.78(3)	0.95/0.80	23.9	21	5.74×10^{-5}
DC ₃ N*	10 – 9	84429.814(3)	0.95/0.81	22.3	21	4.69×10^{-5}
DC ₃ N	11 – 10	92872.375(3)	0.95/0.80	26.7	23	6.24×10^{-5}
CH ₃ CCH*	5 ₁ – 4 ₁	85455.6667(1)	0.95/0.81	19.5	22	1.95×10^{-6}
CH ₃ CCH	5 ₀ – 4 ₀	85457.3003(1)	0.95/0.81	12.3	22	2.03×10^{-6}
CH ₃ CCH	6 ₁ – 5 ₁	102546.0242(1)	0.95/0.79	24.5	26	3.46×10^{-6}
CH ₃ CCH	6 ₀ – 5 ₀	102547.9844(1)	0.95/0.79	17.2	26	3.56×10^{-6}
CH ₃ CCD	6 ₁ – 5 ₁	93454.331(3)	0.95/0.80	22.9	26	2.51×10^{-6}
CH ₃ CCD*	6 ₀ – 5 ₀	93456.044(3)	0.95/0.80	15.7	26	2.59×10^{-6}
CH ₂ DCCH*	6 ₀₆ – 5 ₀₅	97080.728(6)	0.95/0.80	16.3	13	3.03×10^{-6}

Notes. Numbers in parenthesis give the uncertainty on the last digit. ^(a) Extracted from the Cologne Database for Molecular Spectroscopy (Müller et al. 2001).

4.3 Analysis

4.3.1 Spectral line emission

The goal of this work is to study the deuterium fraction of HC₃N and CH₃CCH across the source. To do so, a reliable estimation of the molecular column densities is crucial.

For CH₃CCH and CH₃CCD, we have observed more than one transition. Figure 4.2 presents spectra of the CH₃CCH (5–4) and the CH₃CCD (6–5) K=0 and K=1 transitions, extracted towards the three molecular peaks and the dust peak in L1544 by using a circular aperture with a diameter corresponding to the telescope beam size, 31". The extraction locations and beam size are indicated in Fig. 4.1. The observations show that the line intensity ratio between the K=0 and K=1 transitions varies across the core, for CH₃CCH between 1:1 and 1:1.2, and for CH₃CCD between 1:1 and 1:1.4. In a first approach, we assumed local thermodynamic equilibrium (LTE) and modelled the spectra using `lte_molecule` and `generate_model` from the python package `pyspeckit`. We applied a constant excitation temperature of 10 K, which is consistent with previous measurements of CH₃CCH in this source (e.g. Vastel et al. 2014), and adjusted the column densities to fit the intensity of the K=0 lines. The resulting modelled spectra are shown in red in Fig. 4.2. For both CH₃CCH and CH₃CCD, the line ratio between K=0 and K=1 is estimated to be around 2:1. This is not visible in the observations, which vary between 1:1 and 1:1.4. However, the emission lines are in the optically thin regime, with $\tau_{\text{max}}(T_{\text{ex}} = 10 \text{ K}) \approx 0.27$. Therefore, self-absorption as a cause of the line ratios is unlikely. Following this,

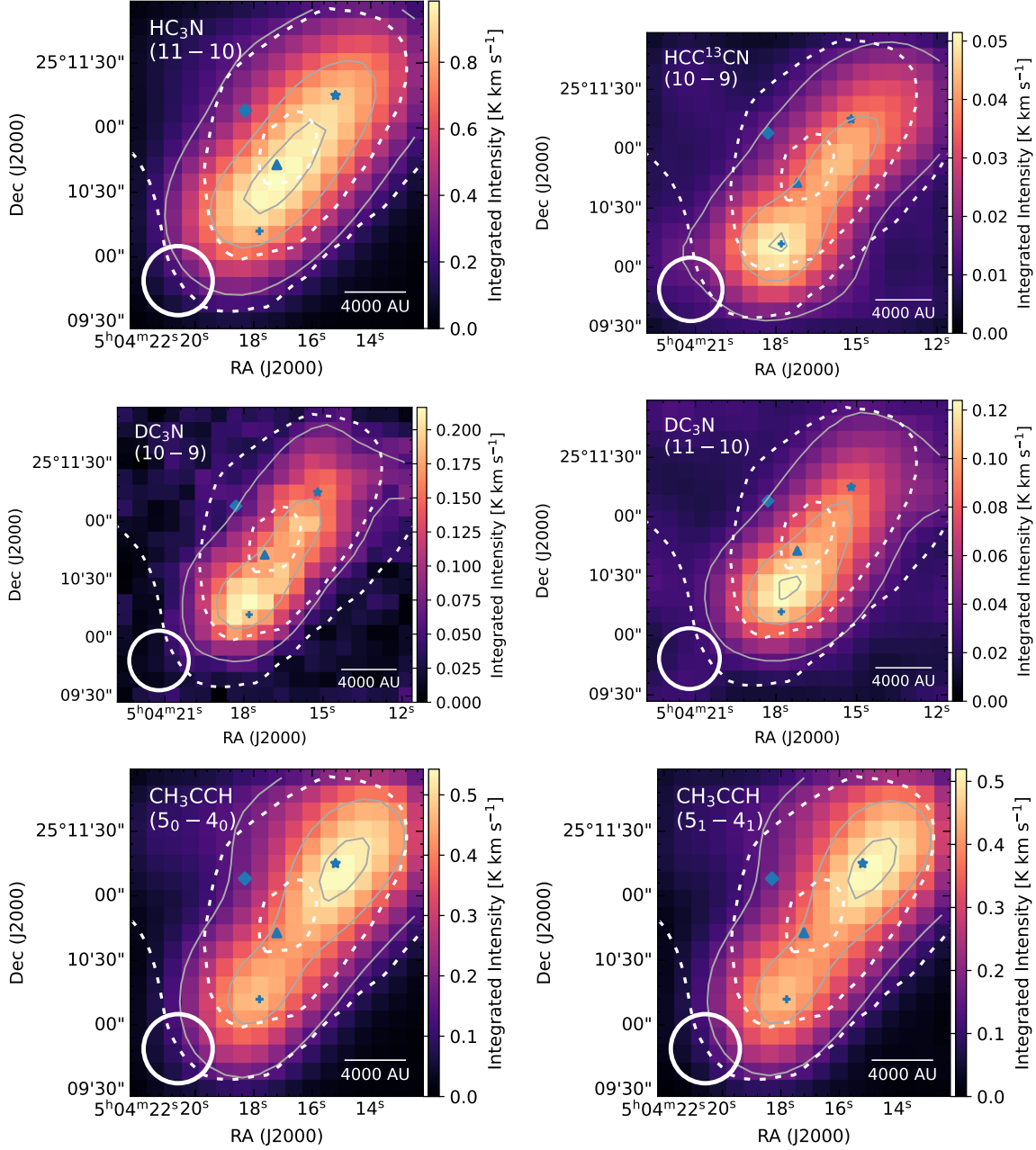


Figure 4.1: Integrated intensity maps of the observed transitions. The solid line contours indicate the 30%, 70%, and 90% level of the peak integrated intensity. The dashed line contours represent 30%, 50%, and 90% of the H_2 column density peak derived from *Herschel* data (Spezzano et al. 2016). The markers in blue represent the dust peak of L1544 (triangle) and the molecular emission peaks of CH_3OH (diamond), CH_3CCH (star), and $\text{c-C}_3\text{H}_2$ (plus sign), where emission spectra (shown in Fig. 4.2) are extracted within a circular aperture with a diameter corresponding to the telescope beam. The white circle in the bottom-left corner indicates the beam size of the IRAM 30 m telescope (31'').

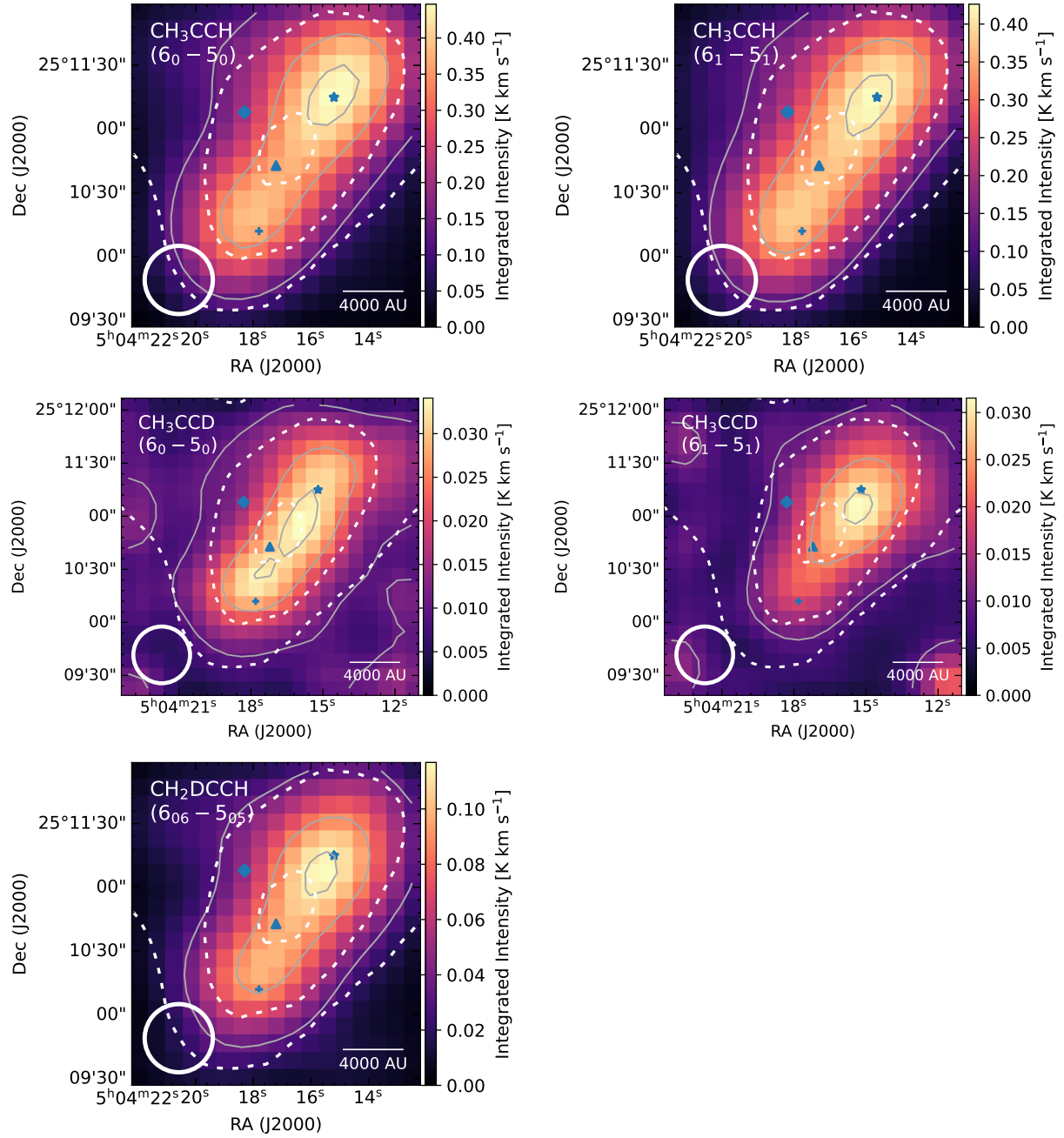


Figure 4.1: continued.

there has to be another effect influencing the line intensities, and given the relative high critical density of CH_3CCH ($n_{\text{crit}} \approx 10^5 \text{ cm}^{-3}$) it is likely that non-LTE effects play a role.

Therefore, a non-LTE approach would be necessary to test if non-LTE effects might be the cause of the observed intensity ratios, and to constrain the molecular column densities. A full non-LTE analysis lies beyond the scope of this paper, because the required three-dimensional model of the temperature and density structure of the core does not exist. Instead, we focus on the analysis of the emission lines extracted at the dust peak of L1544. In the following, we model the observed spectra with radiative transfer in Sec. 4.3.2, before we analyse the column densities and deuterium fractions in Sec. 4.3.3.

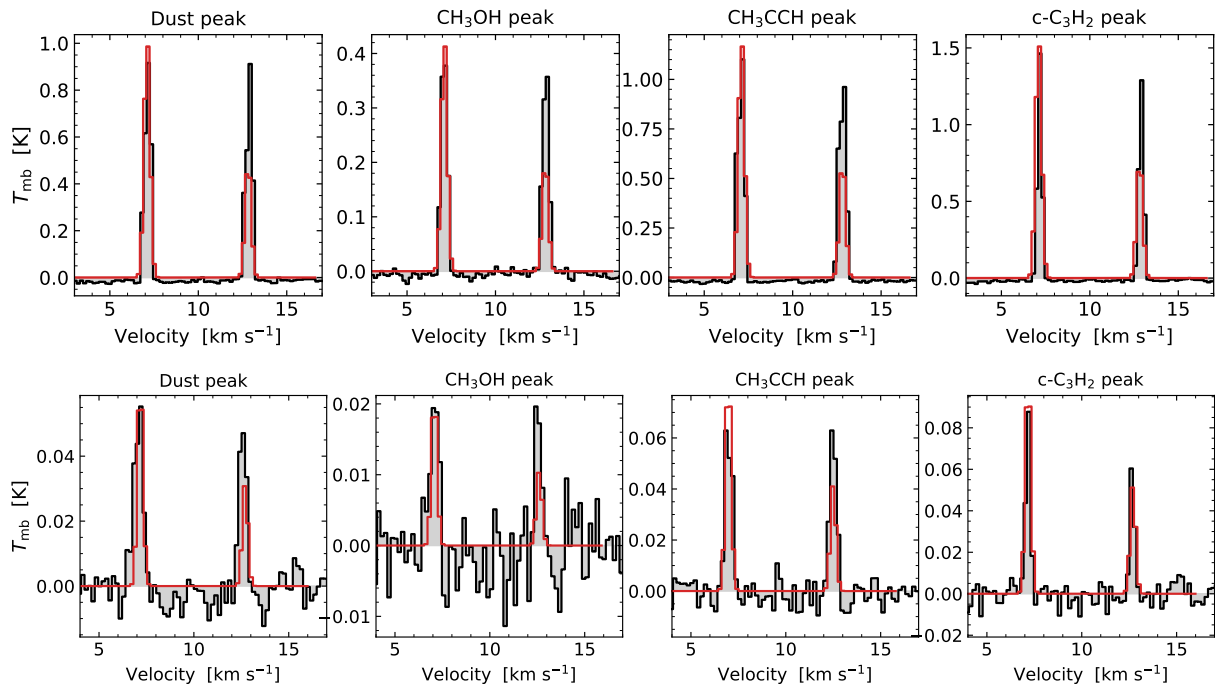


Figure 4.2: Observed spectra (black) of the CH_3CCH (5–4) and the CH_3CCD (6–5) $K=0$ and $K=1$ transitions extracted towards the three molecular peaks in L1544 and the dust peak, using a circular aperture with diameter $31''$. The extraction locations are indicated in Fig. 4.1. Shown in red are synthetic spectra produced with the LTE model generator of the python package *pyspeckit*. The input column densities are $[5, 2, 6, 8](\pm 0.3) \times 10^{13} \text{ cm}^{-2}$ for CH_3CCH and $[3, 1, 4, 5](\pm 0.3) \times 10^{12} \text{ cm}^{-2}$ for CH_3CCD for the dust peak, CH_3OH peak, CH_3CCH peak, and $\text{c-C}_3\text{H}_2$ peak, respectively. The input excitation temperature is set to a constant value of 10 K for both molecules.

4.3.2 Non-LTE modelling at the dust peak

In this section, we simulate the observed molecular spectra with the non-LTE radiative transfer code LOC (Line Transfer with OpenCL, Juvela 2020). We assumed a one-

dimensional model that considers spherical symmetry of the physical properties and chemical abundances. The physical structures are characterised by volume density, kinetic temperature, and infall velocity. For this, we adopted the physical model of L1544 presented in Keto et al. (2015) (hereafter Keto-Caselli model), shown in Fig. 4.3. It describes an unstable quasi-equilibrium Bonnor-Ebert sphere (Ebert 1955; Bonnor 1956) with a peak central H_2 volume density of $n_0 \approx 10^7 \text{ cm}^{-3}$, and a central gas temperature of 6 K. To approximate the molecular abundances, we used profiles of the fractional abundance with respect to H_2 molecules, hereafter referred to as "abundance profiles". We applied both constant abundance profiles (with a complete freeze-out radius of 1800 au, following Caselli et al. 2022), as well as radial abundance profiles predicted by the state-of-the-art gas-grain chemical model `pyRate` (Sipilä et al. 2019). In `pyRate`, we applied the two-phase model of the code, where the gas-phase chemistry and the entire ice chemistry on the grains are active. We used the same initial abundances and chemical networks as presented in Sipilä et al. (2019) and the same standard values for various model parameters as used in Giers et al. (2022). In the chemical simulation, the Keto-Caselli model is used as a static physical model for L1544, assuming an external visual extinction of $A_V = 2 \text{ mag}$ to account for the molecular cloud where the core is embedded. The resulting abundance profiles were extracted at various evolutionary times across the chemical simulation.

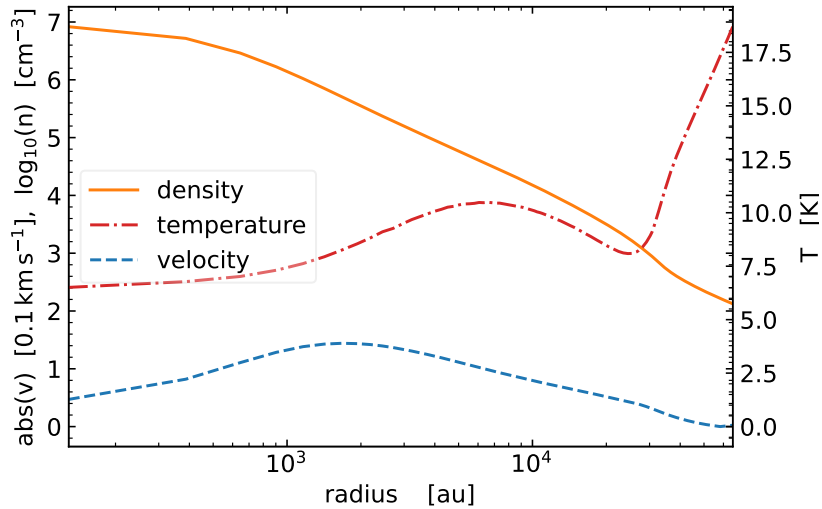


Figure 4.3: Profiles of the gas temperature (red), H_2 number density (orange, in logarithmic scale), and infall velocity (blue, in units of 0.1 km s^{-1}) for the Keto-Caselli model of L1544 (Keto et al. 2015). The velocity in the model is negative, but is shown here as positive to improve the readability.

For the line simulations of HC_3N and CH_3CCH , the latest collision rate coefficients are available on the Leiden Atomic and Molecular Database (LAMDA, Schöier et al. 2005; Faure et al. 2016; Ben Khalifa et al. 2024). In the case of HCC^{13}CN , we used the rates of HC_3N and scaled them with the corresponding reduced mass of the isotopologue. For

the deuterated isotopologues of HC₃N and CH₃CCH, a scaling of the rates is not possible, as the different nuclear spins induce different hyperfine structures. Therefore, we can not apply the non-LTE modelling to these species.

Results

The abundance profiles that provide the best-fit solutions for HC₃N are shown in Fig. 4.4. A comparison between the modelled spectra and the observed lines is shown in Fig. 4.5 for HC₃N, HCC¹³CN and CH₃CCH. The observed spectra, shown in black, are extracted at the dust peak of the core, averaged over an area corresponding to the telescope beam size (31''). To enable a comparison, the synthetic spectra were convolved with this observational beam for each transition frequency. The molecular column densities for the centre of the core derived by LOC are given in Table 4.2.

The spectrum of HCC¹³CN can be reproduced both with a constant abundance and a radial abundance profile from the chemical modelling. With a depletion radius of 1800 au, a constant abundance of $2.5 \cdot 10^{-11}$ fits best. When applying the abundance profiles from pyRate, the observed lines are reproduced within a factor of 2 with profiles extracted at early timesteps of the chemical simulation ($t_1 = 10^5$ yrs, $t_2 = 1.3 \cdot 10^5$ yrs).

The brightness of the HC₃N spectral line can be reproduced with modelling at a constant fractional abundance of $1.5 \cdot 10^{-9}$ with a depletion radius of 1800 au, and, similar to HC₃N, at early timesteps ($t_1 = 10^5$ yrs, $t_2 = 1.3 \cdot 10^5$ yrs). However, the modelling fails to recreate the slightly asymmetric, redshifted shape of the line profile. This asymmetric shape is likely caused by the velocity gradient across the core with higher velocities in the south, where HC₃N peaks (Bianchi et al. 2023). Therefore, there is more material emitting at slightly higher velocities, compared to the systemic velocity of the core in the centre, causing a redshifted asymmetry. The column densities of HC₃N and HCC¹³CN derived by the best-fit models result in an isotopic ratio of $^{12}\text{C}/^{13}\text{C} = 60 \pm 20$, which is in agreement with the value for the local interstellar medium, $^{12}\text{C}/^{13}\text{C} = 68$ (Milam et al. 2005).

All four observed transitions of CH₃CCH can be reproduced with a constant abundance profile of $1.8 \cdot 10^{-9}$ and a depletion radius of 1800 au. The non-LTE modelling is able to

Table 4.2: Column densities of the models with LOC that provide the best fit (see Fig. 4.5).

Molecule	model	N_{centre} ($\times 10^{12} \text{ cm}^{-2}$)
HC ₃ N	const	38
	radial (t_1)	64
	radial (t_2)	21
HCC ¹³ CN	const	0.63
	radial (t_1)	0.94
	radial (t_2)	0.31
CH ₃ CCH	const	45

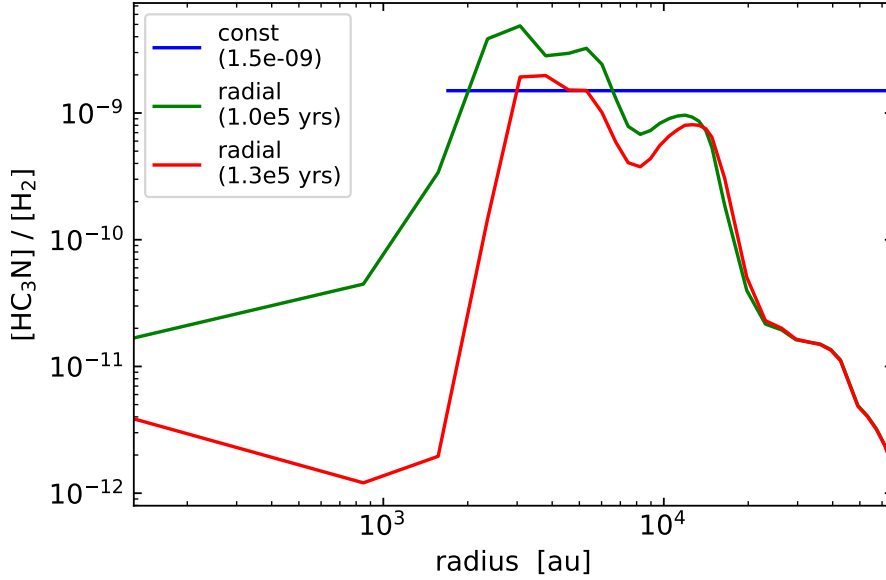


Figure 4.4: Fractional abundance profiles of HC_3N of the best-fit results produced with LOC. For HCC^{13}CN , the radial abundance profiles derived with chemical modelling (green, red) correspond to the profiles of the main isotopologue, scaled down by the carbon isotopic ratio, 68.

reproduce the observed line ratios between the $K=0$ and $K=1$ transitions correctly. Therefore, we are able to use the central column density of CH_3CCH derived by LOC in the further analysis. However, the emission lines cannot be reproduced with any of the abundance profiles derived from chemical modelling. In fact, the radially varying abundance profiles largely underestimate the abundance of CH_3CCH , by roughly three orders of magnitude. This issue was already noted in previous research, where various chemical models failed to predict observed molecular abundances of CH_3CCH . The formation pathways of CH_3CCH have been widely studied, proposing formation via ion-molecule reactions, neutral-neutral reactions, and dissociative recombination in the gas phase (Schiff & Bohme 1979; Turner et al. 1999; Calcutt et al. 2019), and hydrogenation of C_3 on grain surfaces (Hickson et al. 2016; Guzmán et al. 2018). However, models demonstrated that those gas-phase and grain-surface formation pathways are not sufficient to reproduce CH_3CCH abundances in cold molecular clouds (see e.g. Öberg et al. 2013; Hickson et al. 2016). Hence, it is necessary to further investigate potential grain surface reactions of CH_3CCH or alternative gas phase pathways at low temperatures to improve the chemical networks and ultimately the model predictions.

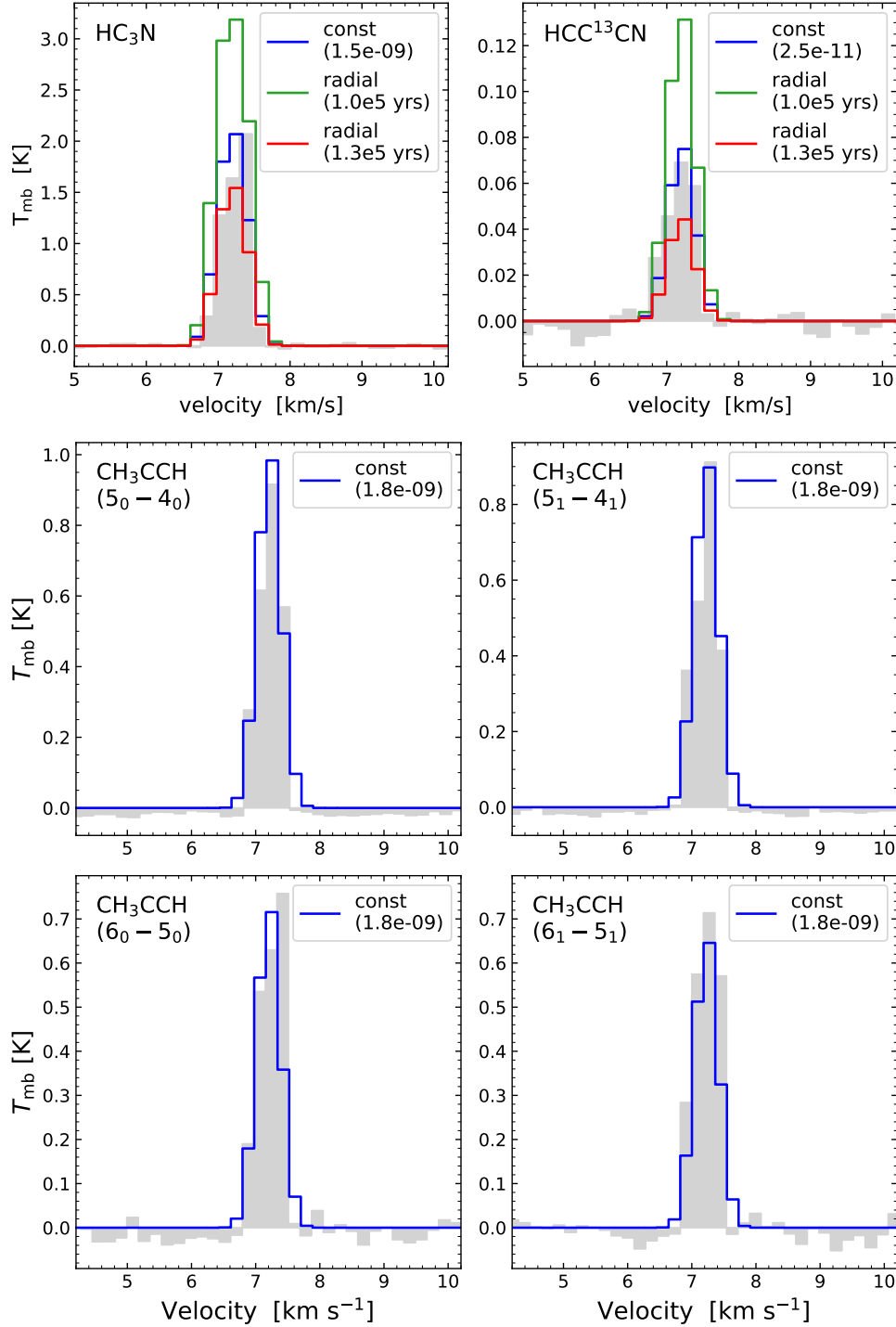


Figure 4.5: Comparison of the observed spectra (grey) of HC_3N , HCC^{13}CN , and CH_3CCH , extracted towards the dust peak of L1544, to the synthetic spectra (blue, green, red) produced with LOC. For the constant abundances, a radius of depletion of 1800 au is assumed.

4.3.3 Column density and deuteration maps

Column density

To derive the corresponding column densities, we assumed optically thin emission, following the derivation presented in Mangum & Shirley (2015). Furthermore, we applied the approximation of a constant excitation temperature across the core (see Caselli et al. 2002b; Redaelli et al. 2019). This results in a column density of:

$$N = \frac{8\pi\nu^3}{c^3} \frac{Q_{\text{rot}}(T_{\text{ex}})}{g_u A_{\text{ul}}} [J_{\nu}(T_{\text{ex}}) - J_{\nu}(T_{\text{bg}})]^{-1} \frac{e^{\frac{E_u}{kT_{\text{ex}}}}}{e^{\frac{h\nu}{kT_{\text{ex}}}} - 1} \int T_{\text{mb}} dv, \quad (4.2)$$

where $Q_{\text{rot}}(T_{\text{ex}})$ is the partition function of the molecule at an excitation temperature T_{ex} , g_u and E_u are the degeneracy and energy of the upper level of the transition, respectively, A_{ul} is the Einstein coefficient for spontaneous emission, $T_{\text{bg}} = 2.73$ K is the temperature of the cosmic microwave background, $J(T)$ is the Rayleigh-Jeans equivalent temperature, and T_{mb} is the main-beam temperature. The corresponding parameters for each transition used in the derivation of the column density are listed in Table 4.1.

As done in Redaelli et al. (2019), the excitation temperatures for CH_3CCH and HCC^{13}CN were chosen to match the respective column density at the dust peak (see Table 4.2) that was derived with a constant molecular abundance profile by the non-LTE modelling in Sec. 4.3.2. Therefore, we applied an excitation temperature of 7 K for HC_3N and 10 K for CH_3CCH , which is in agreement with values used in previous studies (see e.g. Howe et al. 1994; Markwick et al. 2005; Vastel et al. 2014; Hily-Blant et al. 2018; Agúndez et al. 2019; Bianchi et al. 2023). As shown in the previous section, non-LTE effects are responsible for the line intensity ratio observed for CH_3CCH $K=0$ and $K=1$ towards the dust peak. We decided to use the $K=0$ lines to derive the column density map of CH_3CCH , because the column density value of these lines at the dust peak matches with an excitation temperature that is reasonable for L1544. For the $K=1$ lines, however, we would have to use an excitation temperature around 60 K, which is significantly larger than what was routinely observed in this core. An example of the CH_3CCH column density and deuteration maps derived from the $K=1$ transitions is given in Appendix C.1. As the non-LTE approach was not possible for the deuterated isotopologues due to missing collision rate coefficients, we adopted the respective excitation temperature of the main species. To derive the column density of CH_3CCD , we used the $K=0$ transition, similar as for CH_3CCH . The effect of the excitation temperature on the derived column density ratios was found to be small, with a change of a few percent upon a variation of ± 1 K (as stated in Spezzano et al. 2013).

To account for the varying optical depth of the lines across the core, we correct the column density by deriving the optical depth at each pixel individually, using the respective excitation temperature, and then multiplying the maps by (see e.g. Goldsmith & Langer 1999):

$$f_{\tau} = \frac{\tau}{1 - \exp(-\tau)}. \quad (4.3)$$

This results in a correction factor of up to 1.1 in the southern part of the core. The transitions used to derive the column density for each species are marked with an asterisk

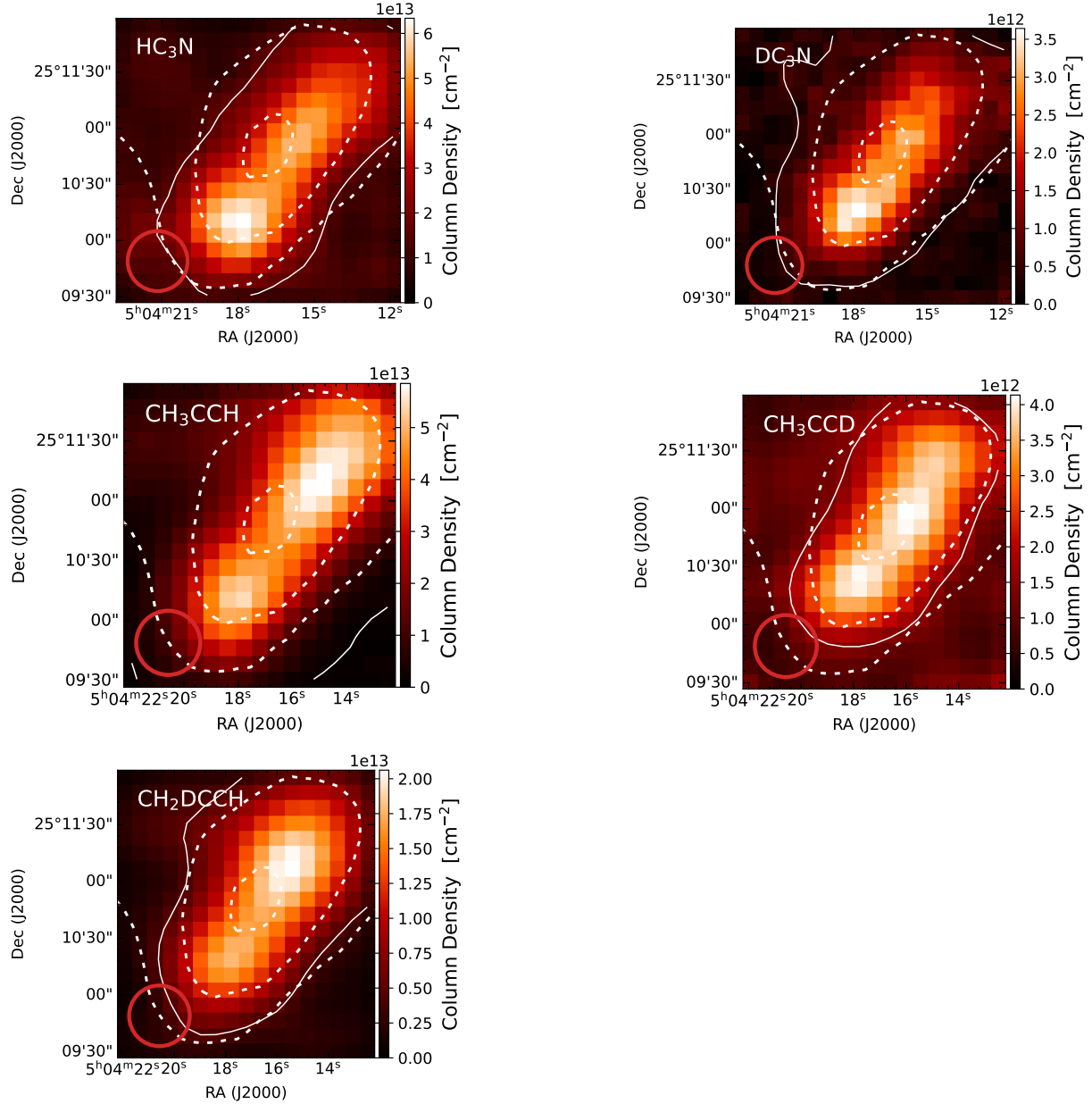


Figure 4.6: Column density maps of HC_3N and CH_3CCH , and their deuterated isotopologues. The column densities are derived with a constant excitation temperature of 7 K and 10 K for HC_3N and CH_3CCH and their isotopologues, respectively, and are corrected for optical depth. To derive the column density of HC_3N , we used the emission of HCC^{13}CN , assuming a constant $^{12}\text{C}/^{13}\text{C}=68$ (Milam et al. 2005). The solid line contours indicate the 3σ level of the integrated intensity. The dashed line contours represent 30%, 50%, and 90% of the H_2 column density peak derived from *Herschel* maps (Spezzano et al. 2016). The circle in the bottom-left corner indicates the beam size of the IRAM 30 m telescope (31'').

in Table 4.1. The corresponding column density maps are presented in Fig. 4.6. In the following, we discuss the results for the each species individually.

HC₃N: The emission of HC₃N is moderately thick in this core, with optical depths of up to 0.9. Therefore, we approximate the column density of HC₃N with the emission of the ¹³C isotopologue, multiplied by the isotopic ratio for the local interstellar medium, ¹²C/¹³C = 68. Given the results of the non-LTE modelling, ¹²C/¹³C = 60 ± 20 at the dust peak, this is a reasonable assumption. The resulting column density map peaks in the south-east of the core, on the carbon-chain peak caused by non-uniform external illumination (see Spezzano et al. 2017), with a maximum value of 6.3(5) × 10¹³ cm⁻², where the number in parenthesis indicates the uncertainty of the last digit. This is slightly lower than previous measurements towards L1544, which is likely due to differences in the excitation temperature (8.0(4) × 10¹³ cm⁻², Hily-Blant et al. 2018; 7(2) × 10¹³ cm⁻², Howe et al. 1994). Additionally, the column density of HC₃N shows a local peak towards the north-west of the core (where CH₃CCH peaks) with a maximum of 5.4(5) × 10¹³ cm⁻², and a dip in the core centre indicating the molecular freeze-out zone. The deuterated isotopologue is distributed similarly to the ¹³C-isotopologue, peaking in the south-east (3.5(2) × 10¹² cm⁻²), and with a local peak in the north-west (3.1(2) × 10¹² cm⁻²). These column density levels are consistent with previous measurements towards L1544 (e.g. Howe et al. 1994).

CH₃CCH: In contrast to HC₃N, the emission of CH₃CCH peaks in the north-west of the core, with a maximum value of 5.85(6) × 10¹³ cm⁻². Here, it is possible that ongoing accretion of material from cloud to core leads to chemically fresh gas and an replenishment of CH₃CCH (see Giers et al. 2025). A weaker, secondary peak is visible in the south-east of the core, on the carbon-chain peak, with values of 5.43(6) × 10¹³ cm⁻². The column density of CH₂DCCH is distributed similarly to the main isotopologue, with a peak column density of 2.0(2) × 10¹³ cm⁻², and a southern peak at 1.7(2) × 10¹³ cm⁻², which is equal to the northern one within errorbars. The column density of CH₃CCD, on the other hand, peaks closer to the core centre (4.1(4) × 10¹² cm⁻²), with a second, equally strong peak (4.0(5) × 10¹² cm⁻²) in the southern part of the core.

Deuterium fraction

The deuterium fraction maps were derived by dividing the column density maps of the deuterated isotopologues by that of the main species. For the ratio of N(CH₂DCCH)/N(CH₃CCH), we accounted for the degeneracy of the D atom by dividing the ratio by a factor of three. The resulting deuteration maps for HC₃N and CH₃CCH are presented in Fig. 4.7. In the following, we describe the results for each molecule individually:

DC₃N/HC₃N: The deuterium fraction map shows moderate levels of deuteration, in the range of 0.04 - 0.07. The distribution is rather homogeneous across the whole core.

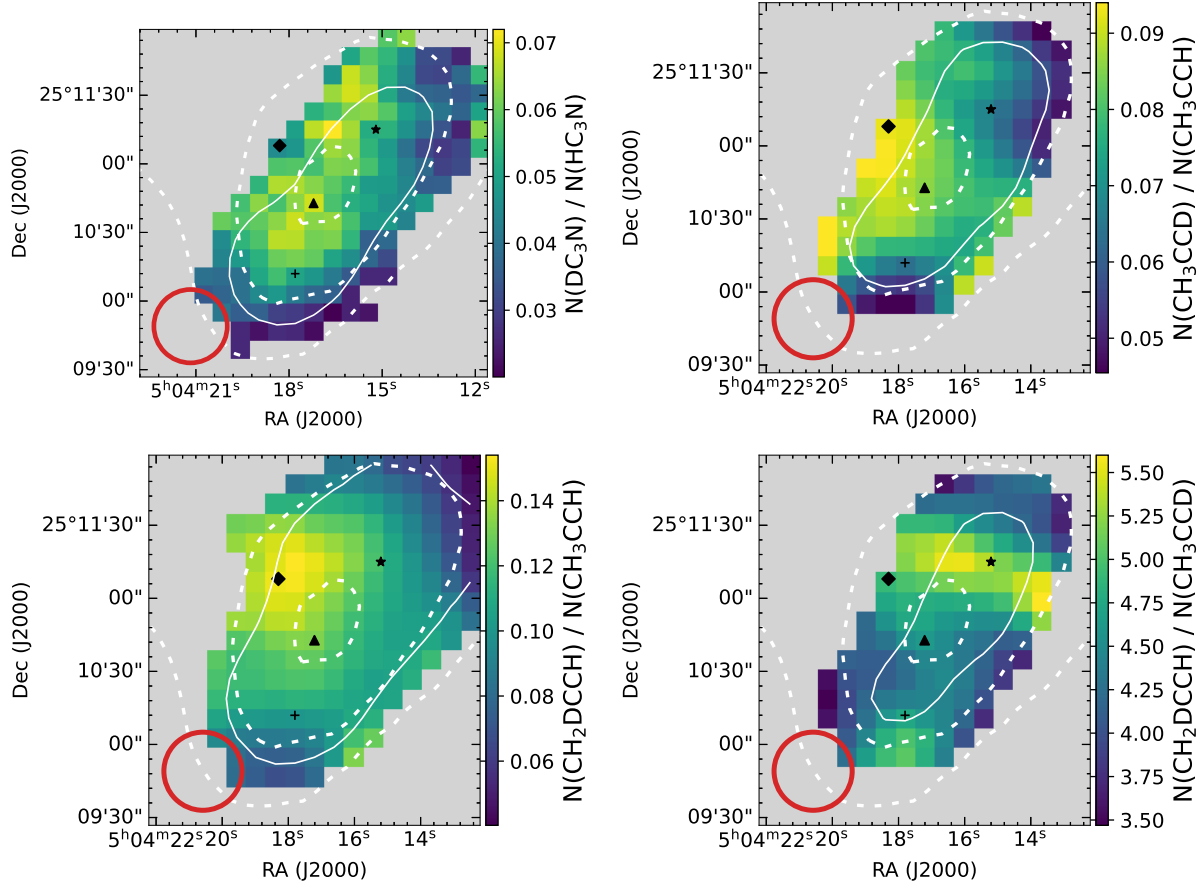


Figure 4.7: Deuteration maps of HC_3N and CH_3CCH . Plotted are only pixels above 3σ . The solid line contour indicates the 5σ confidence level of the ratio. The dashed line contours represent 30, 50, and 90% of the H_2 column density peak derived from *Herschel* maps (Spezzano et al. 2016). The circle in the bottom-left corner indicates the beam size of the IRAM 30 m telescope ($31''$). The markers in black represent the dust peak of L1544 (triangle) and the molecular emission peaks of CH_3OH (diamond), CH_3CCH (star), and $\text{c-C}_3\text{H}_2$ (plus sign).

The highest values (0.06-0.07) are widespread, and located mostly north and south-east of the core centre.

$\text{CH}_3\text{CCD}/\text{CH}_3\text{CCH}$: The map shows that the deuteration is, within the errorbars, distributed rather homogeneously across the core, with values between 0.07 and 0.09. There is a trend visible of higher values (≥ 0.09) concentrated towards the centre and the north-east of the core. The high D/H values at the edges of the map should be treated with caution since they have large uncertainties of around 30%.

CH₂DCCH/CH₃CCH: This deuteration map shows a larger extension across the core than the map of CH₃CCD/CH₃CCH, due to a better signal-to-noise ratio. Additionally, the map exhibits a significantly larger deuterium fraction than CH₃CCD/CH₃CCH, with values between 0.09 and 0.15. In contrast to CH₃CCD/CH₃CCH, the distribution of CH₂DCCH/CH₃CCH shows a clear peak away from the centre towards the north-east of the core, with a peak value of 0.15(3).

CH₂DCCH/CH₃CCD: The ratio map between the two deuterated isotopologues shows that CH₂DCCH is more abundant than CH₃CCD. Within errorbars, the distribution is roughly homogeneous across the core, with a slight trend of larger values towards the north (5.3 ± 0.9) compared to the south (4.1 ± 0.9). The statistically expected ratio of CH₂DCCH/CH₃CCD is three, which results from the statistical considerations for the insertion of deuterium in the CH₃ group and the CCH group. We derive an average ratio of 4.5(8), which is a 50% increase with respect to the statistical value.

4.4 Discussion

4.4.1 HC₃N

The column density maps presented in Fig. 4.6 show that the two carbon chains HC₃N and CH₃CCH and their deuterated counterparts exhibit different morphologies across the core. The emission of both HC₃N and its deuterated isotopologue peak at the carbon-chain peak in the south of the core. The resulting deuterium fraction map, however, shows a rather homogeneous distribution across the core with moderate levels of deuteration. This indicates that HC₃N and DC₃N do not trace the high density regions in the core centre, but a less dense layer of gas in an outer shell of the core. This is supported by the fact that the abundance profiles used to reproduce the emission lines (see Fig. 4.4) peak at a volume density of roughly 10^5 cm^{-3} , which is significantly lower than the critical densities of both transitions ($n_{\text{crit}}(\text{HC}_3\text{N}) \approx 10^7 \text{ cm}^{-3}$, $n_{\text{crit}}(\text{HCC}^{13}\text{CN}) \approx 10^6 \text{ cm}^{-3}$). This was also observed by Bianchi et al. (2023), who state that the carbon-chain abundance is enhanced towards the external part of the core, where the material is more exposed to the interstellar radiation field (ISRF; see also Spezzano et al. 2016). This becomes visible in the slightly asymmetric line shapes of HC₃N and HCC¹³CN at the dust peak and the c-C₃H₂ peak (see Fig. C.2). The asymmetry is caused by the velocity gradient across L1544, which leads to larger velocities at the carbon-chain peak in the southern part (see Spezzano et al. 2016, 2017; Giers et al. 2025), creating redshifted, asymmetric line profiles for carbon-chain molecules, as most of their emission originates in this part. This effect cannot be reproduced with the non-LTE modelling, because the physical structure is assuming the core to be spherically symmetric, and does not take into account asymmetric velocity variations across the core.

The D/H map of HC₃N shows moderate levels of deuteration (0.04-0.07). This is in agreement with previous measurements in L1544 (Howe et al. 1994). Results for the deuterium fraction of HC₃N show roughly similar values in other cores: 0.03-0.11 in TMC-1

(Howe et al. 1994), 0.03(1) in the cold envelope of the class 0 protostar L1527 (Sakai et al. 2009), and 0.03(1) in the dense core L483 (Turner 2001; Sakai et al. 2009; Agúndez et al. 2019). Rivilla et al. (2020) studied DC_3N in a variety of high-mass star-forming clumps, deriving a slightly lower values, ranging from 0.003 to 0.022. They conclude that DC_3N is likely formed by gas-phase ion-molecule reactions in the outer and less dense part of those objects. In the outer envelope, the gas is less dense and deuteration therefore less efficient. In addition, the outer layers are less shielded from the ISRF, which enhances the abundance of atomic carbon and therefore the formation of carbon-chains like HC_3N . The moderate deuteration levels and extended morphology of the D/H map in L1544 gives evidence that also in the low-mass regime, DC_3N is formed in the gas phase.

4.4.2 CH_3CCH

The observed ratios between the peak intensities of the $K=0$ and $K=1$ emission lines of CH_3CCH and CH_3CCD – both for the (5–4) and the (6–5) transitions – cannot be reproduced assuming LTE conditions but instead require a non-LTE treatment. Compared with the observations, the LTE model either overestimates the $K=0$ line or underestimates the $K=1$ line. Low densities in the layers traced by CH_3CCH could cause non-LTE conditions, leading to a subthermal excitation of the transitions, and subsequently generating the emission lines different from the expected LTE profile. Figure 4.2 shows that the line intensity ratios also vary across the core. However, the ratio of the integrated intensity maps of the ($5_0 - 4_0$) over the ($5_1 - 4_1$) transition is roughly constant across the core, with an average value of 1.05(2). Therefore, the effect of the line variations on the derived column density of CH_3CCH and the deuteration maps should be within the errorbars constrained by the derivation. The emission of CH_3CCH peaks in the north-west of L1544, away from the carbon-chain peak in the south-east. The northern part of this core is shaped by the interaction of two filaments (e.g. Spezzano et al. 2016). These structures are believed to transport fresh, chemically young gas into the north/north-west of the core (Giers et al. 2025). This replenishes the abundance of CH_3CCH and causes the molecular peak away from the carbon-chain peak in the south.

In contrast to CH_3CCH , the emission of the deuterated isotopologue CH_3CCD shows two strong peaks, one located in the southern part of the core at the carbon-chain peak, similar to the emission of DC_3N . The second one is close to the dust peak, shifted slightly to the north-west, which coincides with the emission peak of HDCO , singly deuterated H_2CO (see Chacón-Tanarro et al. 2019). The resulting deuterium fraction map of CH_3CCD peaks close to the core centre, slightly offset towards the CH_3OH emission peak in the north-east, with maximum values of 0.09(2). A similar behaviour was also observed for the carbon chain $\text{c-C}_3\text{H}_2$, where the deuteration is most efficient towards the dust peak in the centre of the core (Giers et al. 2022). This, together with the similarities to DC_3N and HDCO , supports the theory that the CCH group of CH_3CCH is deuterated in the gas phase, as was stated in Markwick et al. (2005).

The distribution of CH_2DCCH across L1544 mostly follows the emission of the main species. In the resulting deuteration map, however, the highest values (0.15(3)) are con-

centrated off-centre and towards the north-east of the core, where CH_3OH peaks (see e.g. Spezzano et al. 2016). In this region, CH_3OH traces a local density enhancement (Lin et al. 2022a), a clumpy substructure that introduces a deviation from the Bonnor-Ebert density profile used to describe the core structure in Keto et al. (2015). In fact, the CH_3OH peak seems to be the meeting point of the two larger scale filamentary structures where L1544 is embedded. This intersection of filaments is believed to induce slow collision shocks, causing the local density enhancement (Lin et al. 2022a). The peak of the CH_2DCCH deuteration map overlaps with this higher-density region. The higher density in this area leads to a more efficient freeze out of molecules, including CO, which drives more deuteration in the gas phase (e.g. Caselli et al. 2002b). In addition, the enhanced freeze-out onto grains leads to a higher reactivity and hence drives the reactive desorption of species from dust grains (e.g. Vasyunin et al. 2017). The combination of these effects likely leads to the higher deuteration efficiency that is observed for CH_2DCCH .

In general, CH_2DCCH shows higher levels of deuteration (0.09-0.15) than CH_3CCD (0.07-0.09). Observations by Markwick et al. (2005) show a similar behaviour: Along the TMC-1 ridge, $\text{CH}_3\text{CCD}/\text{CH}_3\text{CCH}$ ranges from 0.04 to 0.10, while $\text{CH}_2\text{DCCH}/\text{CH}_3\text{CCH}$ goes up from 0.08 to 0.17. On the other hand, the dense core L483 and the Class I protostar L1551 IRS5 show equal levels of deuteration for CH_2DCCH and CH_3CCD at moderate values of 0.06 - 0.08 (Agúndez et al. 2019; Marchand et al. 2024). However, Agúndez et al. (2019) find much higher D/H levels of up to 0.18(5) by using the ratio $^{13}\text{CH}_2\text{DCCH}/^{13}\text{CH}_3\text{CCH}$.

A comparison between the two deuterated isotopologues (see right-hand plot in Fig. 4.7) shows that, in L1544, CH_2DCCH is 4.1-5.5 times more abundant than CH_3CCD . In contrast, Markwick et al. (2005) reported a $\text{CH}_2\text{DCCH}/\text{CH}_3\text{CCD}$ ratio ranging between 1.2–2 along the TMC-1 ridge, which is significantly lower than the statistical value of three. The authors attribute these lower values to decreased abundances of H_2D^+ due to ambipolar diffusion, leading to a lower deuterium fractionation. On the other hand, Agúndez et al. (2019) report a ratio of 3 ± 1 for $\text{CH}_2\text{DCCH}/\text{CH}_3\text{CCD}$ in the dense core L483. The authors are able to reproduce the observed abundances of CH_3CCH and its singly and doubly deuterated isotopologues in this core with a pure gas-phase chemical model (Agúndez et al. 2021). From this, they conclude that the formation and deuteration of CH_3CCH in L483 happens in the gas-phase rather than on grain surfaces. They propose that deuterated CH_3CCH is formed by the dissociative recombination of $\text{C}_3\text{H}_5\text{D}^+$ and $\text{C}_3\text{H}_6\text{D}^+$. However, in L1544, the average $\text{CH}_2\text{DCCH}/\text{CH}_3\text{CCD}$ ratio observed is 4.5(8), which is significantly higher than the expected statistical value of three, showing that CH_2DCCH has a higher deuteration efficiency than CH_3CCD in this core. Therefore, there have to be some additional processes that enhance the deuteration of the CH_3CCH methyl group, or decrease the abundance of CH_3CCD . Together with the tentative gradient of $\text{CH}_2\text{DCCH}/\text{CH}_3\text{CCD}$ across the core, this suggests that there are different deuteration mechanisms for the two functional groups with varying efficiency across the core. Actually, a comparison of the emission of deuterated CH_3CCH and of deuterated methanol, shows interesting similarities: both CH_2DCCH and CH_2DOH peak in the northern part of L1544, close to the dust peak (see Chacón-Tanarro et al. 2019), and the deuteration in the methyl group is

significantly larger than in the CCH or the OH group ($\text{CH}_2\text{DOH}/\text{CH}_3\text{OD} \geq 10$, Bizzocchi et al. 2014). The root of this might be a common formation path of CH_2DCCH and CH_2DOH on the surfaces of dust grains, which was not considered in the pure gas-phase model used in Agúndez et al. (2021). The similarities between the CH_3CCD and HDCO emission provides another hint, as H_2CO also can be deuterated both in gas phase and on grain surfaces. However, extensive chemical models, that incorporate both gas-phase and grain-surface reactions of large hydrocarbons, are necessary to study the chemical processes and clarify the relevance of these observed trends for the isotopic fractionation of the same functional group in different molecules.

4.5 Conclusions

In this work, we studied the deuterium fraction of the carbon chains HC_3N and CH_3CCH in the prestellar core L1544. We analysed the emission of the species together with their ^{13}C -bearing and deuterated isotopologues. Using non-LTE calculations, we constrained and derived the column densities of the different species. The corresponding deuterium fraction maps show different behaviour in terms of distribution and level of deuteration, and seem to trace different physical conditions.

HC_3N shows moderate levels of deuteration, with a very homogeneous and extended distribution. This suggests that HC_3N and DC_3N trace the outer layers of the envelope of the core with intermediate-density gas rather than the dense core centre, and supports a gas-phase formation of DC_3N . With non-LTE modelling, we derive a $^{12}\text{C}/^{13}\text{C}$ ratio of 60 ± 20 for HC_3N towards the dust peak of L1544, which is consistent with the value for the local ISM, 68. CH_2DCCH shows a clear deuteration peak away from the core centre, towards the north-east, coinciding with the CH_3OH molecular peak. This is likely linked to the local density enhancement in this region that enhances the CO freeze out and therefore increases deuteration and reactive desorption. CH_3CCD , on the other hand, seems to be most efficiently formed closer to the centre of the core, where deuteration is enhanced also in other carbon chain such as $\text{c-C}_3\text{H}_2$. CH_2DCCH shows significantly higher levels of deuteration than CH_3CCD , and is about 4-5 times more abundant across the core, similar to what was observed for deuterated CH_3OH . This, together with the tentative trend of higher values towards the north, suggests that there are different deuteration mechanisms for the two functional groups (CH_3 and CCH) in L1544, with varying efficiency across the core.

Overall, the results of this work suggest that gas-phase reactions dominate the formation and deuteration of carbon chains in L1544, with spatial variations driven by physical structure, density, and external radiation. However, given the similarities of the deuterium fraction of CH_3CCH and CH_3OH , there might be an additional deuteration pathway of CH_3CCH on the surfaces of dust grains. This shows the importance of understanding the chemical processes happening when using deuterated molecules as tools to study the interface between the dense core and the surrounding cloud.

Chapter 5

Conclusions

In this thesis, I presented the results of three observational projects focused on the chemical segregation in the earliest stages of star formation. I used radio observations of various molecular tracers, analysing their distributions across different cores and the influence of the environment on the chemical structures present. This chapter summarises the main results of each project and discusses possible future directions in this field of research.

5.1 Summary of this thesis

In Chapter 2, I compared the level of deuteration of the four simple molecules CCH, HCN, HNC, and HCO^+ in the envelopes of the prestellar core L1544 and the protostellar core HH211, using data observed with the Onsala 20 m radio telescope. I derive the molecular column densities and deuteration levels by using both the ^{13}C - and ^{18}O -bearing isotopologues. I find similar levels of deuteration in L1544 and HH211, showing that the deuterium fractionation is most likely equally efficient towards both cores. This suggests that the protostellar envelope of HH211 still retains the chemical composition of the original prestellar core. Due to the fact the two cores are embedded in different molecular clouds but still show similar levels of deuteration, I conclude that the environmental conditions do not have a significant effect on the deuterium fractionation within dense cores. Furthermore, this work highlights the uncertainties when dealing with ^{13}C isotopologues and the influence of the applied carbon isotopic ratio on the derived deuteration levels. With a combination of non-LTE radiative transfer modelling and abundances from chemical simulations, I reproduced the observed emission lines, testing new collisional rate coefficients for HNC, HN^{13}C , DNC, and DCN that include the hyperfine structure of the rotational transitions of the molecules. The results show that it is crucial to take into account the hyperfine structures of the molecules to reproduce the observed line shapes. In addition, for molecules such as HCN, it is necessary to include the outer layers of the core to the physical model, to consider the effects of extended structures and to correctly model the emission lines towards the prestellar core.

In Chapter 3, I analysed the chemical segregation of $\text{c-C}_3\text{H}_2$, CH_3OH , and CH_3CCH

in a sample of starless and prestellar cores with unsupervised machine learning, using data observed with the IRAM 30 m radio telescope. I applied the density-based clustering algorithms DBSCAN and HDBSCAN to identify chemical and physical structures within these cores. To enable cross-core comparisons, I characterised the clustering input samples based on their physical environment, discarding the two-dimensional spatial information. With the clustering analysis, I found significant chemical differentiation across the core, reproducing the known segregation between $c\text{-C}_3\text{H}_2$ and CH_3OH and identifying a non-apparent differentiation between $c\text{-C}_3\text{H}_2$ and CH_3CCH . Different environmental conditions of the individual cores were reflected by variations in the relevance of the input features in the respective clustering results. From similarities in the clustering patterns of CH_3OH and CH_3CCH relative to $c\text{-C}_3\text{H}_2$, I concluded that $c\text{-C}_3\text{H}_2$ traces more outer layers or lower-density regions than the other two molecules. In addition, I suggested that in the prestellar core L1544, CH_3CCH traces a landing point of chemically fresh gas that is accreted to the core, highlighting the impact of accretion processes on molecular distributions. This study shows that density-based clustering can provide valuable insights into the chemical and physical structures of starless cores. It demonstrates that already small datasets covering only two or three molecules can yield meaningful results.

In Chapter 4, I studied the deuteration of the two carbon chains HC_3N and CH_3CCH in the prestellar core L1544, using data observed with the IRAM 30 m radio telescope. With non-LTE radiative transfer modelling, I reproduced the spectra of HC_3N , HCC^{13}CN , and CH_3CCH observed towards the dust peak of the core. Using those results, I constrained the excitation temperatures of the observed species to derive the molecular column densities and the deuteration maps. I found moderate levels of deuteration for DC_3N and CH_3CCD , and significantly larger values for CH_2DCCH . The results suggest that gas-phase reactions dominate the formation and deuteration of both HC_3N and CH_3CCH in L1544, with spatial variations driven by differences in physical structure, density, and external radiation. From the significantly larger deuteration fraction of CH_2DCCH compared to CH_3CCD , I suggested that there might be different deuteration mechanisms for the two functional groups, with varying efficiency across the core. Similarities between the deuteration of CH_3CCH and CH_3OH might indicate an additional deuteration pathway of CH_3CCH on the surfaces of dust grains. This work shows the importance of understanding the chemical processes happening when using deuterated molecules as tools to study the interface between the dense core and the surrounding cloud.

5.2 Outlook and future prospects

In my work, I focus on the analysis of chemical segregation in starless, prestellar, and protostellar cores by using deuteration, radiative transfer, and density-based clustering. The projects presented in this thesis give multiple opportunities for further research and deeper discussion. In the following, I will discuss some options.

In Chapter 2, I compared the deuteration of simple molecules in a prestellar and a protostellar core. However, those observations were limited to single-pointing data, where

the large 40'' beam of the Onsala 20 m radio telescope covers an extended area within the cores, averaging out spatial variations. On the other hand, the deuteration maps presented in Chapter 4 reveal significant differences in the distribution and efficiency of deuterated species across the prestellar core L1544. This highlights the importance of molecular emission maps for understanding extended chemical structures in such cores. Contrary to single-pointing spectra, emission maps can provide crucial information about how the physical structure and surrounding environment of a core influence chemical processes such as deuteration. Possible target species to map would for example be the molecules studied in this project (CCH, HCN, HNC, and HCO^+), and their deuterated counterparts (CCD, DCN, DNC, and DCO^+). For HH211, there is generally not much information available about the distribution of molecules across the envelope of the core, as most studies focus on the prominent jet instead or only use single-pointing observations. Molecular emission maps towards this core could for example help tracking down the exact origin of the second velocity component seen in the observed spectral lines, which is currently attributed to large-scale emission extending north-east of the core (see Chapter 2). In general, emission maps covering the extended core would provide further possibilities to study similarities and differences between the chemical structures and segregation in prestellar and protostellar envelopes. For the prestellar core L1544, the morphology and deuteration efficiency of CCD would complement existing data on other carbon chains such as HC_3N or CH_3CCH , which are analysed in this thesis (see Chapter 4). Meanwhile, additional data on the distribution of DCN would enhance our understanding of N-bearing molecules in L1544 and put constraints on their formation and deuteration. This would complement the existing knowledge on nitrogen fractionation and deuteration in L1544 (e.g. Hirota et al. 2003; Redaelli et al. 2019, 2023), where the former is known to be enhanced towards the southern part of the core (Spezzano et al. 2022b).

In general, isotopic fractionation, and particularly deuterium enrichment, is a powerful tool for studying the formation pathways of molecules in the cold and dense conditions of prestellar cores. Fractionation processes are sensitive to temperature, density, and ionization, and can therefore provide insights whether a molecule is formed in the gas phase, on grain surfaces, or through a combination of both. By comparing observed ratios to chemical models, molecular formation reactions can be constrained. This approach could be useful for CH_3CCH , where observations of the doubly-deuterated isotopologues could provide further details on the formation and deuteration of this molecule in L1544 (see Chapter 4), shedding light on possible grain surface routes. Other isotopologues, including isotopes of C, O, or N are useful to constrain the elemental and isotopic composition of the material that will later form stars and planets. Comparisons of molecular content of objects in different evolutionary stages provide insight into the level of inheritance of the chemical complexity across the process of star formation. An extension of the survey analysed in Chapter 2 could add data about the deuteration of other, for example more complex molecules, and cover more cores in the starless, prestellar, or protostellar stage. This would help to better understand the role of the environment on the chemistry and if the observed trend of similar levels of deuteration is specific to the studied two cores L1544 and HH211 or if it can also be found in other cores.

The work presented in chapters 1 and 3 has highlighted how non-LTE radiative transfer modelling is a very useful tool to better understand the link between the physical and chemical structure of a source. Unfortunately, the non-LTE modelling of a species is limited by the availability of its collisional rate coefficients. Collisional rates are publicly available for more than 100 atoms and molecules¹, but especially for deuterated species and/or for larger, more complex molecules, the rates are still missing. Collisional rate coefficients would for example be interesting for the singly (and doubly) deuterated isotopologues of CH₃CCH studied in Chapter 4. This could shed light on the question whether the different functional groups have different deuteration mechanisms and therefore trace different conditions and densities. A collaboration with theorists will provide the collisional rates of singly and doubly deuterated H₂CS in the near future. These molecules were detected in L1544 and analysed by Spezzano et al. (2022a), but chemical models failed to reproduce the abundance of D₂CS. With the combination of the corresponding collisional rates, and chemical and radiative-transfer modelling, we will be able to better constrain the abundances and reaction pathways of H₂CS and isotopologues. Collaborations of observers with theoretical groups are a win-win situation and present opportunities to not only expand the pool of existing collisional rates for observers but also give theoreticians access to observational data. For the community, the radiative transfer modelling of new molecules will be useful to provide additional constraints on the physical structure models of cores and on chemical simulations.

The second project of this thesis (see Chapter 3) focuses on analysing chemical structures with the help of clustering algorithms. In this study, I compared the spatial distribution of three molecules (c-C₃H₂, CH₃OH, CH₃CCH) across three objects (starless: B68, L1521E; prestellar: L1544). A logical next step for this project is to expand the studied sample by including a broader variety of molecules and additional cores. The data used for the project are part of a larger dataset that contains observations of four additional prestellar cores (HMM-1, L429, L694-2, and OphD) with up to 20 detected molecules per core, which makes them promising core candidates for a clustering analysis. Since those cores are located in different molecular clouds than the currently studied objects, they will allow to better assess the influence of the environment of a core on its chemical structures. Furthermore, the core sample will provide more information about what physical or chemical structures individual molecules are most sensitive to. The extended dataset including the additional prestellar cores contains molecules such as HC₃N, CCH, C₄H, H₂CCO, HNCO, HCO⁺, HC¹⁸O⁺, H¹³CO⁺, HCN, HNC, HN¹³C, ¹³CN, N₂H⁺, CCS, C³⁴S, HCS⁺, OCS, ³⁴SO, and SO₂. Nitrogen-bearing species such as N₂H⁺ or ¹³CN are particularly interesting for a clustering analysis because they are known to trace the higher-density regions in the centre of cores. This will provide complementary information to the existing analysis, as the current sample only covers O-bearing (CH₃OH) and carbon-chain molecules (c-C₃H₂, CH₃CCH), which do not trace the highest densities. In a different approach, a clustering analysis could compare the distribution of molecules from the same molecular family,

¹For example in the online databases LAMDA (home.strw.leidenuniv.nl/~moldata/) or EMAA (emaa.osug.fr/).

for example different C-bearing, O-bearing, or S-bearing molecules. Molecules within the same molecular family usually exhibit similar emission morphologies in a core but might still trace different underlying chemical or physical conditions that are not immediately apparent from visual inspection alone. A clustering analysis could help to identify further chemical segregation and find correlations or differences in chemically related species.

A major part of this thesis focuses on the prestellar core L1544. This core is a well-studied object and for over 30 molecules there are data available showing the distribution of the emission across the core (e.g. Spezzano et al. 2017). Despite its status as a prototypical core in a relatively isolated and stable environment, studies in recent years have revealed that L1544 is much more complex than it seems. For instance, the core is embedded in two perpendicular structures that meet in the north-east of the core, causing a weak collision shock and a local density enhancement traced by CH_3OH (Bizzocchi et al. 2014; Spezzano et al. 2016; Lin et al. 2022a). In the north-west of the core, incoming fresh gas replenishes chemically young species such as CH_3CCH (Giers et al. 2025, see Chapter 3). Another notable feature is how the uneven external illumination of the core leads to a chemical segregation of C- and O-bearing molecules (Spezzano et al. 2016; see also Chapter 3). In general, the comprehensive pre-existing knowledge about the chemical and physical structures of the core makes it easier to set observations of new molecules into context and to understand their chemical connections and environmental influences. Therefore, continued mapping observations of molecules, for example of deuterated species (e.g. DNCO , DCS^+) or of increasingly complex molecules (e.g. the ones detected by Jiménez-Serra et al. 2016, such as CH_3OCH_3 or NH_2CHO), would further improve the understanding of this core. In recent years, observations of COMs have become more frequent and accessible, providing new possibilities to study the more complex chemical structures and to hunt for the detection of prebiotic species that shed light on the age-old question of the origins of life. The continued study of L1544 provides the opportunity for the core to serve as a model example for the chemical processes occurring under prestellar conditions, helping to unravel the link between chemical and physical structures in low-mass star-forming regions.

Appendix A

Supplementary material for Chapter 2

Section A.3 was contributed by the collaborators and co-authors Cheikh T. Bop and Francois Lique.

A.1 Observed lines

Table A.1 lists all emission lines observed towards L1544 and HH211, along with the respective transition, frequency and telescope beam size. Fig. A.1 and Fig. A.2 show the corresponding spectra for L1544 and HH211, respectively, with the hyperfine transition given in the top left corner.

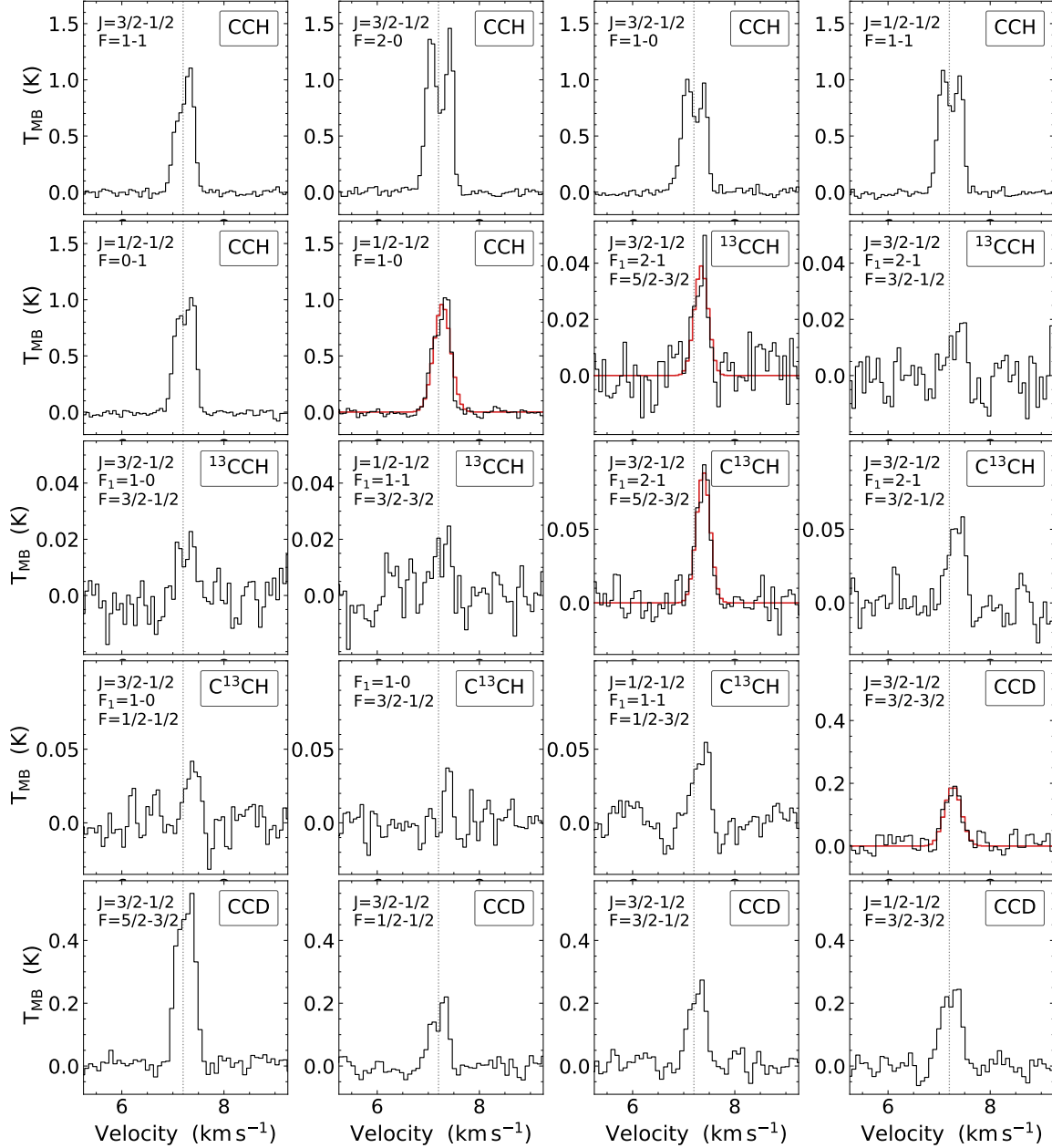


Figure A.1: Ground-state rotational lines observed towards the pre-stellar core L1544. The respective hyperfine transition is given in the upper left corner of each plot. The dotted line indicates the rest velocity of the system (7.2 km s^{-1}). The red curves show the fits used in the LTE analysis.

Table A.1: Properties of the observed lines.

Molecule	Transition	Relative Intensity	Frequency (MHz)	HPBW	Ref.
CCH	$N = 1 - 0, J = 3/2 - 1/2, F = 1 - 1$	0.043	87284.11(1)	43.5	1
	$N = 1 - 0, J = 3/2 - 1/2, F = 2 - 0$	0.417	87316.90(1)	43.5	1
	$N = 1 - 0, J = 3/2 - 1/2, F = 1 - 0$	0.208	87328.59(1)	43.5	1
	$N = 1 - 0, J = 1/2 - 1/2, F = 1 - 1$	0.208	87401.99(1)	43.5	1
	$N = 1 - 0, J = 1/2 - 1/2, F = 0 - 1$	0.083	87407.17(1)	43.5	1
	$N = 1 - 0, J = 1/2 - 1/2, F = 1 - 0$	0.043	87446.47(1)	43.5	1
^{13}CCH	$N = 1 - 0, J = 3/2 - 1/2, F_1 = 2 - 1, F = 5/2 - 3/2$	0.250	84119.33(2)	45.9	2
	$N = 1 - 0, J = 3/2 - 1/2, F_1 = 2 - 1, F = 3/2 - 1/2$	0.153	84124.14(2)	45.9	2
	$N = 1 - 0, J = 3/2 - 1/2, F_1 = 1 - 0, F = 3/2 - 1/2$	0.167	84153.31(2)	45.9	2
	$N = 1 - 0, J = 1/2 - 1/2, F_1 = 1 - 1, F = 3/2 - 3/2$	0.152	84206.87(2)	45.8	2
C^{13}CH	$N = 1 - 0, J = 3/2 - 1/2, F_1 = 2 - 1, F = 5/2 - 3/2$	0.263	85229.335(4)	44.8	2
	$N = 1 - 0, J = 3/2 - 1/2, F_1 = 2 - 1, F = 3/2 - 1/2$	0.165	85232.805(4)	44.8	2
	$N = 1 - 0, J = 3/2 - 1/2, F_1 = 1 - 0, F = 1/2 - 1/2$	0.085	85247.728(4)	44.8	2
	$N = 1 - 0, J = 3/2 - 1/2, F_1 = 1 - 0, F = 3/2 - 1/2$	0.168	85256.988(4)	44.8	2
	$N = 1 - 0, J = 1/2 - 1/2, F_1 = 1 - 1, F = 1/2 - 3/2$	0.081	85303.990(4)	44.7	2
	$N = 1 - 0, J = 1/2 - 1/2, F_1 = 1 - 1, F = 3/2 - 3/2$	0.158	85307.459(4)	44.7	2
CCD	$N = 1 - 0, J = 3/2 - 1/2, F = 3/2 - 3/2$	0.085	72101.811(5)	58.2	3
	$N = 1 - 0, J = 3/2 - 1/2, F = 5/2 - 3/2$	0.333	72107.721(3)	58.2	3
	$N = 1 - 0, J = 3/2 - 1/2, F = 1/2 - 1/2$	0.101	72109.050(5)	58.2	3
	$N = 1 - 0, J = 3/2 - 1/2, F = 3/2 - 1/2$	0.137	72112.295(5)	58.2	3
	$N = 1 - 0, J = 1/2 - 1/2, F = 3/2 - 3/2$	0.137	72187.708(5)	58.1	3
	$N = 1 - 0, J = 1/2 - 1/2, F = 1/2 - 3/2$	0.101	72189.726(6)	58.1	3
HCN	$N = 1 - 0, J = 1/2 - 1/2, F = 3/2 - 1/2$	0.085	72198.193(6)	58.1	3
	$J = 1 - 0, F = 1 - 1$	0.333	88630.4156(2)	43.0	4
	$J = 1 - 0, F = 2 - 1$	0.556	88631.8475(3)	43.0	4
	$J = 1 - 0, F = 0 - 1$	0.111	88633.9357(3)	43.0	4
H^{13}CN	$J = 1 - 0, F = 1 - 1$	0.333	86338.7352(1)	43.9	5
	$J = 1 - 0, F = 1 - 0$	0.556	86340.1666(1)	43.9	5
	$J = 1 - 0, F = 0 - 1$	0.111	86342.2543(3)	43.9	5
DCN	$J = 1 - 0, F = 1 - 1$	0.333	72413.50(1)	57.9	6
	$J = 1 - 0, F = 2 - 1$	0.556	72414.93(1)	57.9	6
	$J = 1 - 0, F = 0 - 1$	0.111	72417.03(1)	57.9	6
HCO^+	$J = 1 - 0$	-	89188.525(4)	42.8	7
H^{13}CO^+	$J = 1 - 0$	-	86754.288(5)	43.7	8
HC^{18}O^+	$J = 1 - 0$	-	85162.223(5)	44.9	8
DCO^+	$J = 1 - 0$	-	72039.3124(8)	58.3	9
HNC	$J = 1 - 0$	-	90663.568(4)	42.2	10
HN^{13}C	$J = 1 - 0$	-	87090.825(4)	43.6	9
DNC	$J = 1 - 0$	-	76305.700(1)	53.9	9

References. (1) Padovani et al. (2009); (2) McCarthy et al. (1995); (3) Cabezas et al. (2021); (4) Ahrens et al. (2002); (5) Fuchs et al. (2004); (6) Brünken et al. (2004); (7) Tinti et al. (2007); (8) Schmid-Burgk et al. (2004); (9) van der Tak et al. (2009); (10) Saykally et al. (1976).

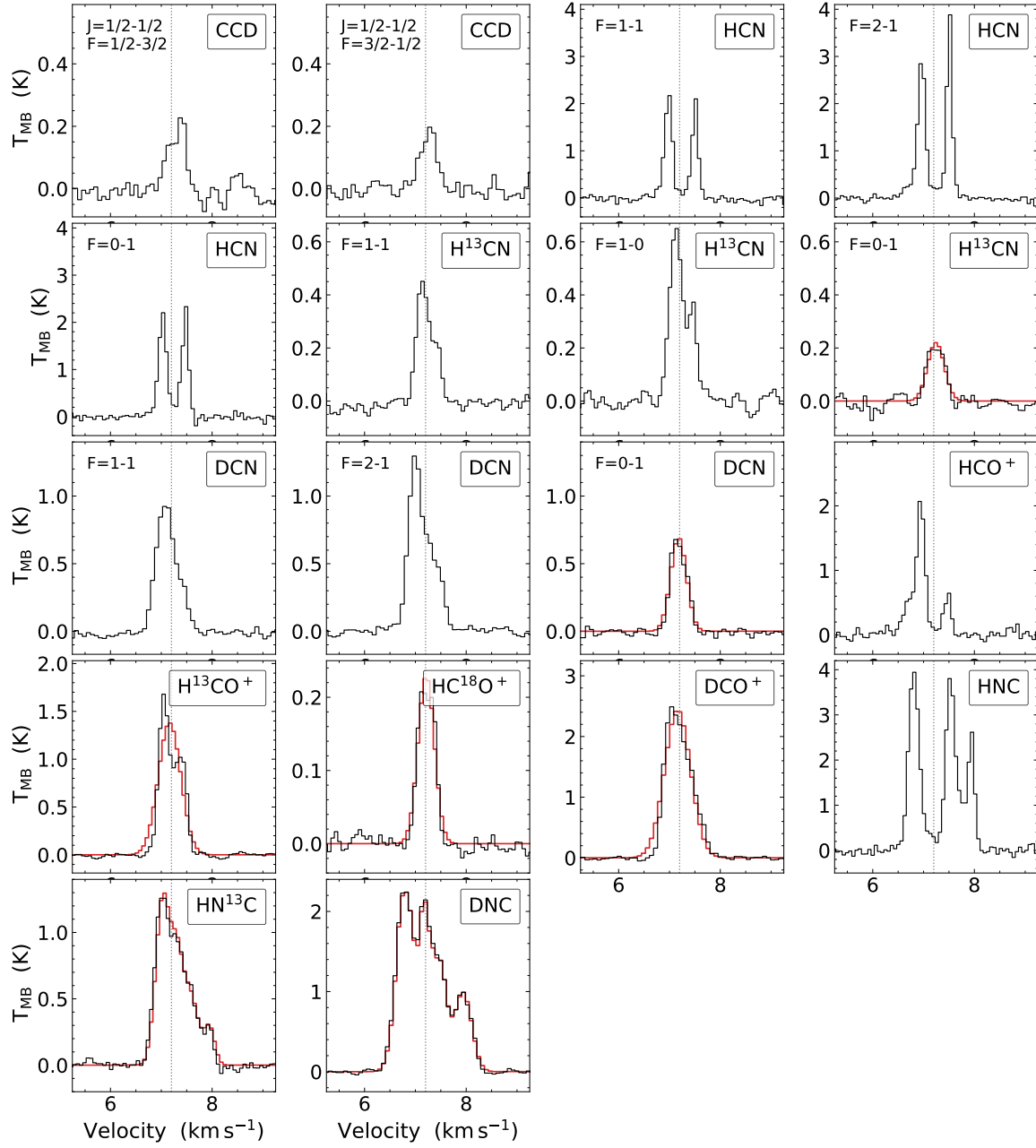


Figure A.1: continued.

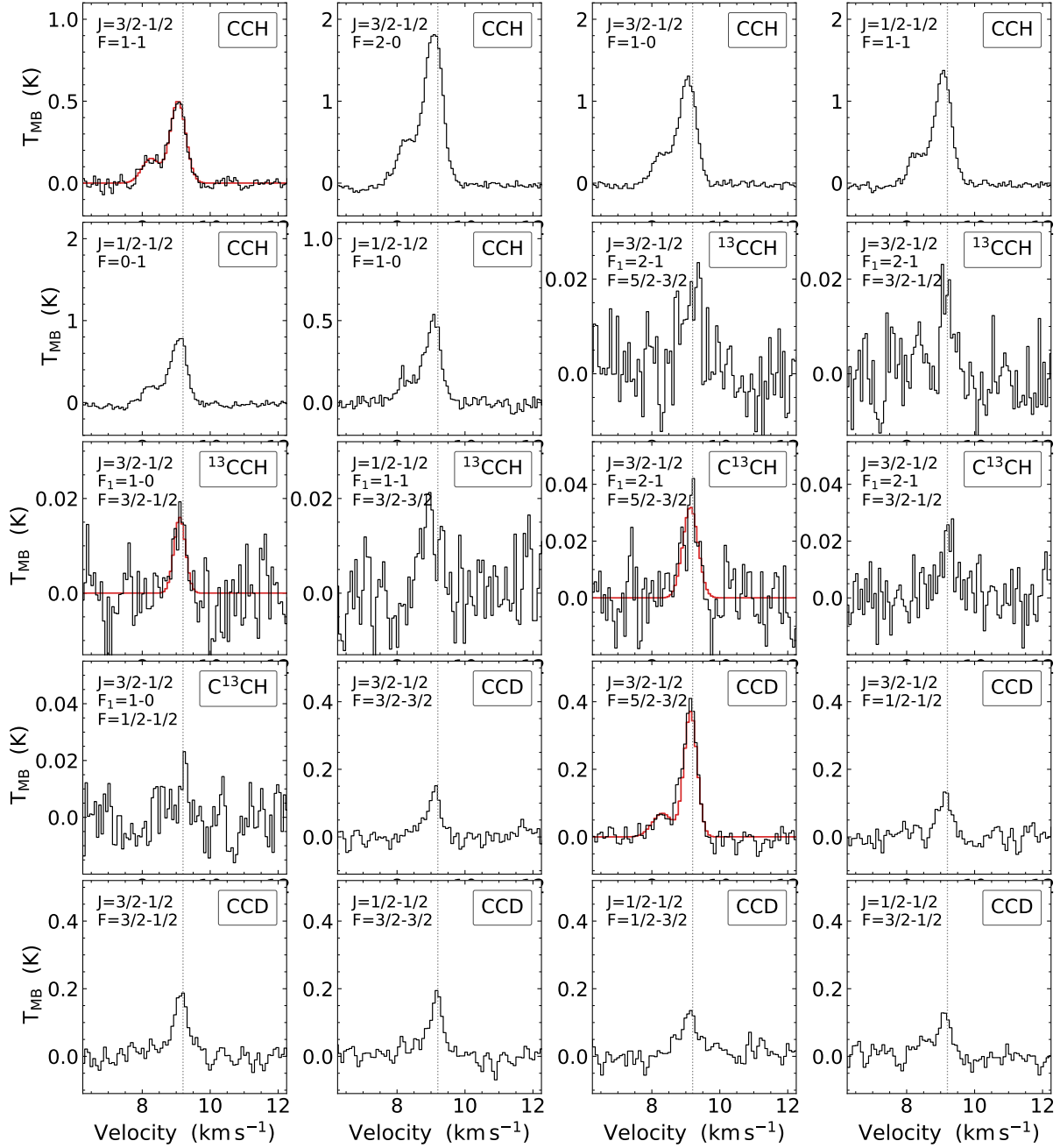


Figure A.2: Ground-state rotational lines observed towards the protostellar core HH211. The respective hyperfine transition is given in the upper left corner of each plot. The dotted line indicates the rest velocity of the system (9.2 km s^{-1}). The red curves show the fits used in the LTE analysis.

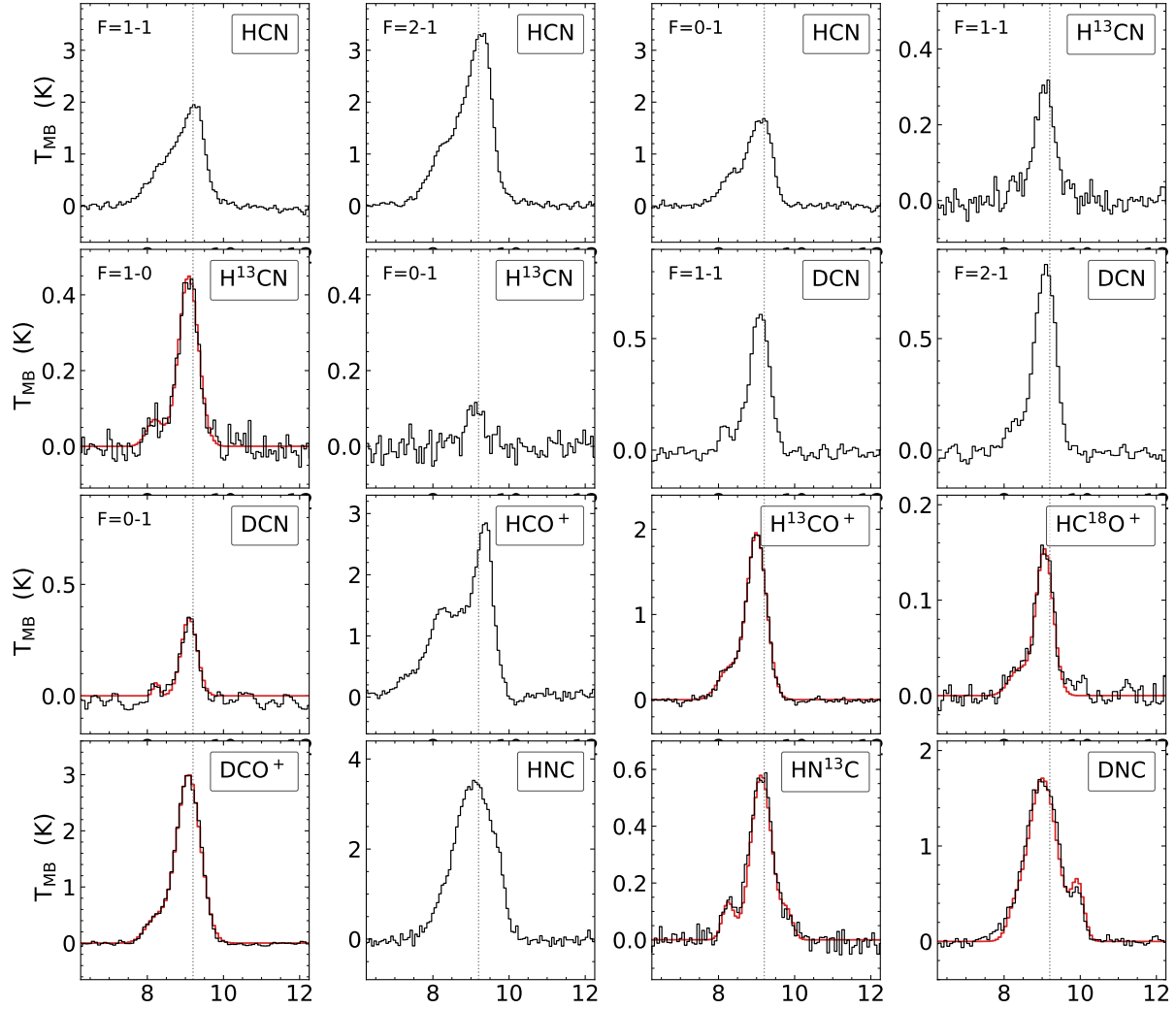


Figure A.2: continued.

A.2 Parameters of the additional velocity component towards HH211

Table A.2 presents the best-fit parameters of the additional velocity component observed towards HH211, along with the corresponding column density. The parameters were obtained by applying a two-component Gaussian fit to the spectral lines. In the case of HN^{13}C and DNC, we applied a three-component Gaussian to account for their more separate hyperfine component at high velocity. The additional velocity component is located at approximately 8.22 km s^{-1} .

Table A.3 compares the level of deuteration between the additional component (add.) and the main component. Within errorbars, the deuterium fractionation is similar for CCH, HCN, and HNC, indicating that main and additional component trace the same gas along the line-of-sight. For HCO^+ , the D/H ratios differ by a factor of two between the two components. However, with this little amount of data on the additional velocity component, we cannot make any conclusive statement about this. Most likely, the additional velocity components of HCO^+ and isotopologues are tracing the same gas as the other molecules in this study, plus some additional gas or structure.

Table A.2: Best-fit parameters and column densities of the additional velocity component observed towards HH211.

Molecule	$T_{\text{mb,peak}}$ (K)	V_{LSR}^a (km s^{-1})	Δv (km s^{-1})	$\int T_{\text{mb}} dv$ (K km s^{-1})	N ($\times 10^{12} \text{ cm}^{-2}$)
CCH	0.15(1)	8.22	0.62(9)	0.10(2)	70(20)
CCD	0.070(3)	8.27	0.7(1)	0.053(9)	6(1)
H^{13}CN	0.07(2)	8.16	0.5(2)	0.04(2)	0.17(8)
DCN	0.06(7)	8.20	0.2(3)	0.01(2)	0.5(1.0)
H^{13}CO^+	0.36(1)	8.24	0.66(5)	0.25(2)	0.30(5)
HC^{18}O^+	0.027(4)	8.35	0.6(2)	0.017(5)	0.020(7)
DCO^+	0.43(2)	8.18	0.65(4)	0.30(2)	0.46(7)
HN^{13}C	0.13(2)	8.22	0.36(6)	0.05(1)	0.15(4)
DNC	0.40(7)	8.22	0.6(2)	0.25(8)	1.0(4)

Notes. ^(a) The uncertainties of the fits are smaller than the velocity resolution. Therefore, the error on V_{LSR} is given by the observed spectral resolution, 0.07 km s^{-1} .

Table A.3: Comparison of column density ratios in the main component and the additional velocity component observed towards HH211.

	HH211 (main)	HH211 (add.)
$N(\text{CCD})/N(\text{CCH})$	0.10(3)	0.09(3)
$N(\text{DCN})/N(\text{H}^{13}\text{CN} \times 68)$	0.07(2)	0.04(8)
$N(\text{DCO}^+)/N(\text{H}^{13}\text{CO}^+ \times 68)$	0.037(5)	0.022(5)
$N(\text{DCO}^+)/N(\text{HC}^{18}\text{O}^+ \times 557)$	0.09(2)	0.04(2)
$N(\text{DNC})/N(\text{HN}^{13}\text{C} \times 68)$	0.08(3)	0.10(4)

A.3 Collisional rate coefficients

Within the Born-Oppenheimer approximation, the interaction potential is the same for HCN-H₂ and DCN-H₂ (HNC-H₂, DNC-H₂, HN¹³C-H₂) and only depends on the mutual distances of the atoms. Hence, the only difference between the interaction potentials of two different isotopologues with H₂ is the position of the centre of mass taken for the origin of the Jacobi coordinates used to describe the geometries of the system. Then, we employed the HCN-H₂ (Denis-Alpizar et al. 2013) and HNC-H₂ (Dumouchel et al. 2011) potential energy surfaces (PESs) corrected to consider the effect of isotopic substitution.

In the following, the rotational level of the targets (DCN, HNC, HN¹³C, and DNC) will be denoted j . Only the coupling of the molecular rotation with the nuclear spin ($I = 1$) of the nitrogen atom will be considered. The coupling results in a weak splitting of each rotational level j , into three hyperfine levels (except for the $j = 0$ level which is split into a single level). Each hyperfine level is designated by a quantum number F ($F = I + j$) varying between $|I - j|$ and $I + j$. Only collisional excitation by para-H₂ in its ground rotational state has been considered.

Rotational (i.e. nuclear-spin free) scattering matrix, cross sections and rate coefficients for the excitation of DCN, HNC, HN¹³C and DNC (the isotopologues of HCN and HNC) induced by collisions with H₂ were computed by Navarro-Almaida et al. (2022) for temperatures up to 30 K. These data were calculated using the quantum mechanical close-coupling approach (Green 1975). For more computational details, we refer the readers to the work of Navarro-Almaida et al. (2022).

To account for the hyperfine structure due to the nitrogen nuclear spin of the isotopologues mentioned above, we used the scattering matrix computed by Navarro-Almaida et al. (2022) and applied a recoupling method (Alexander & Dagdigan 1985; Lanza & Lique 2014). Therefore, we derived hyperfine resolved rate coefficients for the 25 low-lying energy levels, that is $(j, F) \leq (8, 9)$, for temperatures up to 30 K.

Fig. A.3 displays the temperature dependence of the hyperfine resolved excitation rate coefficients of HCN, DCN, HNC, DNC and HN¹³C induced by collision with H₂ for the $(j, F) = (2, 1) \rightarrow (1, F')$ transitions, where $F' = 0, 1, 2$. The upper panels show that the rate coefficients for DCN-H₂ collisions exhibit systematically a higher magnitude than those

for HCN-H₂ collisions for all transitions, the differences being of the order of roughly 25% – 35%. At the opposite, the collisional data for DNC-H₂ collisions can either overestimate or underestimate the ones for HNC-H₂ collisions depending on the transitions (see lower panels). It is also interesting to note that the HN¹³C-H₂ collisional data increase with increasing temperature while those for HNC-H₂ (DNC-H₂) collisions decrease. All this comparison clearly demonstrates the need of using isotopologues specific collisional data for modelling HCN/HNC isotopologues' observations.

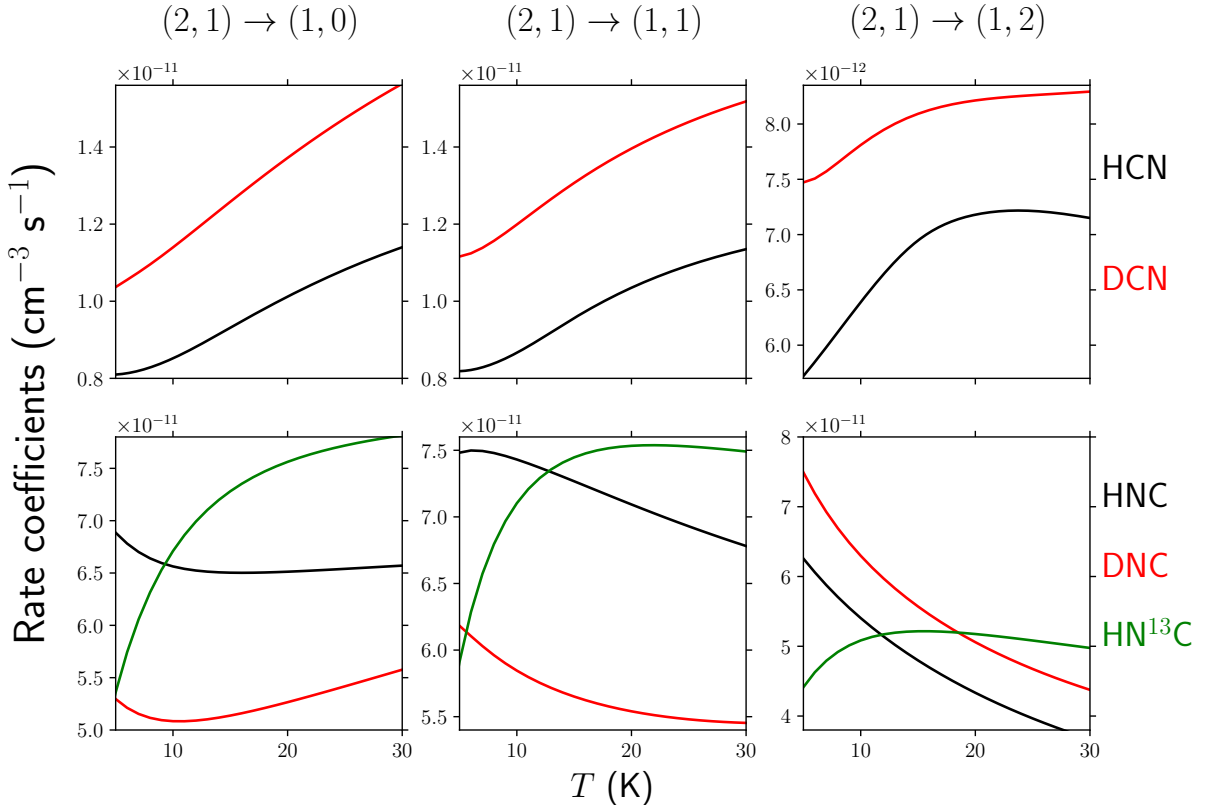


Figure A.3: Temperature dependence of hyperfine resolved rate coefficients of HCN and DCN (upper panels) and HNC, DNC and HN¹³C (lower panels) for the $(j, F) = (2, 1) \rightarrow (1, F')$ transitions. The data of HCN are from Magalhães et al. (2018).

A.4 Derivation of the physical structure of HH211

To derive the physical structure of HH211, we used the *Herschel* SPIRE dust continuum emission maps of the core from the *Herschel* Science Archive (HSA). Using the three SPIRE bands at 250, 350 and 500 μm , we derived the H₂ column density and the dust temperature map, following Harju et al. (2017). Therefore, we smoothed the 250 μm and 350 μm images to the resolution of the 500 μm image (35"), and then fitted a modified blackbody to each pixel. We applied a dust emissivity spectral index of $\beta = 2$, and a dust

emissivity coefficient per unit mass of gas of $\kappa_{250\mu\text{m}} = 0.1 \text{ cm}^2 \text{ g}^{-1}$ (Hildebrand 1983). The two resulting maps are shown in Fig. A.4, with the direction and size of the jet indicated by red and blue arrows.

From the H_2 column density map we derived a radial column density profile, shown in Fig. A.5. For this, we masked out the structures north and south-east of the core, and the filamentary structure trailing the core in the south-west (see Fig. A.4). To characterise the radial profile, we fitted a Plummer-like profile (Plummer 1911; Whitworth & Ward-Thompson 2001), that is modified by a constant term to account for the background column density (Launhardt et al. 2013):

$$N_{\text{H}_2}(r) = \frac{N_{\text{H}_2,0}}{[1 + (r/R_{\text{flat}})^2]^{(p-1)/2}} + N_{\text{out}}, \quad (\text{A.1})$$

where $N_{\text{H}_2,0}$ is the central column density, R_{flat} the characteristic radius of the flat inner portion of the density profile, and power-law index p . The three fit parameters were then used to derive the radial volume density profile (André et al. 2016):

$$n_{\text{outer}}(r) = \frac{n_{\text{H}_2,0}}{[1 + (r/R_{\text{flat}})^2]^{p/2}}. \quad (\text{A.2})$$

The central volume density $n_{\text{H}_2,0}$ is related to the central column density by $n_{\text{H}_2,0} = N_{\text{H}_2,0}/(A_P R_{\text{flat}})$, where A_P is a constant factor related to the inclination angle of the core to the plane of sky, $A_P = \frac{1}{\cos i} \times B\left(\frac{1}{2}, \frac{p-1}{2}\right)$, and B is the Euler Beta function. Assuming $i = 0^\circ$, we derived $N_0 = 9.7(6) \times 10^{22} \text{ cm}^{-2}$, $r_{\text{flat}} = 8(1) \times 10^{-2} \text{ pc}$, $p = 3.0(4)$, $N_{\text{out}} = 6(1) \times 10^{21} \text{ cm}^{-2}$, and $n_0 = 1.9(3) \times 10^5 \text{ cm}^{-3}$.

For the inner regions of the core, we adapted the frequently used physical model for IRAS 16293-2422 A/B (Crimier et al. 2010), where the radial density distribution is described by a power law, $n(\text{H}_2) = r^{-1.8}$. After normalising the profile with the central pixel of the *Herschel* column density map, we arrived at a volume density profile of

$$n_{\text{inner}}(r) = n_1 \cdot \left(\frac{r}{r_1}\right)^{-1.5}, \quad (\text{A.3})$$

where $n_1 = 1.89 \times 10^5 \text{ cm}^{-3}$ and $r_1 = 0.0128 \text{ pc}$ are the starting values of n_{outer} . This Shu-like density distribution is for example also used in Quénard et al. (2018).

To parametrise the radial temperature profile shown in Fig A.5, we applied an empirical fit following Arzoumanian et al. (2011):

$$T(r) = T_{\text{out}} - \frac{\Delta T}{\left(1 + \left(\frac{r}{r_{\text{flat}}}\right)^2\right)^{q/2}}. \quad (\text{A.4})$$

We fixed the values for T_{out} and ΔT to 21 K and 8 K, respectively, resulting in best-fit parameters of $r_{\text{flat}} = 0.17(1) \text{ pc}$ and $q = 0.38(3)$. The profile is decreasing towards the

centre, with a central temperature of around 13 K, and does not show an internal heating source. This is caused by the low spectral resolution of the *Herschel* temperature map, where one pixel covers around 6000 au. For example, Launhardt et al. (2013) show that embedded protostars only influence the temperature in the inner 5000 au of a globule.

To account for the central heating, we adopted the radial dust temperature profile for protostellar envelopes introduced by Motte & André (2001):

$$T_{\text{dust}}(r, L_*) \approx 38 \text{ K} \times \left(\frac{r}{100 \text{ AU}} \right)^{-0.4} \left(\frac{L_*}{1 L_\odot} \right)^{0.2}. \quad (\text{A.5})$$

Assuming that internal heating by the accreting protostar dominates the thermal balance of the envelope, and the envelope is optically thin to the bulk of the infrared radiation. To derive the internal luminosity of the protostar, we followed Dunham et al. (2008), using the 70 μm flux as an estimate:

$$L_{\text{int}} = 3.3 \times 10^8 F_{70}^{0.94} L_\odot. \quad (\text{A.6})$$

Using the 70 μm flux of HH211 published by Dunham et al. (2015), we derived a protostellar luminosity of $L_* = 1.2(2) L_\odot$. For the purpose of our radiative transfer simulations, we assumed the gas temperature in the core to be equal to the dust temperature. This is reasonable, as gas and dust are closely coupled at densities above 10^4 cm^{-3} (Goldsmith 2001).

For the infall velocity profile, we followed Crimier et al. (2010) and assumed a free fall velocity gradient,

$$v(r) = \sqrt{\frac{2gM}{r}}, \quad (\text{A.7})$$

where the mass M is given by the protostellar mass, $0.05 M_\odot$ (Lee et al. 2018). Pineda et al. (2019) show that HH211 is in fact not in free fall, but is rotating. However, as we are working with a 1D model of the source, the freefall velocity profile is somewhat mimicing the rotation by the steep increase of velocity towards the centre (see Pineda et al. in prep).

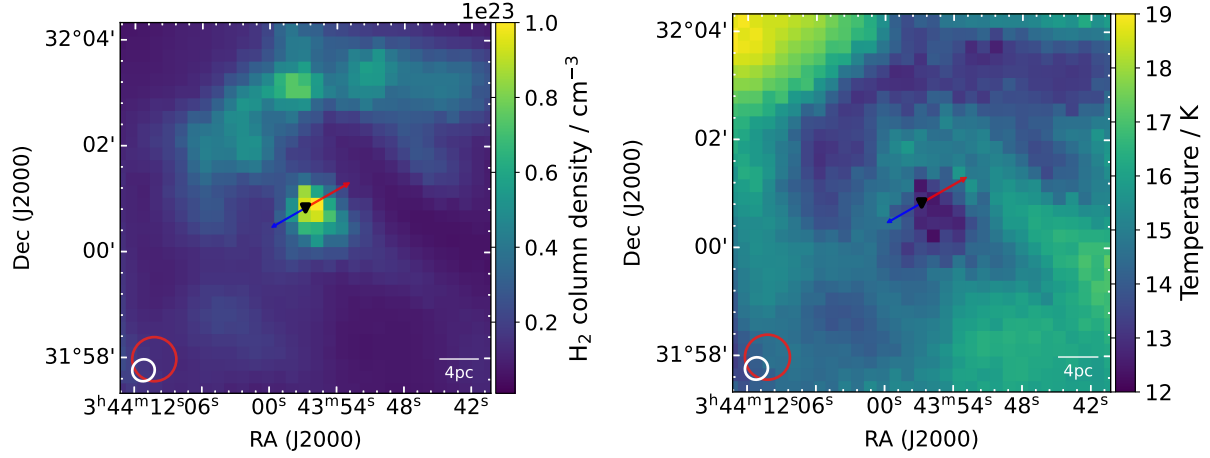


Figure A.4: Analysis of *Herschel* SPIRE data towards HH211. *Left*: H_2 column density map of HH211, derived using the *Herschel* SPIRE images at 250, 350 and 500 μm . The red and blue arrows indicate the direction and size of the molecular outflow/jet from the protostellar core (marked by a black triangle). The *Herschel* beam size is shown as a white circle, the average beam size of our observations is shown as red circle. *Right*: Dust temperature map of HH211, derived using the *Herschel* SPIRE images at 250, 350 and 500 μm .

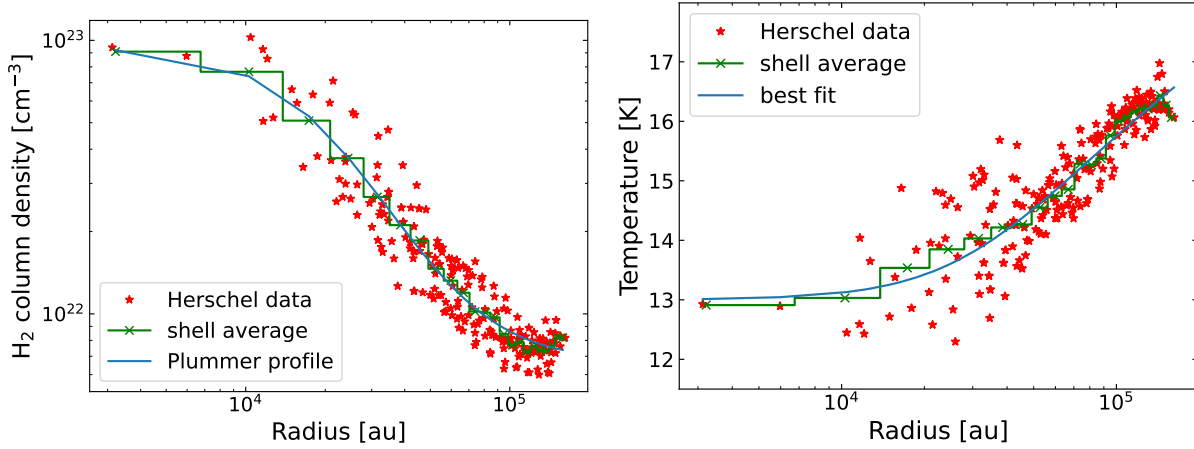


Figure A.5: Analysis of the H_2 column density map and the temperature map derived from *Herschel* data. *Left*: Radial H_2 column density profile (red), overlaid with the fitted Plummer-like profile (blue). Given in green are the data points averaged over shells of 11 arcsec, the size corresponding to 1/3 of the *Herschel* beam size. *Right*: Radial temperature profile (red), overlaid with the best fit (blue) and the averaged shells (green).

A.5 Molecular abundance profiles of the best-fit modelling results

Fig. A.6 and A.7 show the corresponding molecular abundance profiles of the best-fit modelling results derived with LOC and presented in this work, for L1544 and HH211, respectively.

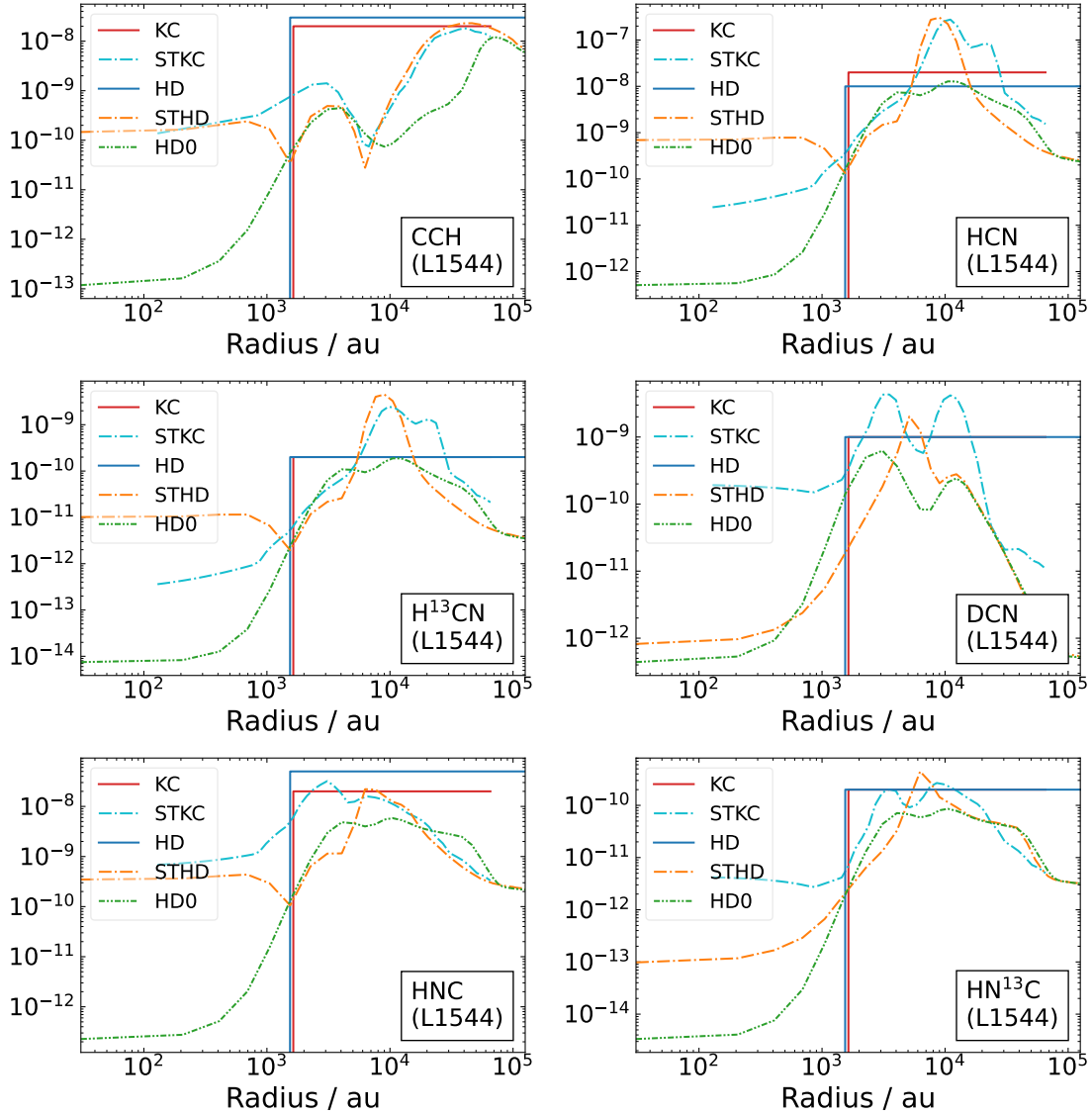


Figure A.6: Fractional molecular abundance (with respect to H_2) profiles of the best-fit results produced with LOC for the spectra observed towards L1544. In the case of the ^{13}C and ^{18}O isotopologues, the abundance profiles correspond to the profiles of the main species, scaled down by the isotopic ratio (68 and 557, respectively).

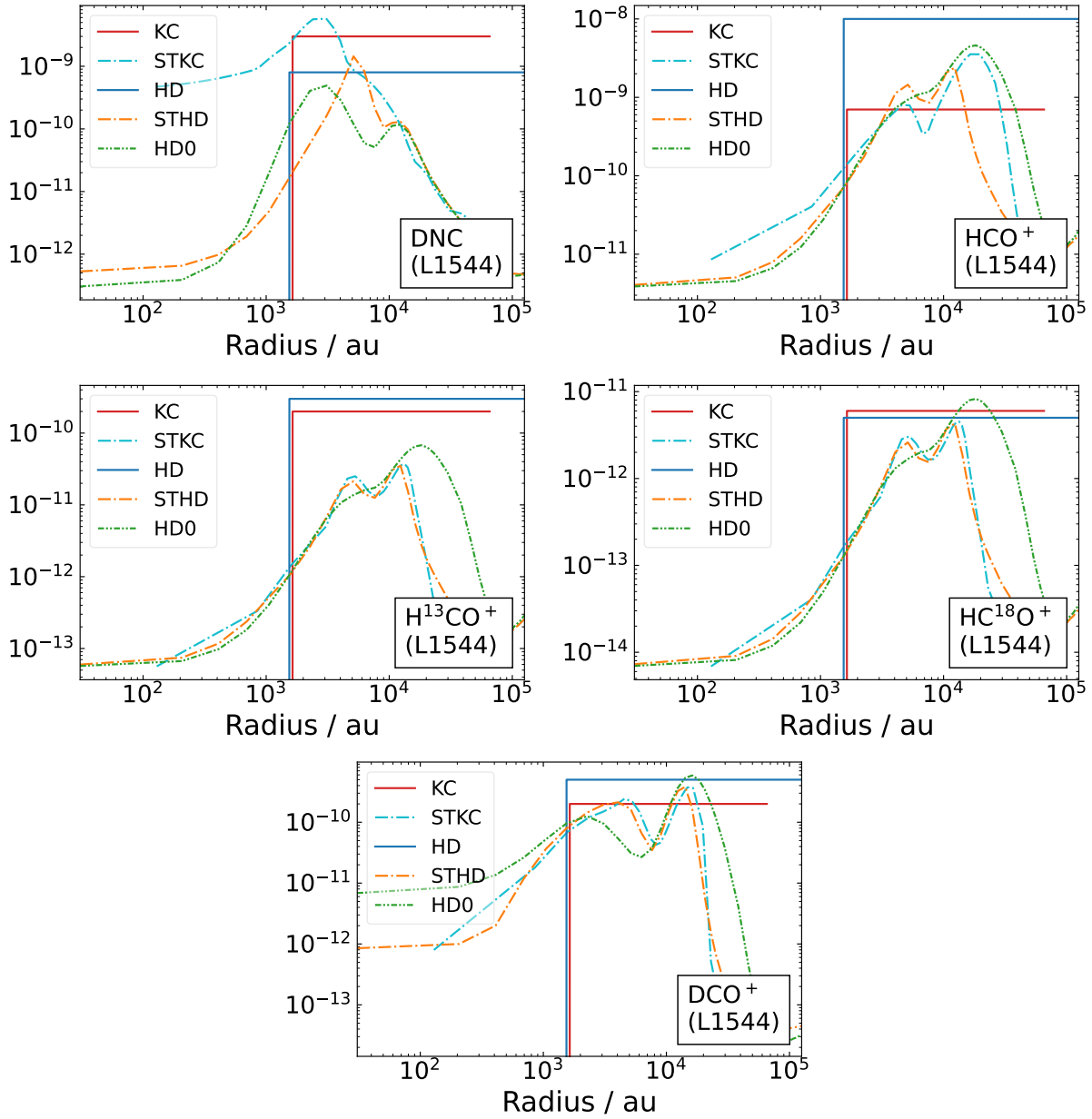


Figure A.6: continued.

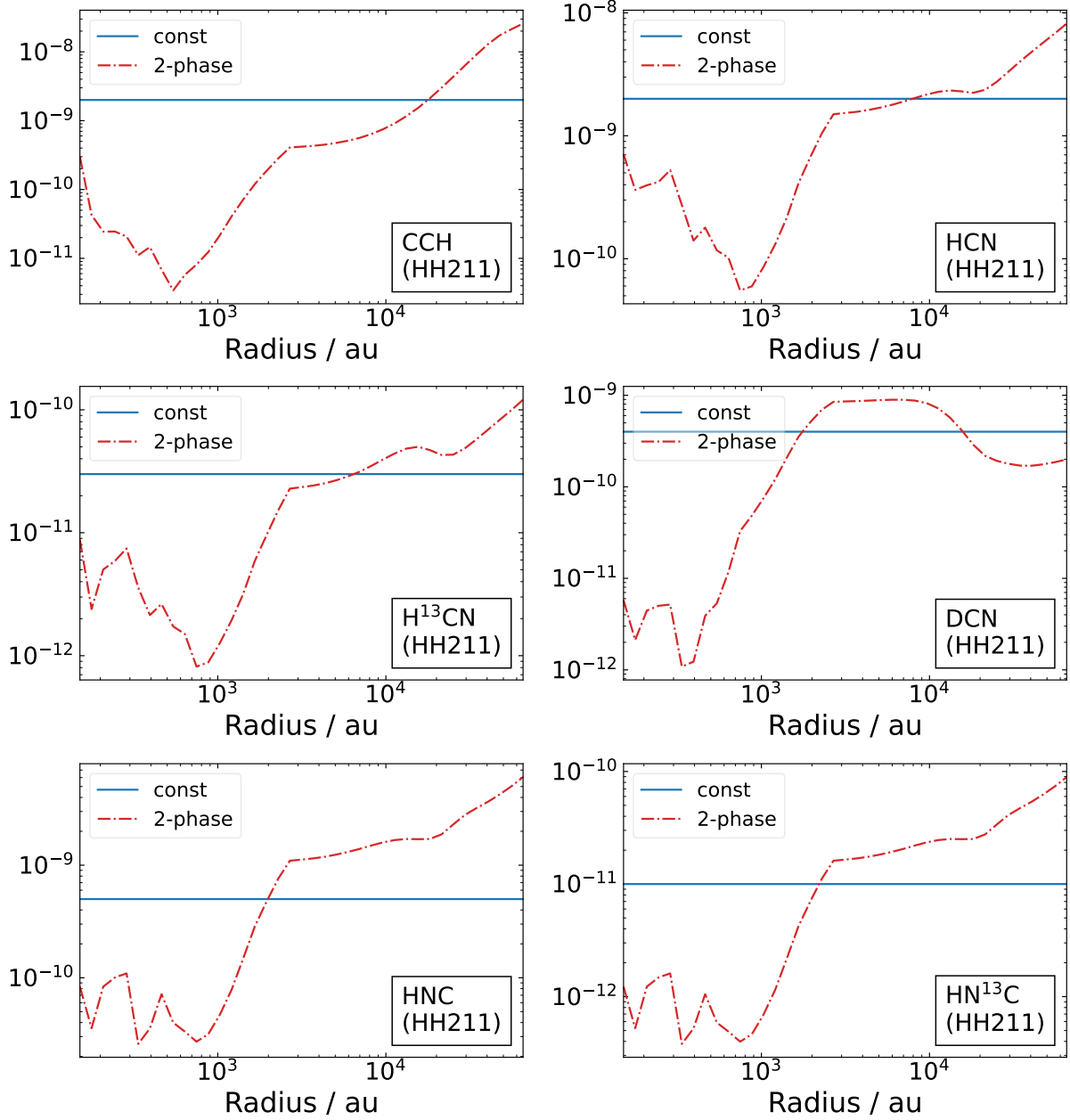


Figure A.7: Fractional molecular abundance (with respect to H_2) profiles of the best-fit results produced with LOC for the spectra observed towards HH211. In the case of the ^{13}C and ^{18}O isotopologues, the abundance profiles correspond to the profiles of the main species, scaled down by the isotopic ratio (68 and 557, respectively).

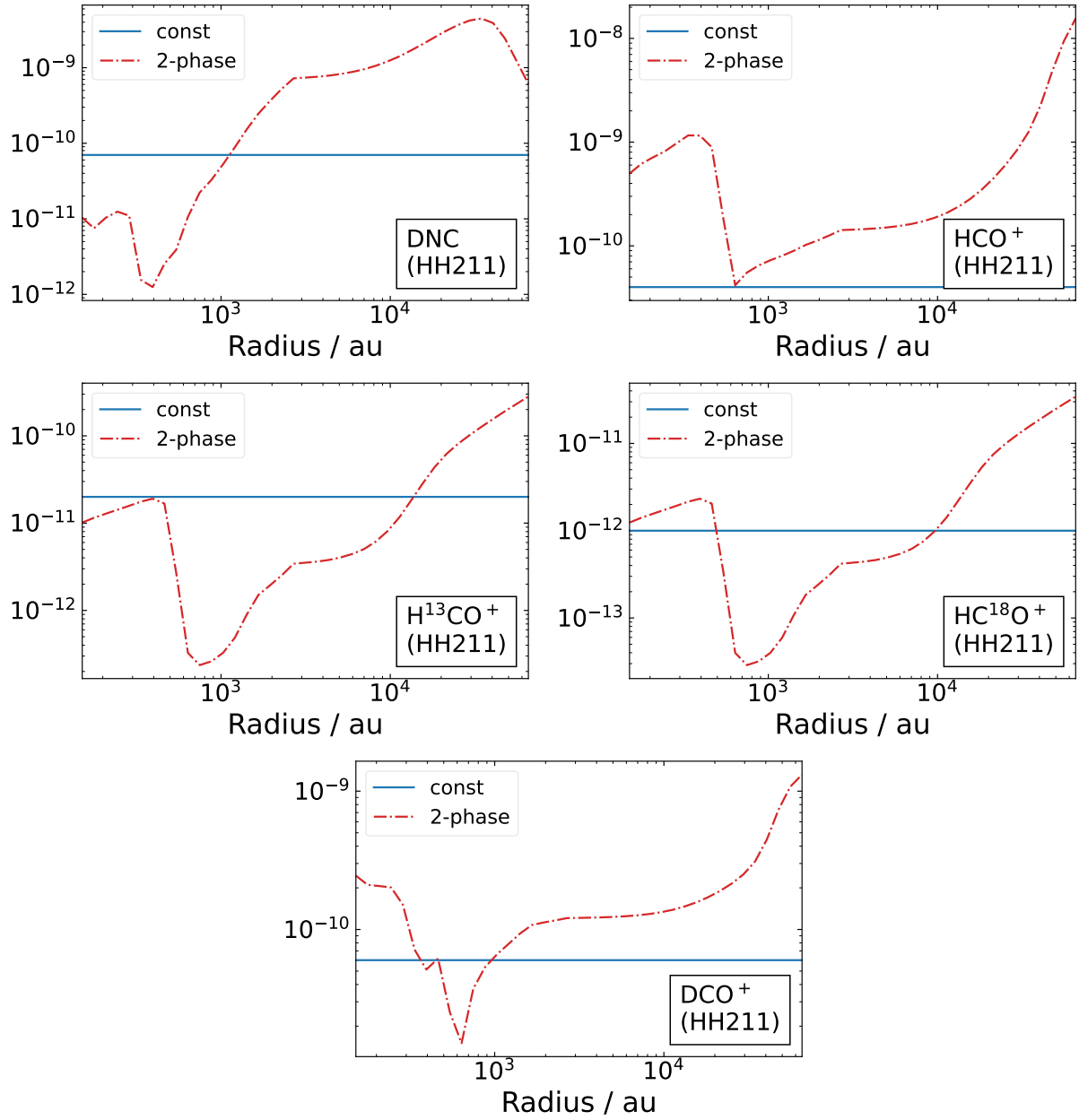


Figure A.7: continued.

Appendix B

Supplementary material for Chapter 3

B.1 Integrated intensity maps

The integrated intensity maps of $c\text{-C}_3\text{H}_2$, CH_3OH , and CH_3CCH , observed towards B68, L1521E, and L1544, are shown in Fig. B.1.

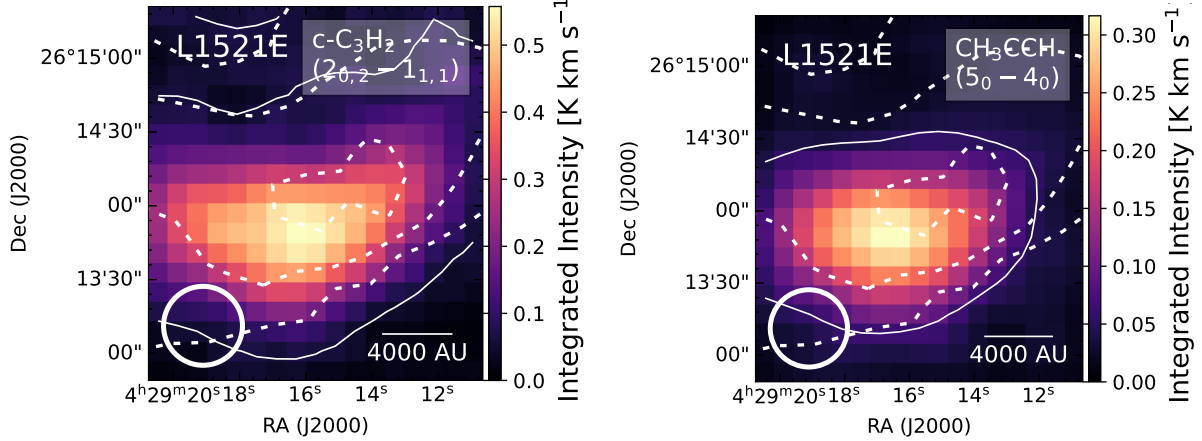


Figure B.1: Integrated intensity maps of CH_3OH , observed towards B68 (Spezzano et al. 2020), L1521E (Nagy et al. 2019), and L1544 (Spezzano et al. 2016). The solid line contours indicate the 3σ level of the integrated intensity, except for CH_3OH in L1544, where they indicate the 9σ level. The dashed line contours represent 90%, 50%, and 30% of the H_2 column density peak derived from *Herschel* maps (Spezzano et al. 2016, 2020). The white circle in the bottom left corner indicates the beam size of the IRAM 30 m telescope (32").

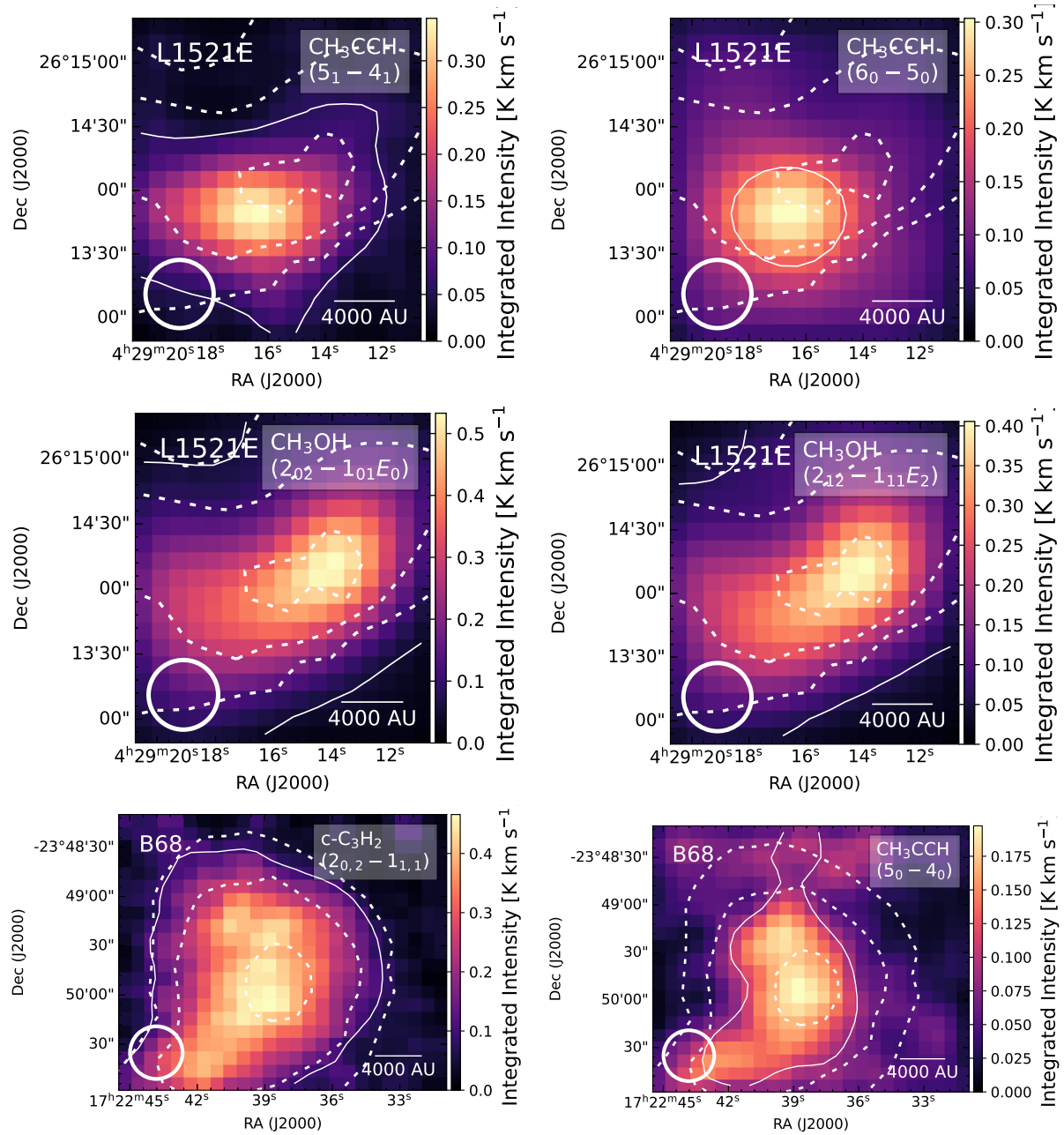


Figure B.1: continued.

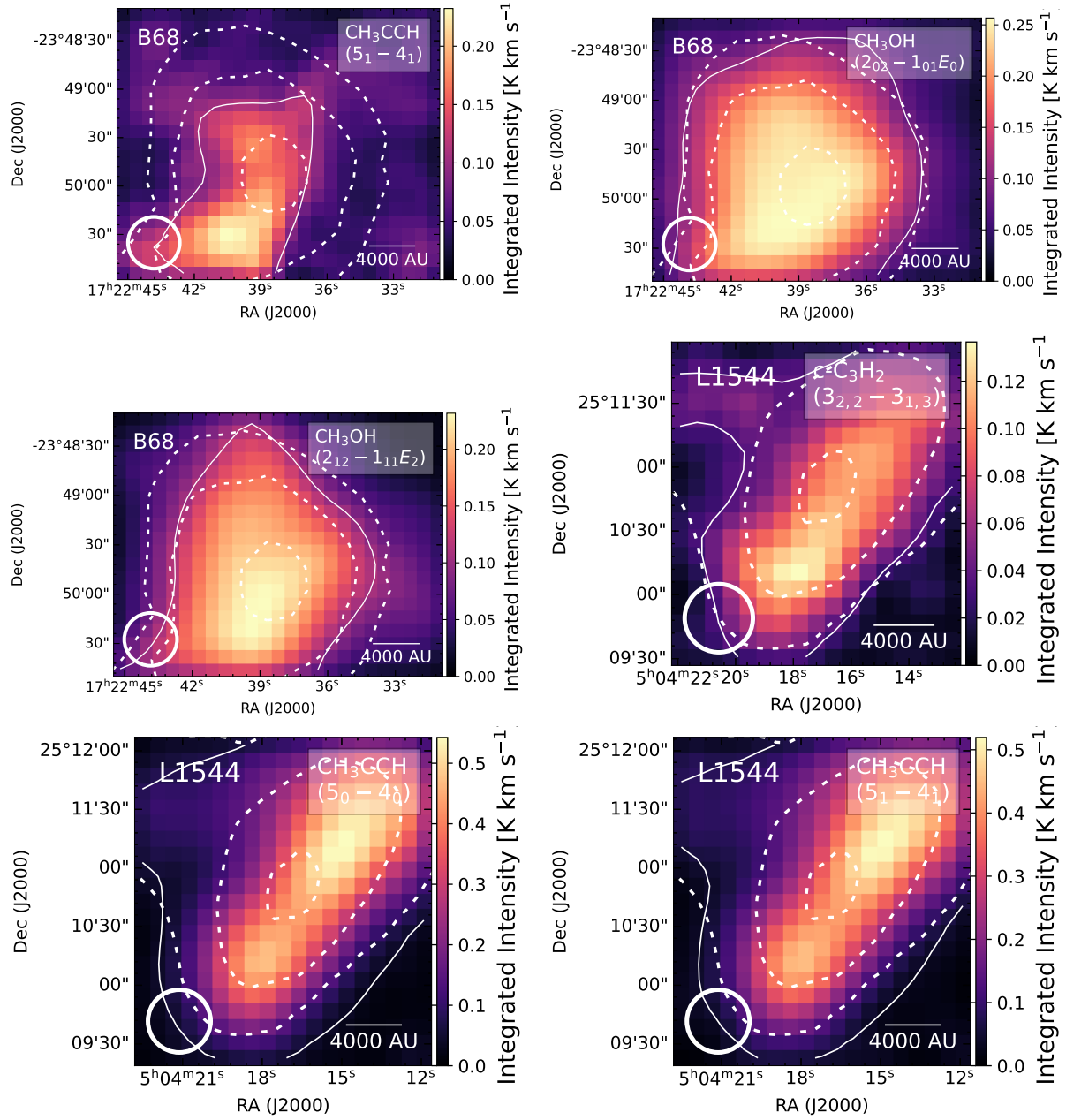


Figure B.1: continued.

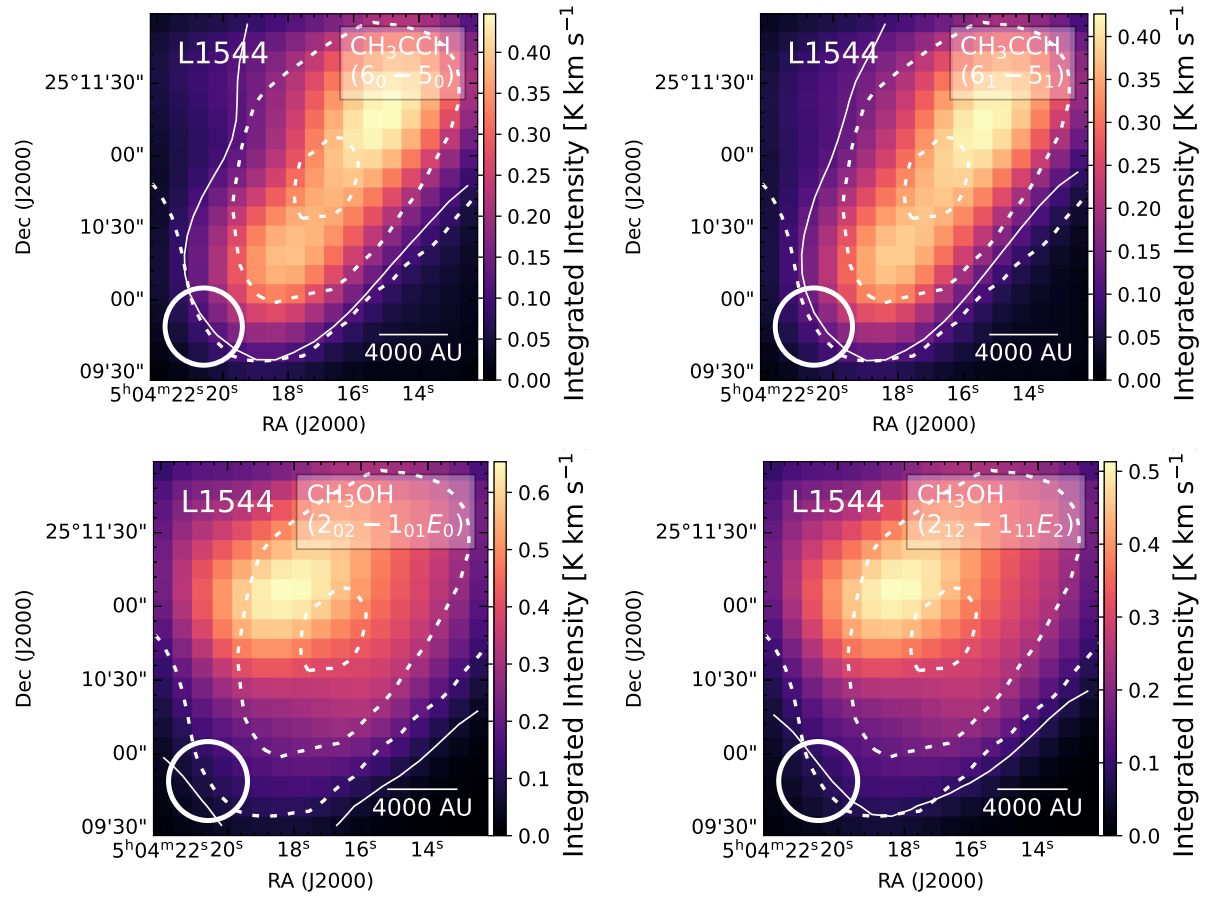


Figure B.1: continued.

B.2 H₂ column density gradient maps

Fig. B.2 shows the H₂ column density gradient maps (*right*) derived from *Herschel* SPIRE maps (see Spezzano et al. 2016, 2020) and the corresponding H₂ column density maps (*left*), for B68 (*top*), L1521E (*middle*), and L1544 (*bottom*), respectively.

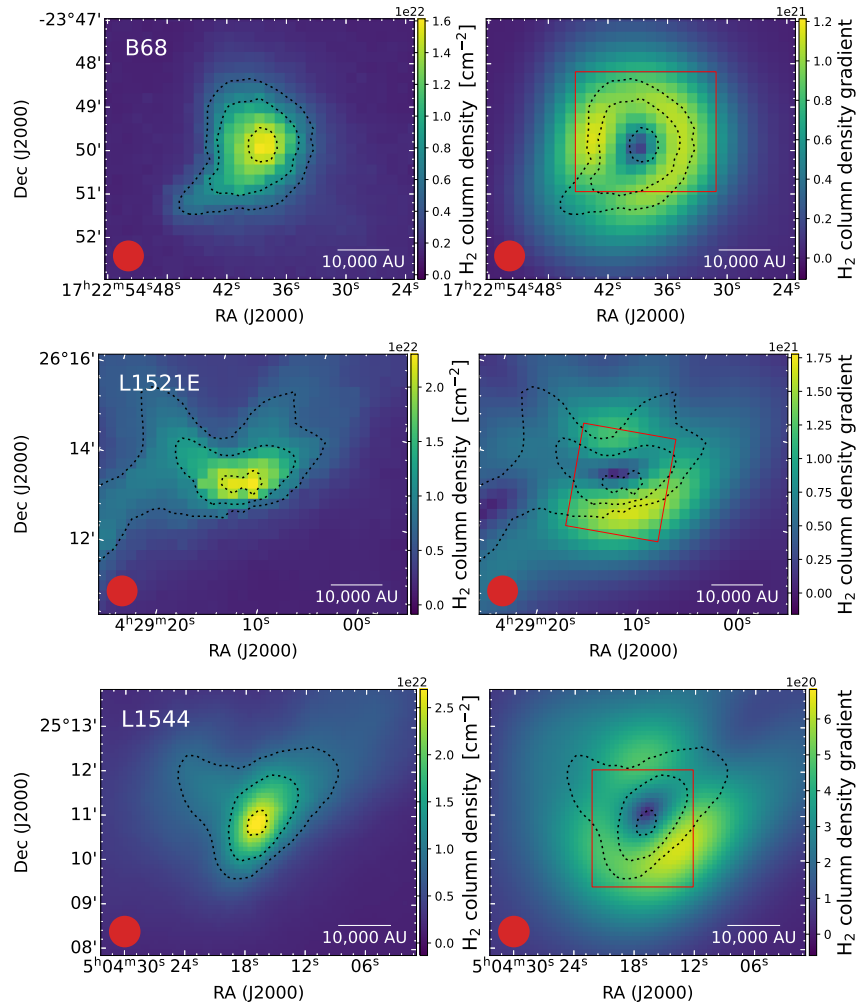


Figure B.2: H₂ column density maps (*left*) derived from *Herschel* SPIRE maps (Spezzano et al. 2016, 2020) and the corresponding H₂ column density gradient maps (*right*), for B68 (*top*), L1521E (*middle*), and L1544 (*bottom*). The dashed line contours represent 90, 50, and 30% of the H₂ column density peak. The red rectangle marks the location and size of the emission maps observed towards each core. The red circle in the bottom left corner indicates the *Herschel* beam size (40'').

B.3 Detailed clustering results

The clustering results of Case 1 and Case 2 are visualised in Fig. B.3 - Fig. B.6 for feature combinations 2, 3, 4, 9, and 10 in L1521E and L1544. The remaining results for Cases 1 and 2, together with the results for Cases 3 and 4 are published on Zenodo, in Figs. C.3 - C.14. The molecular ratio in a cluster and the number of data points assigned to it are given in Tables B.1, B.2, and B.3 for all combinations and Cases, for B68, L1521E, and L1544, respectively.

Table B.1: Cluster content for all Cases and feature combinations for the starless core B68. Given are number of points per cluster and the percentage of $\text{c-C}_3\text{H}_2$ over CH_3OH (Case 1), $\text{c-C}_3\text{H}_2$ over CH_3CCH (Case 2), CH_3OH over CH_3CCH (Case 3), and $\text{c-C}_3\text{H}_2$ over CH_3OH over CH_3CCH (Case 4). Given in brackets is the total number of data points per dataset. Ratios that differ from the initial input ratio by $\geq 10\%$ are given in boldface. The clusters are ordered by size and enumerated from 1 to 5. The corresponding colours in Figs. 3.2 - 3.3, and B.3 - B.6 are: 1=blue, 2=red, 3=cyan, 4=yellow, 5=purple.

	B68							
combi 1	Case 1 (502)		Case 2 (443)		Case 3 (439)		Case 4 (693)	
	#	%	#	%	#	%	#	%
cluster 1	250	44/56	212	55/45	156	74/26	304	32/27/42
cluster 2	52	54/46	26	85/15	135	45/55	39	38/41/21
cluster 3			7	57/43	20	75/25	22	45/ 9/45
cluster 4							11	27/ 0/73
cluster 5								
combi 2	Case 1		Case 2		Case 3		Case 4	
cluster 1	99	46/54	330	55/45	172	72/28	334	34/27/39
cluster 2	66	21/79	6	67/13	21	86/14	41	37/ 7/56
cluster 3	45	42/58			20	0/100		
cluster 4	32	81/19			16	0/100		
cluster 5	30	73/27						
combi 3	Case 1		Case 2		Case 3		Case 4	
cluster 1	114	43/57	264	53/47	264	58/42	469	34/28/38
cluster 2	107	48/52	33	67/33	21	100/0	47	49/ 6/45
cluster 3	50	42/58			11	82/18		
cluster 4	50	24/76			10	90/10		
cluster 5								
combi 4	Case 1		Case 2		Case 3		Case 4	
cluster 1	362	48/52	391	60/40	96	83/17	473	40/19/41
cluster 2	32	41/59	20	50/50	59	54/46	57	32/16/53
cluster 3	20	50/50	10	0/100	37	30/70	40	25/50/25
cluster 4	18	56/44	8	50/50	21	100/0	26	46/23/31
cluster 5	17	29/71			19	53/47	20	20/80/ 0
combi 5	Case 1		Case 2		Case 3		Case 4	
cluster 1	470	49/51	182	56/44	216	85/15	638	36/26/38
cluster 2	24	83/17	38	50/50	13	0/100	11	82/ 0/18
cluster 3					11	0/100	7	100/ 0/ 0
cluster 4					10	0/100		

Table B.1: continued.

	B68							
cluster 5	10 90/10							
combi 6	Case 1		Case 2		Case 3		Case 4	
cluster 1	204 38/62	174 61/39	106 73/27	333 34/21/45				
cluster 2	84 48/52	35 66/34	52 71/29	23 43/ 9/48				
cluster 3		18 33/67	41 78/22					
cluster 4			32 22/78					
cluster 5								
combi 7	Case 1		Case 2		Case 3		Case 4	
cluster 1	266 44/56	224 68/32	240 72/28	415 37/19/43				
cluster 2	29 0/100	70 51/49	41 78/22	40 0/20/80				
cluster 3	16 100/0	25 52/48	18 56/44	35 74/26/ 0				
cluster 4		13 8/92	17 0/100	28 36/29/36				
cluster 5			10 0/100					
combi 8	Case 1		Case 2		Case 3		Case 4	
cluster 1	435 49/51	398 60/40	273 69/31	149 32/17/50				
cluster 2	19 47/53	22 45/55	40 90/10	134 49/ 6/45				
cluster 3		8 50/50	38 39/61	63 35/14/51				
cluster 4		8 0/100	17 0/100					
cluster 5								
combi 9	Case 1		Case 2		Case 3		Case 4	
cluster 1	135 39/61	224 54/46	153 73/27	168 33/17/51				
cluster 2	122 37/63	30 73/27	68 75/25	101 30/19/51				
cluster 3	26 35/65			50 30/26/44				
cluster 4				20 35/ 5/60				
cluster 5				14 57/29/14				
combi 10	Case 1		Case 2		Case 3		Case 4	
cluster 1	146 51/49	248 65/35	287 68/32	142 45/12/43				
cluster 2	89 48/52	33 58/42	60 40/60	123 38/24/38				
cluster 3	44 45/55	29 34/66	30 33/67	78 31/38/31				
cluster 4	22 45/55	15 13/87	21 0/100	33 33/30/36				
cluster 5			18 83/17					

Table B.2: Same as Table B.1, but for the starless core L1521E.

	L1521E							
combi 1	Case 1 (436)		Case 2 (437)		Case 3 (481)		Case 4 (677)	
	# %	# %	# %	# %	# %	# %	# %	# %
cluster 1	97 52/48	109 33/67	104 38/62	191 35/30/35				
cluster 2	71 58/42	49 55/45	96 29/71	158 34/32/34				
cluster 3	27 63/37	48 58/42	16 69/31	34 62/29/ 9				
cluster 4	21 90/10	48 83/17	15 47/53	20 0/100/ 0				
cluster 5	20 0/100		10 90/10	17 0/94/ 6				
combi 2	Case 1		Case 2		Case 3		Case 4	
cluster 1	120 57/43	141 35/65	359 54/46	183 39/27/33				

Table B.2: continued.

	L1521E					
cluster 2	60 23/77	101 64/36	19 0/100	164 24/39/37		
cluster 3	42 50/50		17 0/100			
cluster 4						
cluster 5						
combi 3	Case 1	Case 2	Case 3	Case 4		
cluster 1	83 43/57	154 38/62	173 48/52	260 28/39/33		
cluster 2	51 65/35	106 54/46	34 71/29	86 41/20/40		
cluster 3	50 40/60	24 29/71	31 6/94	30 37/57/ 7		
cluster 4	34 56/44		25 40/60			
cluster 5	20 30/70		23 39/61			
combi 4	Case 1	Case 2	Case 3	Case 4		
cluster 1	96 54/46	120 45/55	206 55/45	169 30/46/24		
cluster 2	75 37/63	61 46/54	102 36/64	85 34/36/29		
cluster 3	45 40/60	33 36/64	47 32/68	69 28/48/25		
cluster 4		29 38/62		54 26/44/30		
cluster 5				43 28/16/56		
combi 5	Case 1	Case 2	Case 3	Case 4		
cluster 1	112 54/46	109 66/34	141 48/52	180 41/23/36		
cluster 2	51 27/73	89 31/69	56 80/20	106 22/41/38		
cluster 3	42 38/62	18 0/100	47 40/60	39 31/ 8/62		
cluster 4	13 100/0	15 47/53		22 27/36/36		
cluster 5	11 100/0	14 64/36				
combi 6	Case 1	Case 2	Case 3	Case 4		
cluster 1	142 49/51	137 47/53	159 47/53	239 31/34/35		
cluster 2	59 44/56	79 56/44	32 81/19	57 51/23/26		
cluster 3	41 71/29	26 27/73	32 62/38	31 23/52/26		
cluster 4			31 42/58	20 50/ 5/45		
cluster 5			27 0/100	16 25/ 6/69		
combi 7	Case 1	Case 2	Case 3	Case 4		
cluster 1	84 48/52	101 48/52	88 42/58	124 34/35/31		
cluster 2	74 62/38	79 62/38	82 46/54	87 39/13/48		
cluster 3	60 50/50	40 17/83	40 78/22	61 34/28/38		
cluster 4			35 26/74	37 35/19/46		
cluster 5				32 22/38/41		
combi 8	Case 1	Case 2	Case 3	Case 4		
cluster 1	94 46/54	368 45/55	184 51/49	470 28/33/39		
cluster 2	48 48/52	25 32/68	37 97/ 3			
cluster 3	41 15/85		26 23/77			
cluster 4	38 34/66		26 38/62			

Table B.2: continued.

	L1521E							
cluster 5	18 50/50							
combi 9	Case 1		Case 2		Case 3		Case 4	
cluster 1	131 39/61	98 55/45	116 72/28	186 33/32/35				
cluster 2	60 33/67	65 35/65	75 43/57	116 26/40/34				
cluster 3	22 45/55	28 54/46	24 54/46	44 41/27/32				
cluster 4	18 17/83	25 32/68	23 0/100	27 15/26/59				
cluster 5		23 30/70	10 30/70					
combi 10	Case 1		Case 2		Case 3		Case 4	
cluster 1	273 42/58	76 47/53	106 72/28	344 28/28/44				
cluster 2	45 42/58	47 57/43	68 49/51	56 21/54/25				
cluster 3	31 48/52	39 36/64	35 34/66	50 24/52/24				
cluster 4	24 50/50	31 55/45	15 40/60	40 28/57/15				
cluster 5	20 60/40	25 32/68	15 40/60	23 22/48/30				

Table B.3: Same as Table B.1, but for the prestellar core L1544.

	L1544							
combi 1	Case 1 (445)		Case 2 (404)		Case 3 (477)		Case 4 (663)	
	# %	# %	# %	# %	# %	# %	# %	# %
cluster 1	178 39/61	184 38/62	168 67/33	298 26/31/43				
cluster 2	49 0/100	59 53/47	49 96/ 4	47 0/ 4/96				
cluster 3		28 82/18	18 39/61	30 37/43/20				
cluster 4		17 0/100	17 0/100	17 0/100/ 0				
cluster 5		13 62/38						
combi 2	Case 1		Case 2		Case 3		Case 4	
cluster 1	218 46/64	261 41/59	132 66/34	202 35/17/48				
cluster 2	36 8/92	80 59/41	45 73/27	78 23/72/ 5				
cluster 3	32 41/59		41 88/12	67 27/24/49				
cluster 4	19 16/84		38 11/89	63 11/ 8/81				
cluster 5	18 17/83		24 0/100	23 0/100/ 0				
combi 3	Case 1		Case 2		Case 3		Case 4	
cluster 1	159 42/58	224 47/53	111 68/32	182 24/35/41				
cluster 2	151 28/72	50 22/78	110 65/35	63 38/52/10				
cluster 3	29 69/31	19 95/5	52 4/96	61 25/31/44				
cluster 4	23 91/9		36 0/100	24 8/17/75				
cluster 5			22 100/0	12 0/100/ 0				
combi 4	Case 1		Case 2		Case 3		Case 4	
cluster 1	146 36/64	160 37/63	104 83/17	630 28/31/41				
cluster 2	55 65/35	59 56/44	52 79/21	14 0/100/ 0				

Table B.3: continued.

	L1544							
cluster 3	27 0/100	9 0/100	43 77/23	11 27/73/ 0				
cluster 4	20 50/50	9 78/22	29 0/100					
cluster 5		6 67/33	19 0/100					
combi 5	Case 1	Case 2	Case 3	Case 4				
cluster 1	178 33/67	339 43/57	415 61/39	367 19/32/48				
cluster 2	80 19/81	9 0/100	24 0/100	16 44/56/ 0				
cluster 3	10 30/70	8 75/25		14 14/21/64				
cluster 4	7 14/86	7 100/0						
cluster 5		6 100/0						
combi 6	Case 1	Case 2	Case 3	Case 4				
cluster 1	211 27/73	179 35/65	318 57/43	377 19/36/46				
cluster 2	15 7/93	24 38/62	13 0/100	24 33/67/ 0				
cluster 3	11 9/91		8 0/100	16 31/19/50				
cluster 4	9 33/67							
cluster 5	6 50/50							
combi 7	Case 1	Case 2	Case 3	Case 4				
cluster 1	362 32/68	185 38/62	250 63/37	477 21/32/46				
cluster 2	11 100/0	30 37/63	24 42/58	15 33/67/ 0				
cluster 3	6 0/100		18 0/100					
cluster 4			12 0/100					
cluster 5			11 64/36					
combi 8	Case 1	Case 2	Case 3	Case 4				
cluster 1	351 36/64	136 40/60	390 59/41	372 20/32/48				
cluster 2	18 50/50	64 50/50	17 59/41	35 14/37/49				
cluster 3		14 21/79		35 37/46/17				
cluster 4				13 15/23/62				
cluster 5								
combi 9	Case 1	Case 2	Case 3	Case 4				
cluster 1	183 34/66	168 38/62	261 52/48	382 26/34/41				
cluster 2	38 55/45	57 44/56	42 36/64	54 37/35/28				
cluster 3	23 39/61							
cluster 4								
cluster 5								
combi 10	Case 1	Case 2	Case 3	Case 4				
cluster 1	319 34/66	138 40/60	391 56/44	486 25/33/42				
cluster 2	20 50/50	31 42/58	15 87/13	25 36/52/12				
cluster 3	13 100/0	30 50/50	10 60/40					
cluster 4	11 0/100	15 53/47	8 0/100					
cluster 5	7 100/0	13 15/85	6 50/50					

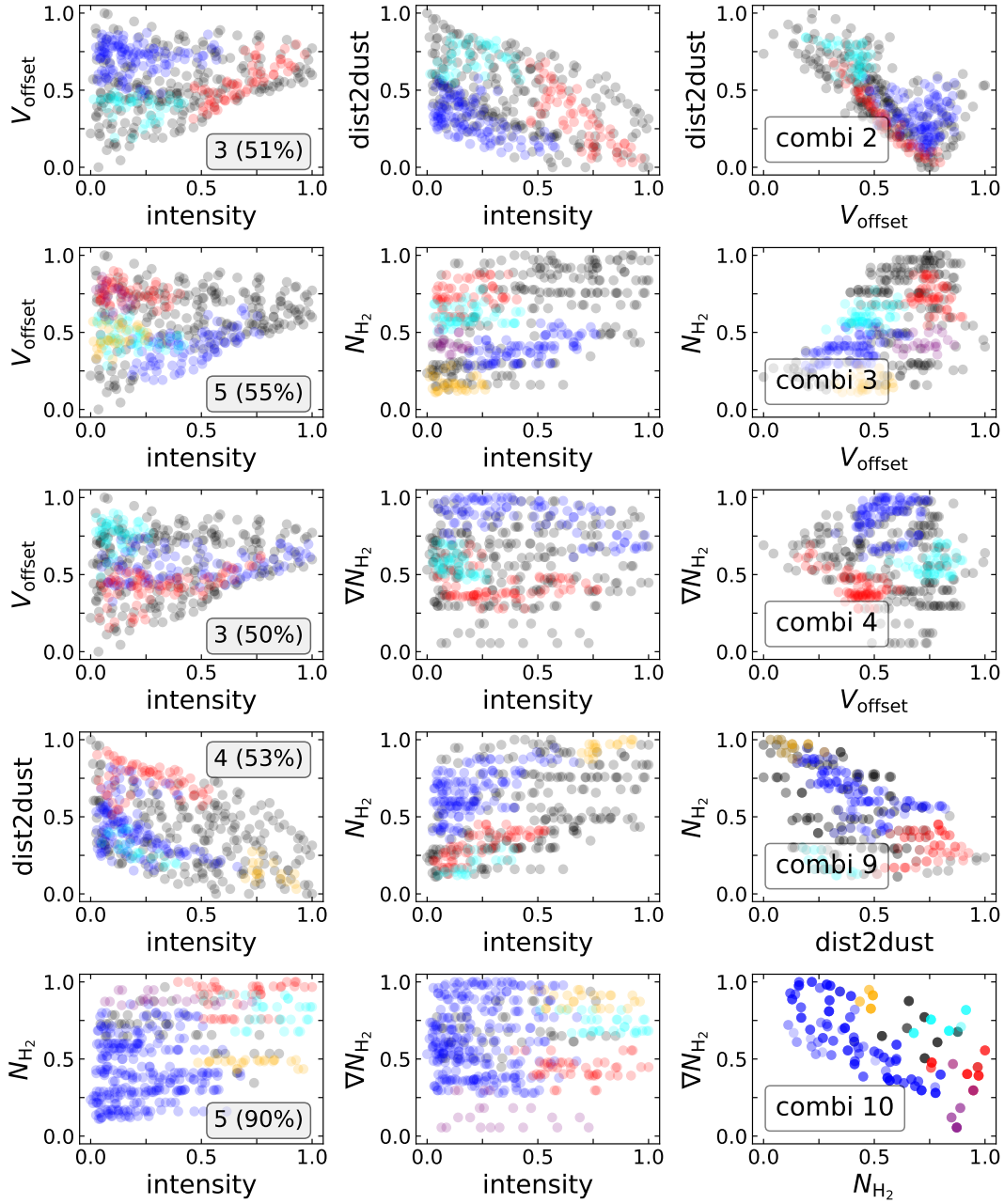
L1521E: $c\text{-C}_3\text{H}_2$ vs CH_3OH (Case 1)

Figure B.3: Clustering results for the starless core L1521E for the dataset of Case 1, for feature combinations 2, 3, 4, 9, and 10. Each row represents a different combination of features (see Table. 3.3). The annotations provide information on how many clusters were found by the algorithm (2-5) and what percentage of data points are assigned to the clusters. Noise points (=points not assigned to any cluster) are plotted in black. The colours of the clusters are ordered by cluster size: the biggest cluster is given in blue, followed by red, cyan, yellow, and purple.

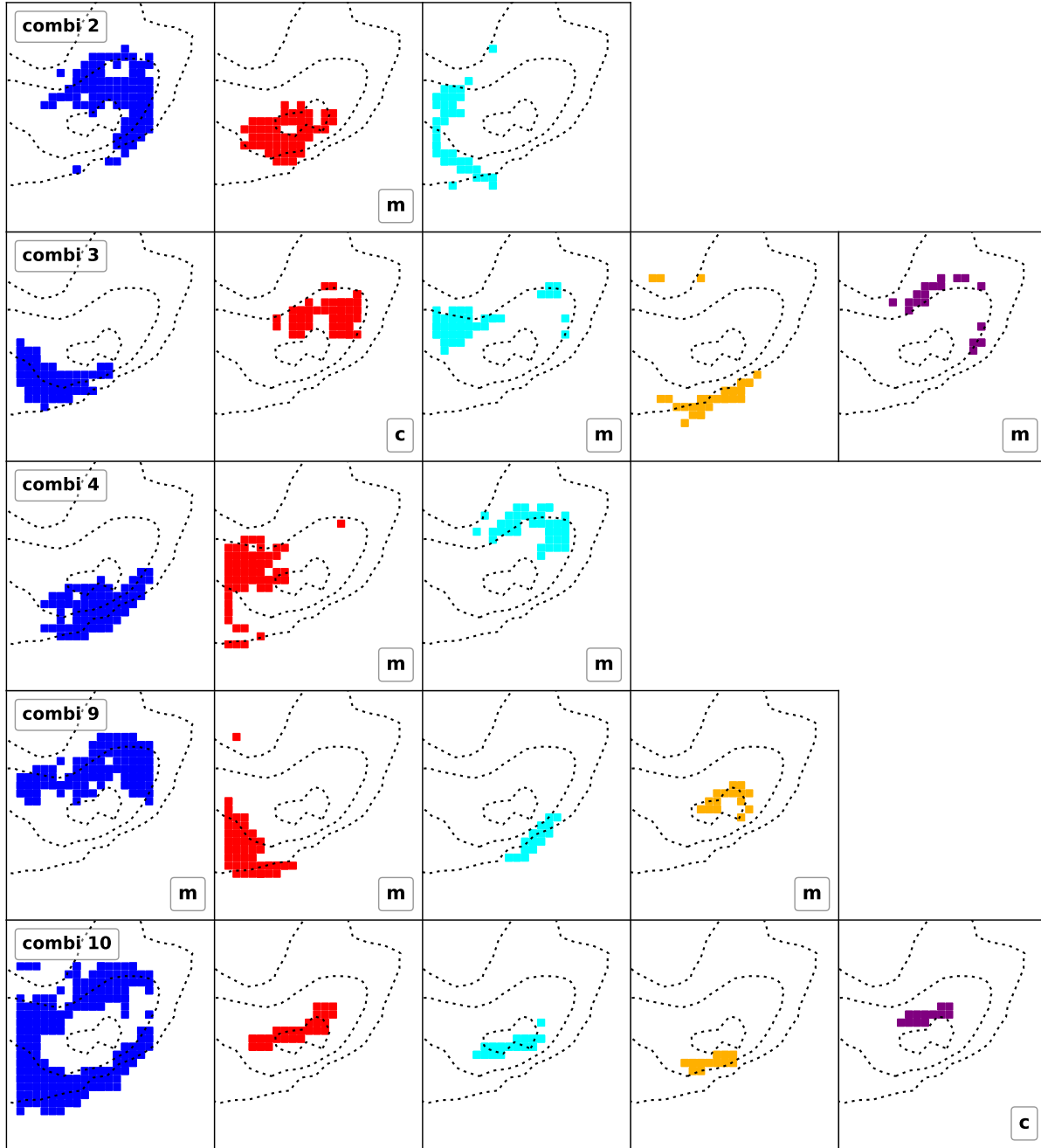


Figure B.3: continued. Corresponding spatial distribution of each cluster across the starless core L1521E. The annotations indicate if a cluster contains > 60% of one molecule (c: $c\text{-C}_3\text{H}_2$; m: CH_3OH ; p: CH_3CCH). The dashed line contours represent 30, 50, 90% of the H_2 column density peak derived from *Herschel* maps Spezzano et al. (2020).

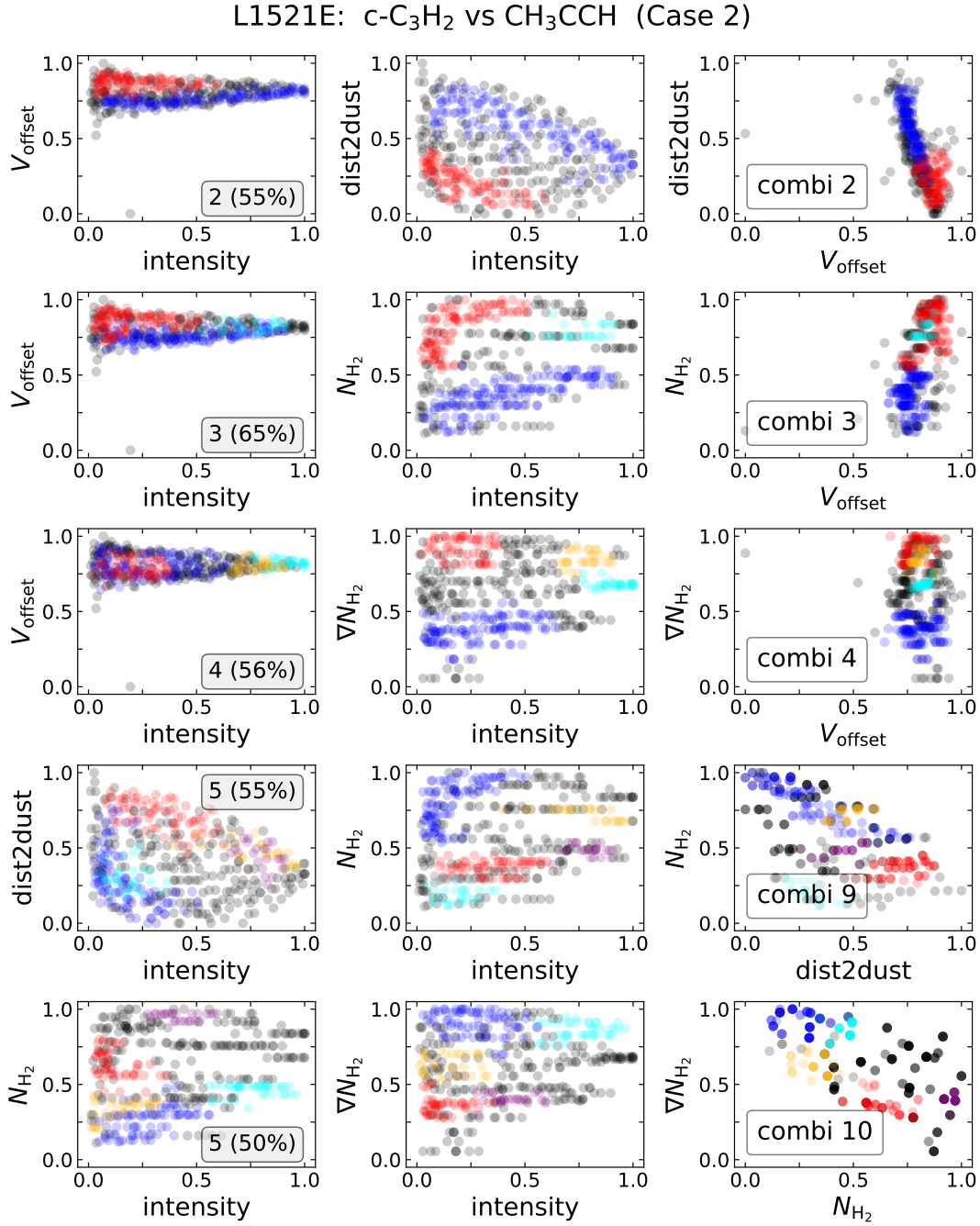


Figure B.4: Clustering results for the starless core L1521E for the dataset of Case 1 (*left*) and the dataset of Case 2 (*right*), for feature combinations 2, 3, 4, 9, and 10. Each row represents a different combination of features (see Table. 3.3). The annotations provide information on how many clusters were found by the algorithm (2-5) and what percentage of data points are assigned to the clusters. Noise points (=points not assigned to any cluster) are plotted in black. The colours of the clusters are ordered by cluster size: the biggest cluster is given in blue, followed by red, cyan, yellow, and purple.

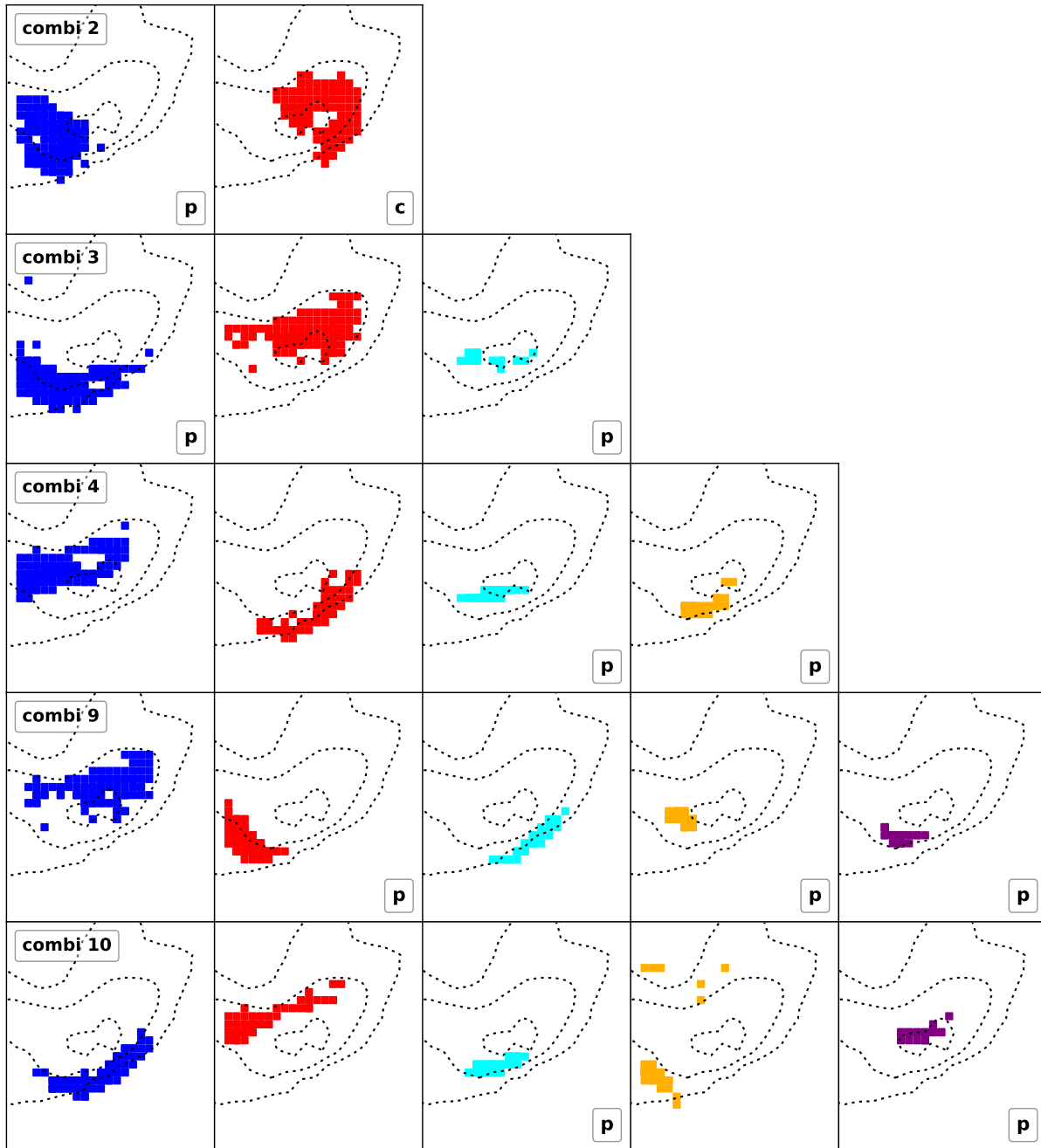


Figure B.4: continued. Corresponding spatial distribution of each cluster across the starless core L1521E. Corresponding spatial distribution of each cluster across the core. The annotations indicate if a cluster contains $> 60\%$ of one molecule (c: c-C₃H₂; m: CH₃OH; p: CH₃CCH). The dashed line contours represent 30, 50, 90% of the H₂ column density peak derived from *Herschel* maps Spezzano et al. (2020).

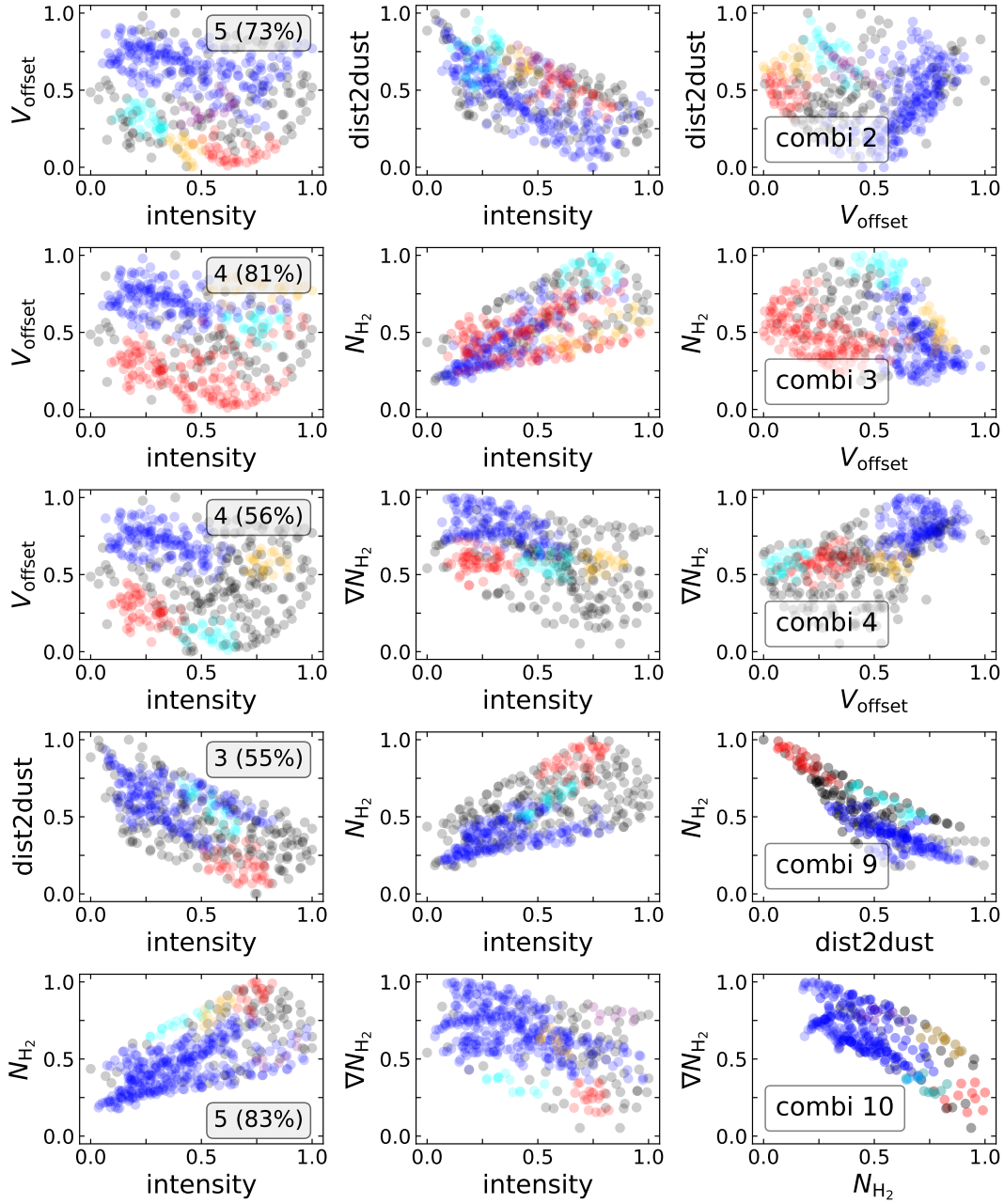
L1544: $\text{c-C}_3\text{H}_2$ vs CH_3OH (Case 1)

Figure B.5: Clustering results for the prestellar core L1544 for the dataset of Case 1, for feature combinations 2, 3, 4, 9, and 10. Each row represents a different combination of features (see Table. 3.3). The annotations provide information on how many clusters were found by the algorithm (2-5) and what percentage of data points are assigned to the clusters. Noise points (=points not assigned to any cluster) are plotted in black. The colours of the clusters are ordered by cluster size: the biggest cluster is given in blue, followed by red, cyan, yellow, and purple.

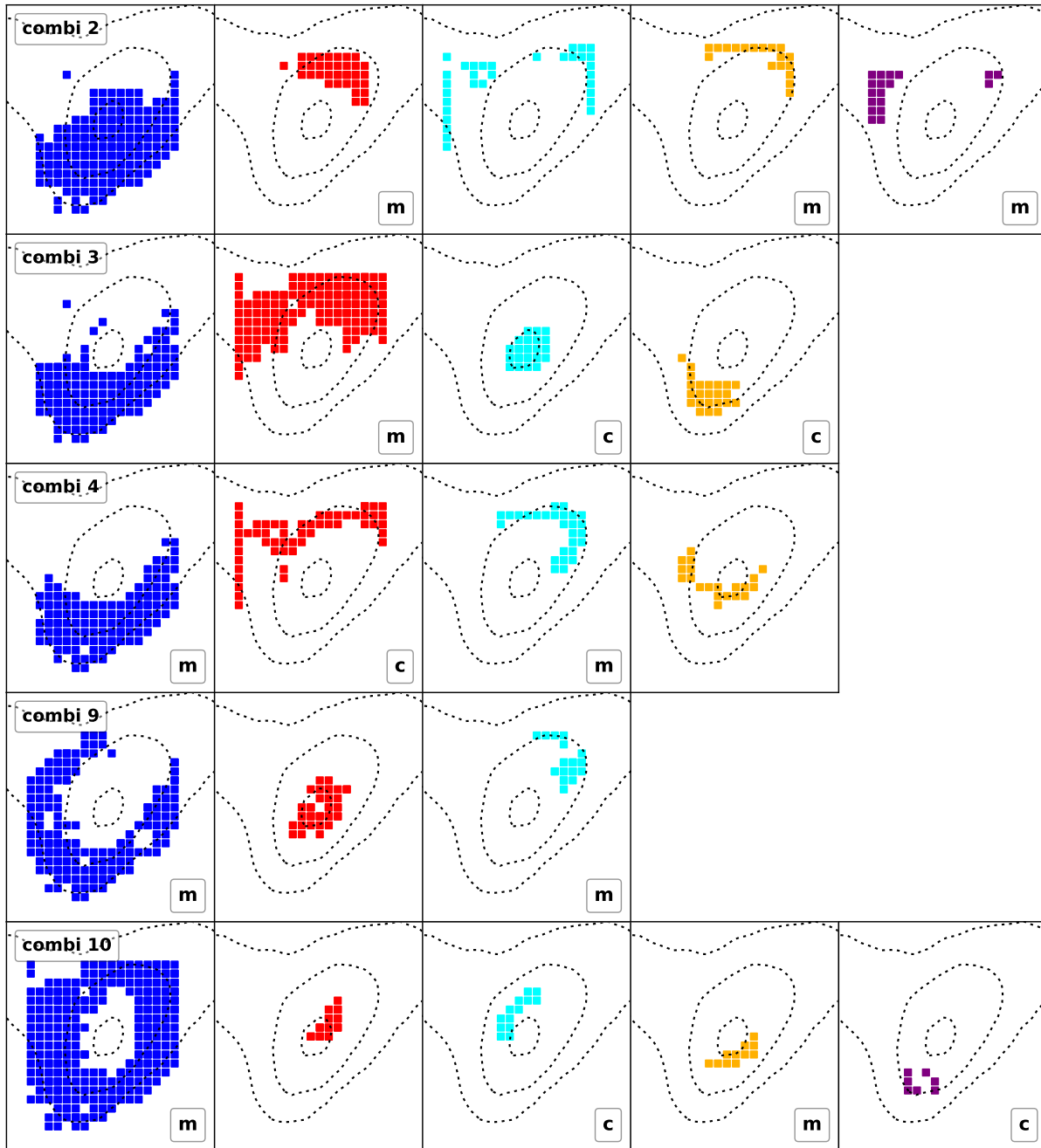


Figure B.5: continued. Corresponding spatial distribution of each cluster across the prestellar core L1544. Corresponding spatial distribution of each cluster across the core. The annotations indicate if a cluster contains $>60\%$ of one molecule (c: $c\text{-C}_3\text{H}_2$; m: CH_3OH ; p: CH_3CCH). The dashed line contours represent 30, 50, 90% of the H_2 column density derived from *Herschel* maps Spezzano et al. (2020).

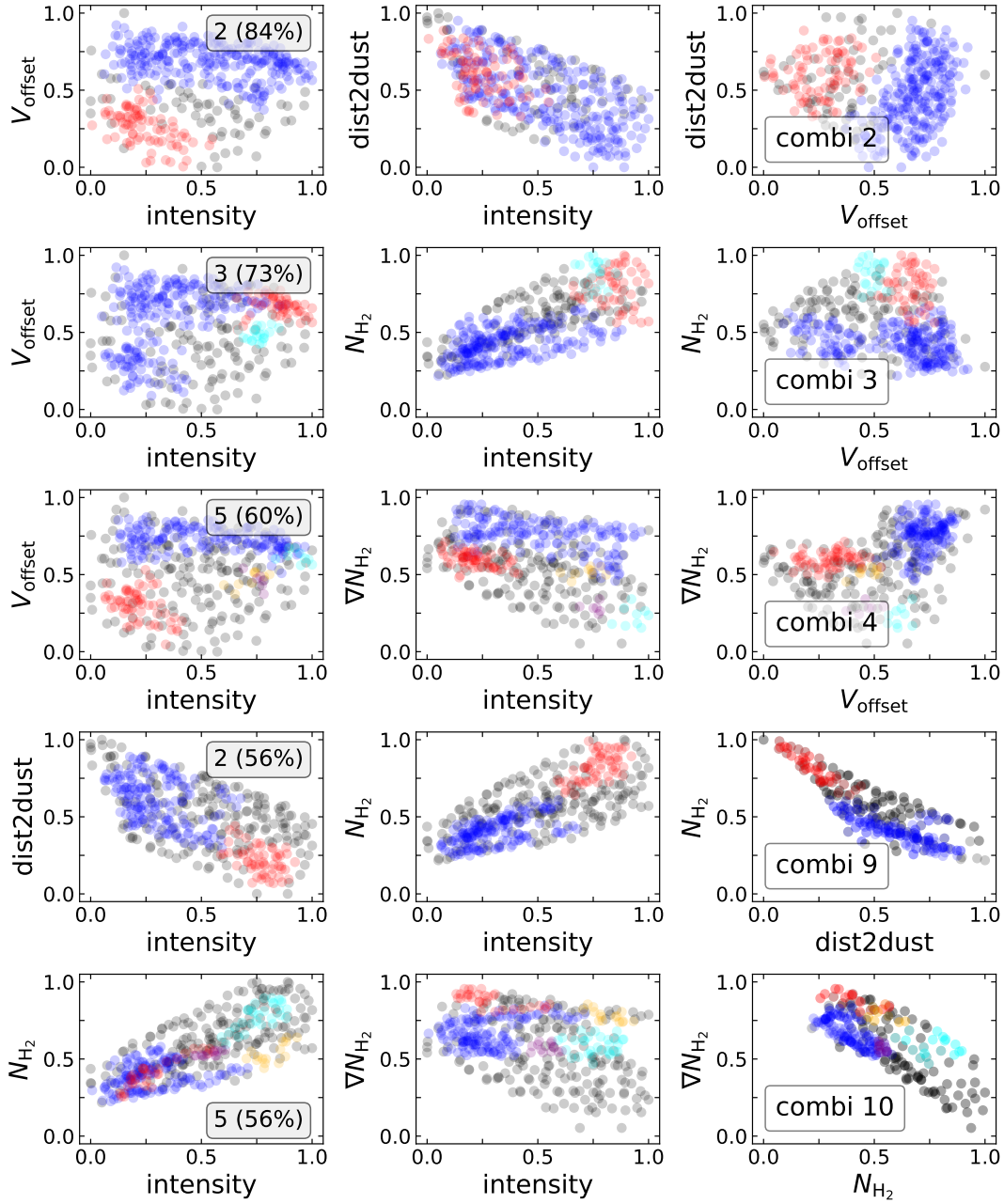
L1544: $c\text{-C}_3\text{H}_2$ vs CH_3CCH (Case 2)

Figure B.6: Clustering results for the prestellar core L1544 for the dataset of Case 2, for feature combinations 2, 3, 4, 9, and 10. Each row represents a different combination of features (see Table 3.3). The annotations provide information on how many clusters were found by the algorithm (2-5) and what percentage of data points are assigned to the clusters. Noise points (=points not assigned to any cluster) are plotted in black. The colours of the clusters are ordered by cluster size: the biggest cluster is given in blue, followed by red, cyan, yellow, and purple.

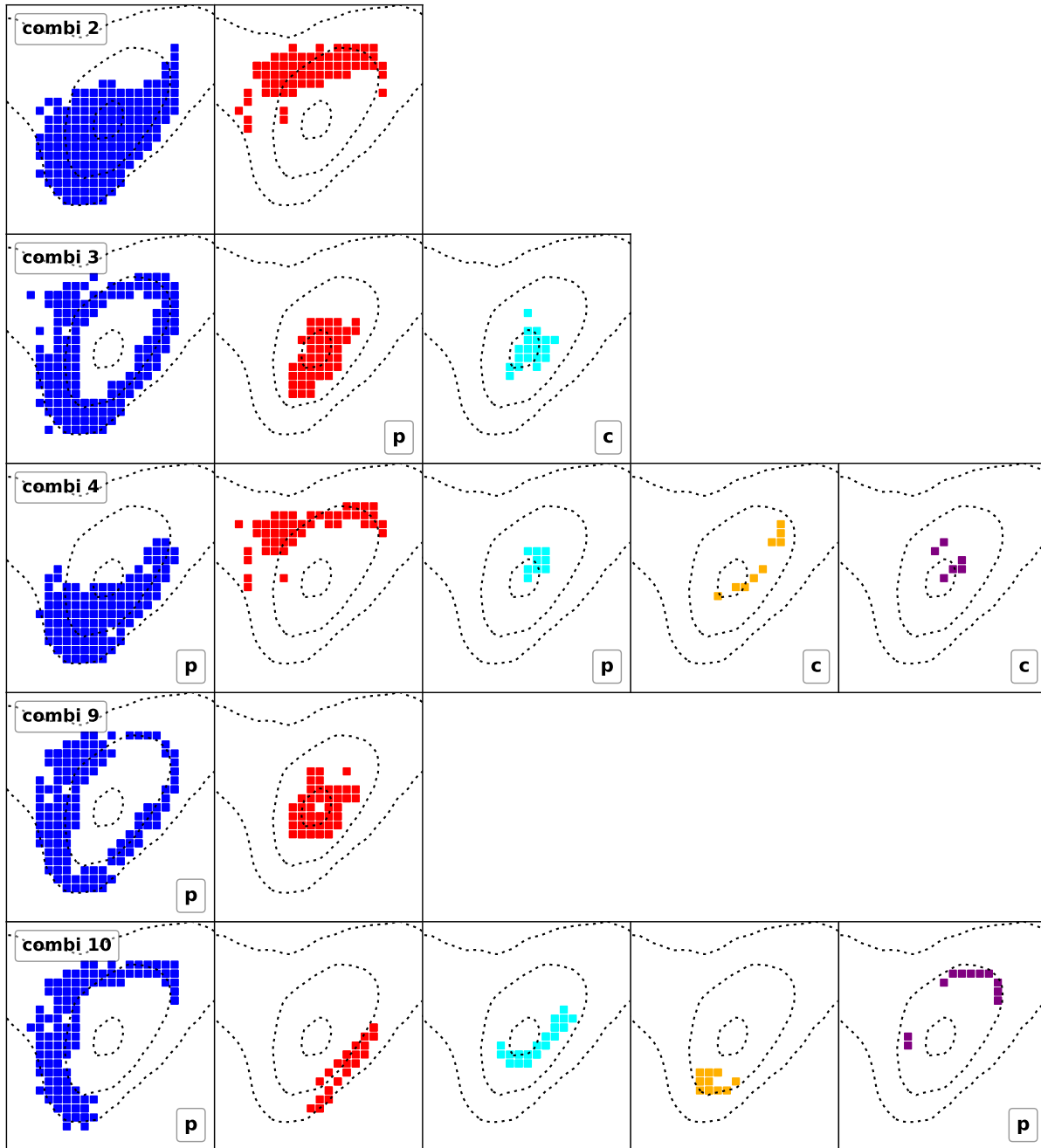


Figure B.6: continued. Corresponding spatial distribution of each cluster across the prestellar core L1544. Corresponding spatial distribution of each cluster across the core. The annotations indicate if a cluster contains $> 60\%$ of one molecule (c: c-C₃H₂; m: CH₃OH; p: CH₃CCH). The dashed line contours represent 30, 50, 90% of the H₂ column density peak derived from *Herschel* maps Spezzano et al. (2020).

B.4 Spectra at dust peak

The spectra of $c\text{-C}_3\text{H}_2$ and CH_3CCH observed towards the dust peaks of B68, L1521E, L1544, OphD, HMM1, L694-2, and L429 are shown in Fig. B.7. The observed datacubes are convolved with the *Herschel* beam size ($40''$), then a circular aperture with radius $8''$ is used to extract the spectra.

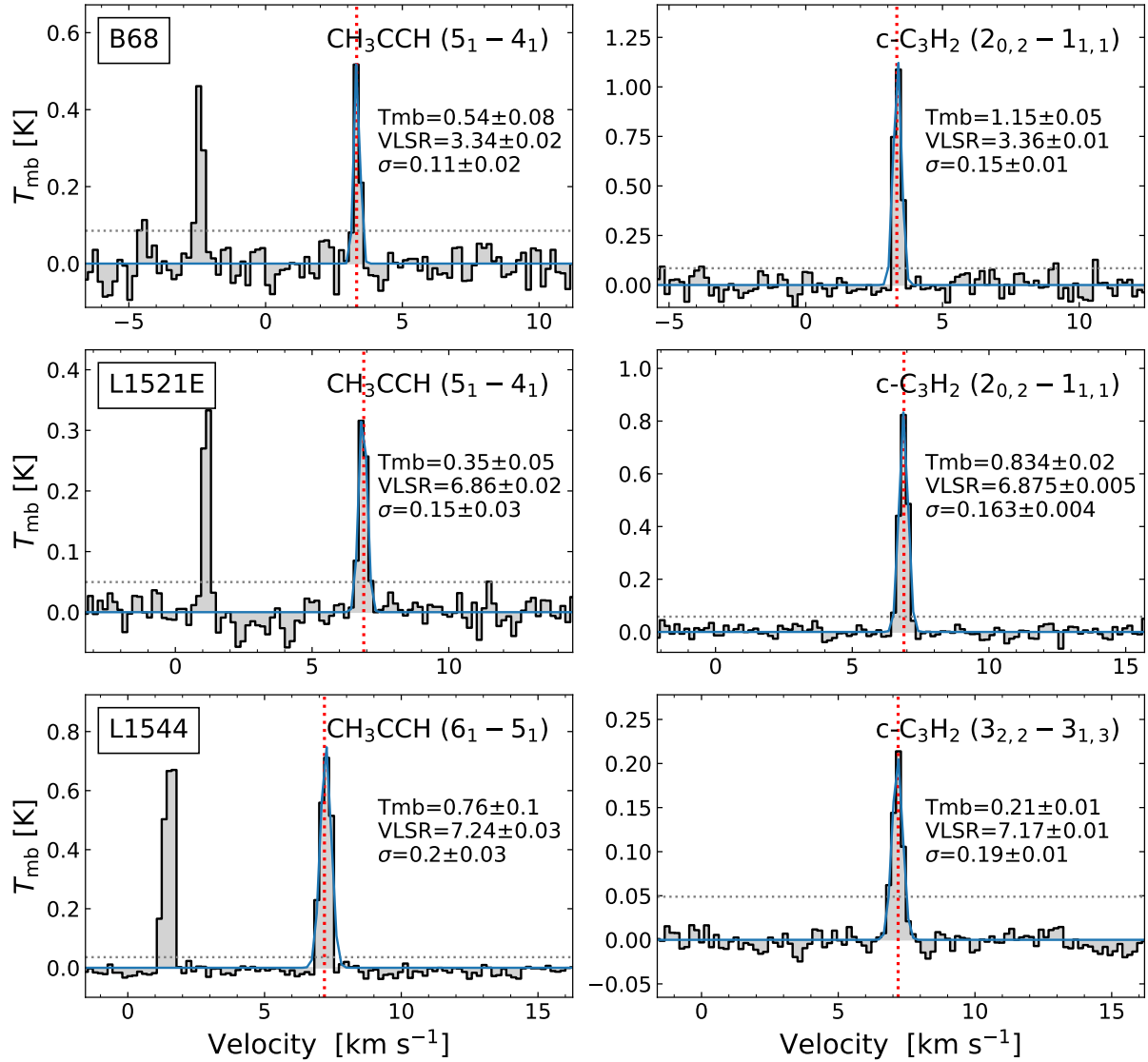


Figure B.7: Spectra of CH_3CCH (*left*) and $c\text{-C}_3\text{H}_2$ (*right*) at the dust peak of each core (black), extracted within a circular aperture of radius $8''$, and the corresponding Gaussian fit (cyan). The 3σ level is indicated by the grey dotted line. The systemic velocity with respect to the line chosen for analysis is shown by the red dotted line. The Gaussian fit parameters are annotated for each line.

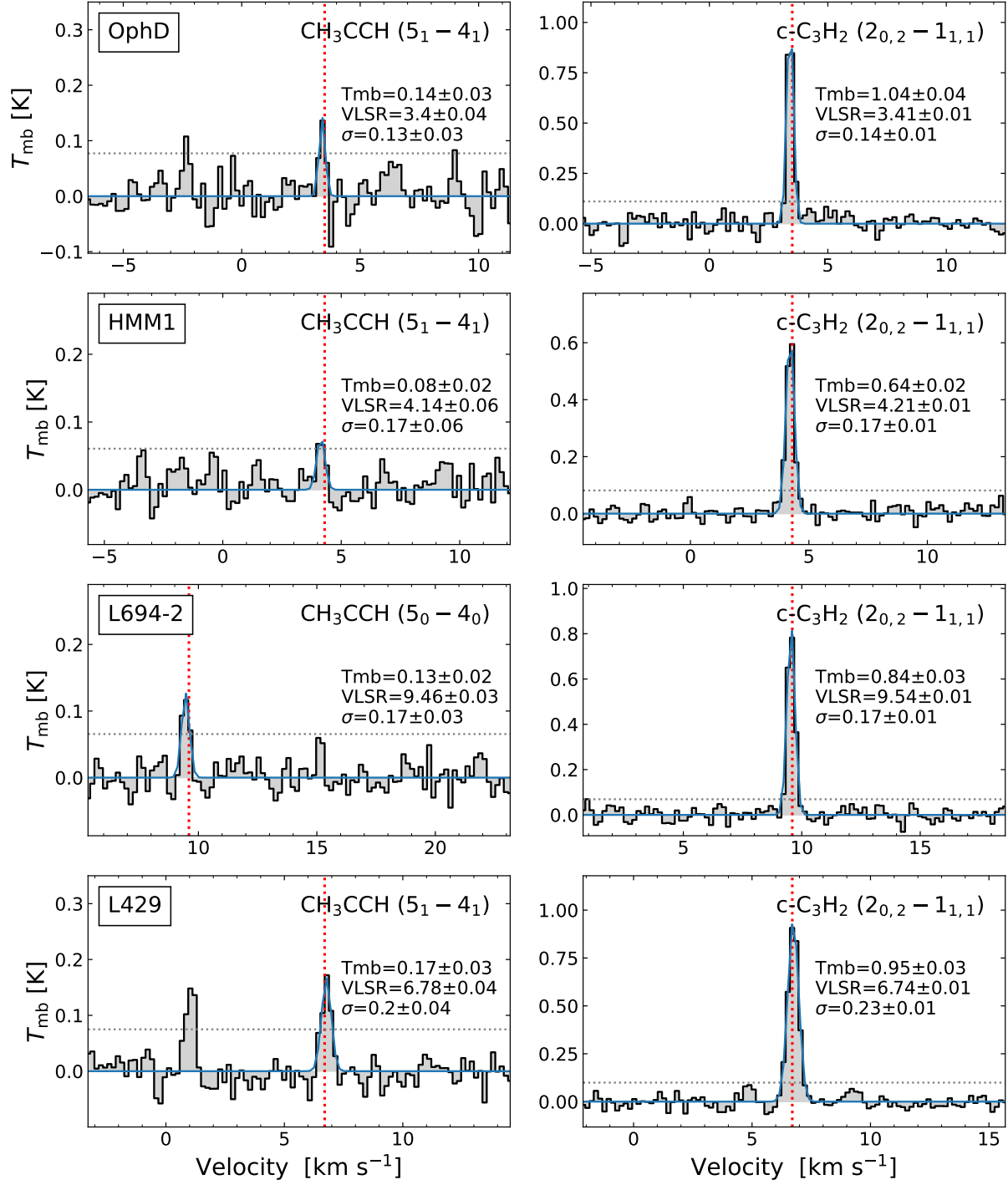


Figure B.7: continued.

Appendix C

Supplementary material for Chapter 4

C.1 Deuteration of CH₃CCH

Figure C.1 shows the column density map of CH₃CCH derived from the (5₁ – 4₁) transition and the resulting deuteration maps. The distribution of the emission of the K=1 transition is identical to the K=0 transition, resulting in an identical morphology the deuterium fraction. However, the peak column density derived from K=1 with an excitation temperature of 10 K is a factor of two higher compared to the values derived from K=0. This results in substantially lower deuteration levels, with peak values of 0.045(5) and 0.07(1) for N(CH₃CCD)/N(CH₃CCH) and N(CH₂DCCH)/N(CH₃CCH), respectively. However, the relation between the two deuterated isotopologues does not change, and therefore, the conclusions drawn from them are not affected.

C.2 Spectra at molecular peaks

Figure C.2 presents spectra of the observed transitions, extracted towards the three molecular emission peaks (CH₃OH, CH₃CCH, c-C₃H₂) and the dust peak of L1544, by using a circular aperture with a diameter corresponding to the telescope beam size (31"). CH₃CCH and CH₂DCCH are only represented with one transition each, as they are discussed separately in Sect. 4.3 and Fig. 4.2. The extraction locations of the spectra are marked in Fig. 4.1. The corresponding linewidths and velocities of the spectra at the four extraction locations are derived with Gaussian fitting and given in Table C.1.

The spectra show that the brightest emission of all observed molecules is detected towards the c-C₃H₂ peak located in the south of the core. In this region, the core is less protected from the ISRF, which drives photochemistry and leads to an active carbon chemistry and an increased abundance of C-bearing species. As all observed molecules in this study are essentially carbon chains, they are most efficiently formed in such conditions. However, the linewidths observed towards the CH₃CCH peak (and the dust peak) are

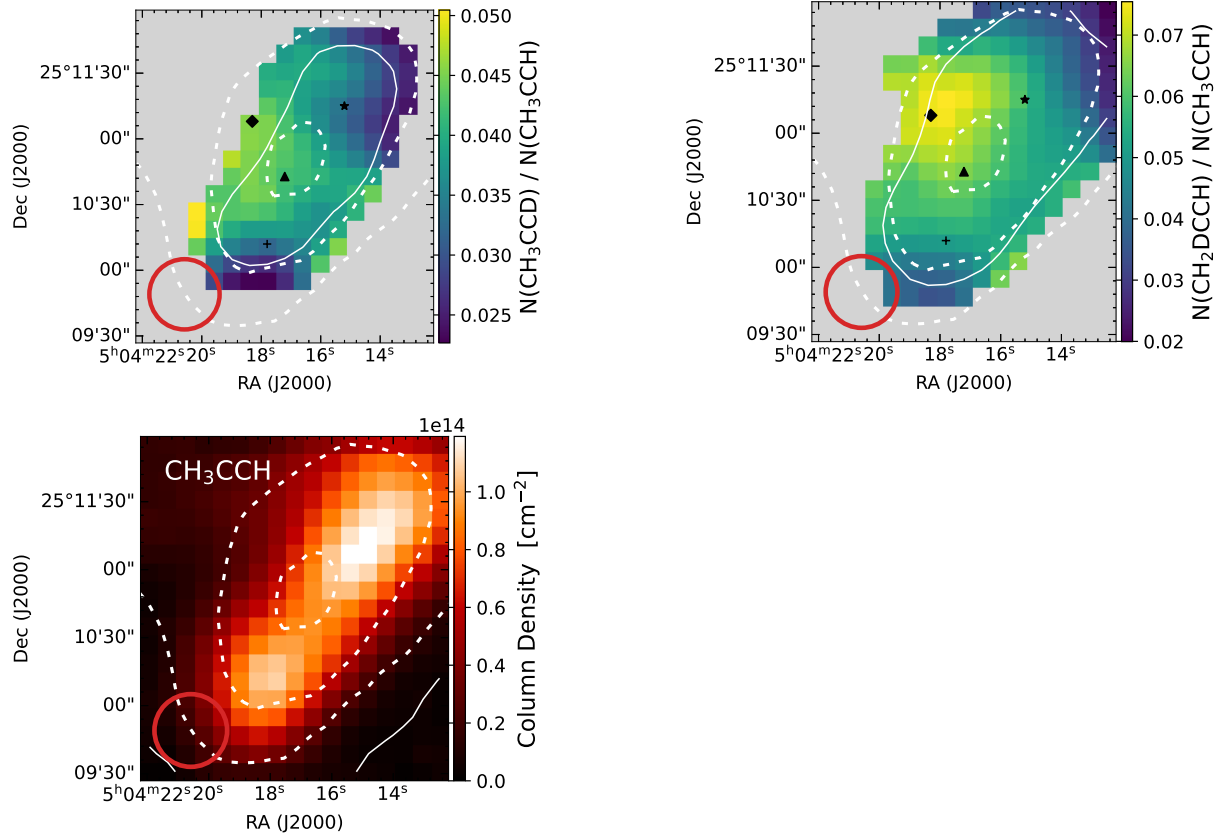


Figure C.1: *Top*: Deuteration maps derived from the CH_3CCH ($5_1 - 4_1$) transition. The solid contour indicates the 5σ confidence level of the ratios. The dashed line contours represent 90%, 50%, and 30% of the H_2 column density peak derived from *Herschel* maps (Spezzano et al. 2016). The markers in black represent the dust peak (triangle) and the molecular emission peaks of CH_3OH (diamond), CH_3CCH (star), and $\text{c-C}_3\text{H}_2$ (plus sign). *Bottom*: Column density map of CH_3CCH derived with the ($5_1 - 4_1$) transition. The solid contour indicates the 3σ level of the column density.

generally larger than towards the $\text{c-C}_3\text{H}_2$ peak. This indicates the location where the flow of material from the outer regions of the core (in the direction of the main filament) reaches the dense core and undergoes a local slow shock (as in the case of CH_3OH at the CH_3OH peak, see Lin et al. 2022a), causing broader linewidths. The spectra of DC_3N seem to show rather flat line profiles towards the CH_3CCH peak. As DC_3N is tracing less dense gas in the outer envelope of the core, the broader linewidth is likely caused by the larger turbulence present in those outer parts (e.g. Fuller & Myers 1992).

Table C.1: Comparison of the linewidths and velocities of the observed transitions at different locations in L1544 (dust peak, CH₃OH peak, CH₃CCH peak, and c-C₃H₂ peak, see Fig. 4.1). The parameters are derived from the spectra presented in Fig. C.2 using Gaussian fitting, and given in units of km s⁻¹.

Molecule	Transition	Dust peak		CH ₃ OH peak		CH ₃ CCH peak		c-C ₃ H ₂ peak	
		FWHM	V_{LSR}	FWHM	V_{LSR}	FWHM	V_{LSR}	FWHM	V_{LSR}
HC ₃ N	11 – 10	7.250(7)	0.197(7)	7.239(5)	0.180(5)	7.200(6)	0.204(6)	7.268(3)	0.152(3)
HCC ¹³ CN	10 – 9	7.19(2)	0.21(2)	7.17(4)	0.19(4)	7.13(2)	0.24(2)	7.239(6)	0.141(6)
DC ₃ N	10 – 9	7.22(1)	0.19(1)	7.18(2)	0.18(2)	7.141(9)	0.203(9)	7.285(4)	0.130(3)
CH ₃ CCH	5 ₀ – 4 ₀	7.24(3)	0.19(3)	7.19(2)	0.18(2)	7.17(3)	0.20(3)	7.26(2)	0.13(2)
CH ₃ CCD	6 ₀ – 5 ₀	7.17(4)	0.25(4)	7.18(5)	0.21(5)	7.07(3)	0.20(3)	7.26(2)	0.12(1)
CH ₂ DCCCH	6 ₀₆ – 5 ₀₅	7.151(9)	0.195(9)	7.10(2)	0.18(2)	7.062(8)	0.193(8)	7.229(6)	0.136(5)

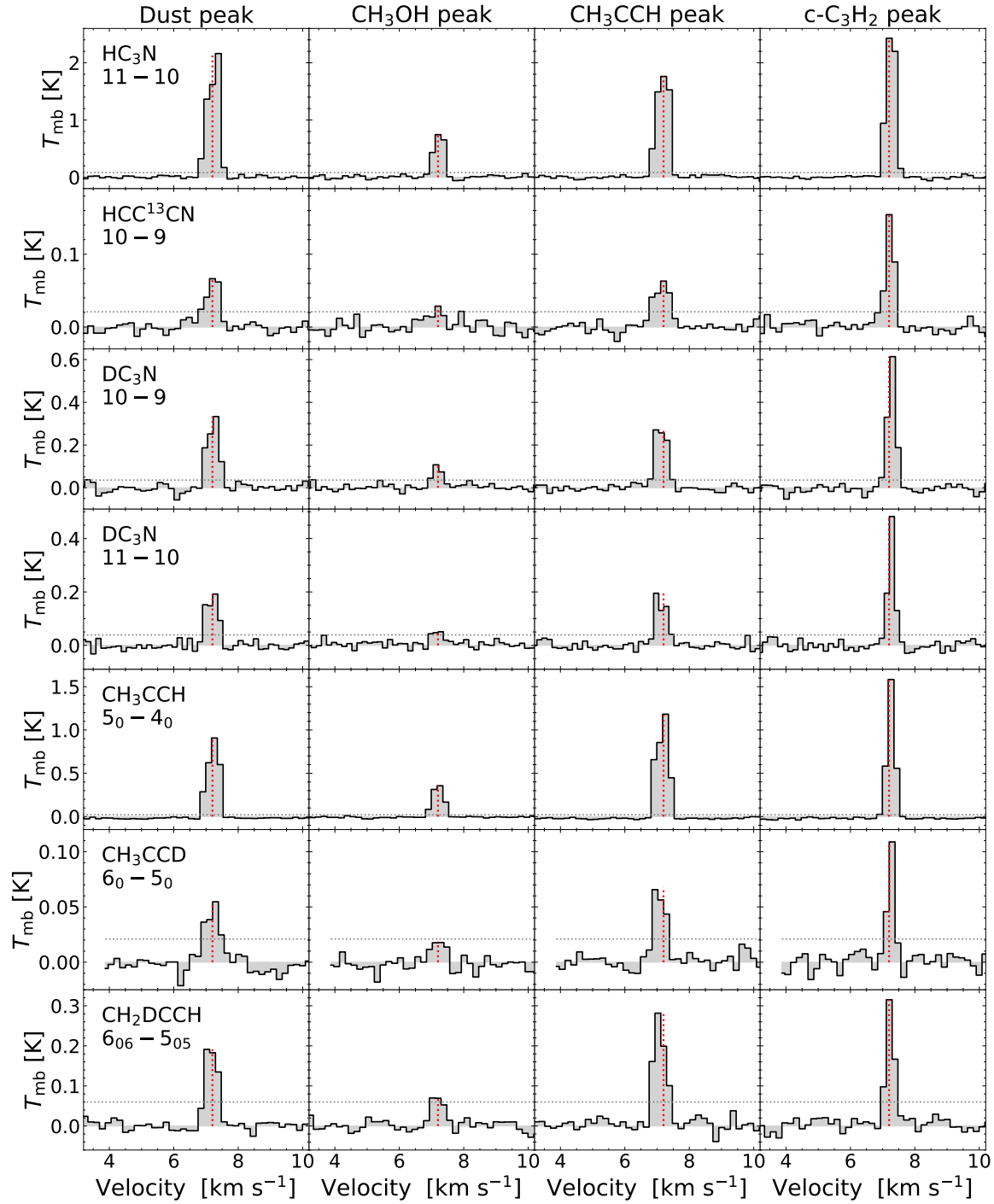


Figure C.2: Spectra of the observed transitions, extracted at the dust peak, the CH_3OH peak, the CH_3CCH peak, and the $\text{c-C}_3\text{H}_2$ peak of L1544. The extraction locations of the spectra and the telescope beam used as aperture are indicated in Fig. 4.1. The horizontal line indicates the 3σ confidence level, the vertical line shows the systemic velocity of the source.

Bibliography

- Abadi, M., Agarwal, A., Barham, P., et al. 2016, arXiv e-prints, arXiv:1603.04467
- Adams, W. S. 1941, ApJ, 93, 11
- Agúndez, M., Marcelino, N., Cernicharo, J., Roueff, E., & Tafalla, M. 2019, A&A, 625, A147
- Agúndez, M., Roueff, E., Cabezas, C., Cernicharo, J., & Marcelino, N. 2021, A&A, 649, A171
- Ahrens, V., Lewen, F., Takano, S., et al. 2002, Zeitschrift Naturforschung Teil A, 57, 669
- Alexander, M. H. & Dagdigian, P. J. 1985, The Journal of chemical physics, 83, 2191
- ALMA Partnership, Brogan, C. L., Pérez, L. M., et al. 2015, ApJl, 808, L3
- Altwegg, K., Balsiger, H., Bar-Nun, A., et al. 2015, Science, 347, 1261952
- Alves, F. O. & Franco, G. A. P. 2007, A&A, 470, 597
- André, P., Di Francesco, J., Ward-Thompson, D., et al. 2014, in Protostars and Planets VI, ed. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning, 27–51
- André, P., Men'shchikov, A., Bontemps, S., et al. 2010, A&A, 518, L102
- André, P., Revéret, V., Könyves, V., et al. 2016, A&A, 592, A54
- Andre, P., Ward-Thompson, D., & Barsony, M. 2000, in Protostars and Planets IV, ed. V. Mannings, A. P. Boss, & S. S. Russell, 59
- Arce, H. G. & Sargent, A. I. 2004, ApJ, 612, 342
- Arzoumanian, D., André, P., Didelon, P., et al. 2011, A&A, 529, L6
- Awad, M. & Khanna, R. 2015, Machine Learning (Berkeley, CA: Apress), 1–18
- Bauer, A. & Burie, J. 1969, C. R. Acad. Sci. Paris, B 268, 800
- Belitsky, V., Lapkin, I., Fredrixon, M., et al. 2015, A&A, 580, A29

- Ben Khalifa, M., Darna, B., & Loreau, J. 2024, *A&A*, 683, A53
- Bergin, E. A. & Tafalla, M. 2007, *ARA&A*, 45, 339
- Bianchi, E., Remijan, A., Codella, C., et al. 2023, *ApJ*, 944, 208
- Bizzocchi, L., Caselli, P., Spezzano, S., & Leonardo, E. 2014, *A&A*, 569, A27
- Bodenheimer, P. H. 2011, *Principles of Star Formation*
- Bonnor, W. B. 1956, *MNRAS*, 116, 351
- Bron, E., Daudon, C., Pety, J., et al. 2018, *A&A*, 610, A12
- Brünken, S., Fuchs, U., Lewen, F., et al. 2004, *Journal of Molecular Spectroscopy*, 225, 152
- Buhl, D. & Snyder, L. E. 1973, in *Molecules in the Galactic Environment*, ed. M. A. Gordon & L. E. Snyder, 187
- Busemann, H., Young, A. F., O’D. Alexander, C. M., et al. 2006, *Science*, 312, 727
- Cabezas, C., Agúndez, M., Marcelino, N., et al. 2021, *A&A*, 654, A45
- Calcutt, H., Willis, E. R., Jørgensen, J. K., et al. 2019, *A&A*, 631, A137
- Campello, R. J. G. B., Moulavi, D., & Sander, J. 2013, in *Advances in Knowledge Discovery and Data Mining*, ed. J. Pei, V. S. Tseng, L. Cao, H. Motoda, & G. Xu (Berlin, Heidelberg: Springer Berlin Heidelberg), 160–172
- Caselli, P., Keto, E., Bergin, E. A., et al. 2012, *ApJl*, 759, L37
- Caselli, P., Pineda, J. E., Sipilä, O., et al. 2022, *ApJ*, 929, 13
- Caselli, P., Pineda, J. E., Zhao, B., et al. 2019, *ApJ*, 874, 89
- Caselli, P., Walmsley, C. M., Tafalla, M., Dore, L., & Myers, P. C. 1999, *ApJl*, 523, L165
- Caselli, P., Walmsley, C. M., Zucconi, A., et al. 2002a, *ApJ*, 565, 331
- Caselli, P., Walmsley, C. M., Zucconi, A., et al. 2002b, *ApJ*, 565, 344
- Ceccarelli, C., Caselli, P., Bockelée-Morvan, D., et al. 2014, in *Protostars and Planets VI*, ed. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning, 859
- Chacón-Tanarro, A., Caselli, P., Bizzocchi, L., et al. 2019, *A&A*, 622, A141
- Chantzios, J., Spezzano, S., Caselli, P., et al. 2018, *ApJ*, 863, 126
- Chapillon, E., Dutrey, A., Guilloteau, S., et al. 2012, *ApJ*, 756, 58

- Cleeves, L. I., Bergin, E. A., Alexander, C. M. O. D., et al. 2014, *Science*, 345, 1590
- Colombo, D., Rosolowsky, E., Ginsburg, A., Duarte-Cabral, A., & Hughes, A. 2015, *MNRAS*, 454, 2067
- Colzi, L., Fontani, F., Rivilla, V. M., et al. 2018, *MNRAS*, 478, 3693
- Colzi, L., Sipilä, O., Roueff, E., Caselli, P., & Fontani, F. 2020, *A&A*, 640, A51
- Crapsi, A., Caselli, P., Walmsley, C. M., et al. 2005, *ApJ*, 619, 379
- Crapsi, A., Caselli, P., Walmsley, M. C., & Tafalla, M. 2007, *A&A*, 470, 221
- Crimier, N., Ceccarelli, C., Maret, S., et al. 2010, *A&A*, 519, A65
- Dagdigian, P. J. 2018, *MNRAS*, 479, 3227
- Dalgarno, A. & Lepp, S. 1984, *ApJl*, 287, L47
- Denis-Alpizar, O., Kalugina, Y., Stoecklin, T., Vera, M. H., & Lique, F. 2013, *The Journal of chemical physics*, 139, 224301
- Denis-Alpizar, O., Stoecklin, T., Dutrey, A., & Guilloteau, S. 2020, *MNRAS*, 497, 4276
- Draine, B. T. 2011, *Physics of the Interstellar and Intergalactic Medium*
- Drozdovskaya, M. N., Coudert, L. H., Margulès, L., et al. 2022, *A&A*, 659, A69
- Drozdovskaya, M. N., Schroeder I, I. R. H. G., Rubin, M., et al. 2021, *MNRAS*, 500, 4901
- Dubernet, M. L., Alexander, M. H., Ba, Y. A., et al. 2013, *A&A*, 553, A50
- Dullemond, C. P., Juhasz, A., Pohl, A., et al. 2012, *RADMC-3D: A multi-purpose radiative transfer tool*, *Astrophysics Source Code Library*, record ascl:1202.015
- Dumouchel, F., Klos, J., & Lique, F. 2011, *Physical Chemistry Chemical Physics*, 13, 8204
- Dunham, M. M., Allen, L. E., Evans, Neal J., I., et al. 2015, *ApJs*, 220, 11
- Dunham, M. M., Crapsi, A., Evans, Neal J., I., et al. 2008, *ApJs*, 179, 249
- Dunham, Jr., T. 1937, *PASP*, 49, 26
- Ebert, R. 1955, *ZAp*, 37, 217
- EMAA. 2021, UGA, CNRS, CNRS-INSU, OSUG
- Emprechtinger, M., Caselli, P., Volgenau, N. H., Stutzki, J., & Wiedner, M. C. 2009, *A&A*, 493, 89

- Enoch, M. L., Young, K. E., Glenn, J., et al. 2006, *ApJ*, 638, 293
- Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. 1996, *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, KDD-96* (AAAI Press), 226
- Evans, Neal J., I. 1999, *ARA&A*, 37, 311
- Faure, A., Lique, F., & Wiesenfeld, L. 2016, *MNRAS*, 460, 2103
- Ferrer Asensio, J., Spezzano, S., Caselli, P., et al. 2022, *A&A*, 667, A119
- Fontani, F., Cesaroni, R., Caselli, P., & Olmi, L. 2002, *A&A*, 389, 603
- Fotopoulou, S. 2024, *Astronomy and Computing*, 48, 100851
- Fuchs, U., Bruenken, S., Fuchs, G. W., et al. 2004, *Zeitschrift Naturforschung Teil A*, 59, 861
- Fuller, G. A. & Myers, P. C. 1992, *ApJ*, 384, 523
- Furuya, K., Aikawa, Y., Sakai, N., & Yamamoto, S. 2011, *ApJ*, 731, 38
- Galli, P. A. B., Loinard, L., Bouy, H., et al. 2019, *A&A*, 630, A137
- Galli, P. A. B., Loinard, L., Ortiz-Léon, G. N., et al. 2018, *ApJ*, 859, 33
- Gerin, M., Combes, F., Wlodarczak, G., Encrenaz, P., & Laurent, C. 1992, *A&A*, 253, L29
- Giers, K., Spezzano, S., Alves, F., et al. 2022, *A&A*, 664, A119
- Giers, K., Spezzano, S., Caselli, P., et al. 2023, *A&A*, 676, A78
- Giers, K., Spezzano, S., Lin, Y., et al. 2025, *A&A*, 699, A103
- Ginsburg, A., Koch, E., Robitaille, T., et al. 2019, *radio-astro-tools/spectral-cube: v0.4.4*
- Ginsburg, A. & Mirocha, J. 2011, *PySpecKit: Python Spectroscopic Toolkit*, *Astrophysics Source Code Library*, record ascl:1109.001
- Ginsburg, A., Sokolov, V., de Val-Borro, M., et al. 2022, *AJ*, 163, 291
- Goicoechea, J. R., Lique, F., & Santa-Maria, M. G. 2022, *A&A*, 658, A28
- Goldsmith, P. F. 2001, *ApJ*, 557, 736
- Goldsmith, P. F. & Langer, W. D. 1999, *ApJ*, 517, 209
- Gratier, P., Majumdar, L., Ohishi, M., et al. 2016, *ApJS*, 225, 25

- Green, S. 1975, *The Journal of Chemical Physics*, 62, 2271
- Gueth, F. & Guilloteau, S. 1999, *A&A*, 343, 571
- Guzmán, A. E., Guzmán, V. V., Garay, G., Bronfman, L., & Hechenleitner, F. 2018, *ApJS*, 236, 45
- Harju, J., Sipilä, O., Brünken, S., et al. 2017, *ApJ*, 840, 63
- Hatchell, J., Millar, T. J., & Rodgers, S. D. 1998, *A&A*, 332, 695
- Herbst, E. & Klemperer, W. 1973, *ApJ*, 185, 505
- Herbst, E. & van Dishoeck, E. F. 2009, *ARA&A*, 47, 427
- Hernández Vera, M., Lique, F., Dumouchel, F., Hily-Blant, P., & Faure, A. 2017, *MNRAS*, 468, 1084
- Hickson, K. M., Wakelam, V., & Loison, J.-C. 2016, *Molecular Astrophysics*, 3, 1
- Hildebrand, R. H. 1983, *QJRAS*, 24, 267
- Hily-Blant, P., Bonal, L., Faure, A., & Quirico, E. 2013, *Icarus*, 223, 582
- Hily-Blant, P., Faure, A., Vastel, C., et al. 2018, *MNRAS*, 480, 1174
- Hily-Blant, P., Walmsley, M., Pineau Des Forêts, G., & Flower, D. 2010, *A&A*, 513, A41
- Hirota, T., Ikeda, M., & Yamamoto, S. 2003, *ApJ*, 594, 859
- Hirota, T., Ito, T., & Yamamoto, S. 2002, *ApJ*, 565, 359
- Hotelling, H. 1936, *Biometrika*, 28, 321
- Howe, D. A., Millar, T. J., Schilke, P., & Walmsley, C. M. 1994, *MNRAS*, 267, 59
- Imai, M., Sakai, N., López-Sepulcre, A., et al. 2018, *ApJ*, 869, 51
- IRAM. 2025, <https://publicwiki.iram.es>
- Jensen, S. S., Spezzano, S., Caselli, P., Grassi, T., & Haugbølle, T. 2023, *A&A*, 675, A34
- Jiménez-Serra, I., Vasyunin, A. I., Caselli, P., et al. 2016, *ApJ*, 830, L6
- Jørgensen, J. K., Schöier, F. L., & van Dishoeck, E. F. 2004, *A&A*, 416, 603
- Juvela, M. 2020, *A&A*, 644, A151
- Keto, E. & Caselli, P. 2008, *ApJ*, 683, 238
- Keto, E. & Caselli, P. 2010, *MNRAS*, 402, 1625

- Keto, E., Caselli, P., & Rawlings, J. 2015, MNRAS, 446, 3731
- Kong, S., Caselli, P., Tan, J. C., Wakelam, V., & Sipilä, O. 2015, ApJ, 804, 98
- Koumpia, E., Semenov, D. A., van der Tak, F. F. S., Boogert, A. C. A., & Caux, E. 2017, A&A, 603, A88
- Kuffmeier, M., Jensen, S. S., & Haugbølle, T. 2023, European Physical Journal Plus, 138, 272
- Lada, C. J., Bergin, E. A., Alves, J. F., & Huard, T. L. 2003, ApJ, 586, 286
- Langer, W. D., Graedel, T. E., Frerking, M. A., & Armentrout, P. B. 1984, ApJ, 277, 581
- Langer, W. D., Schloerb, F. P., Snell, R. L., & Young, J. S. 1980, ApJ, 239, L125
- Lanza, M. & Lique, F. 2014, The Journal of chemical physics, 141, 164321
- Larson, R. B. 1969, MNRAS, 145, 271
- Larson, R. B. 1981, MNRAS, 194, 809
- Launhardt, R., Stutz, A. M., Schmiedeke, A., et al. 2013, A&A, 551, A98
- Lee, C.-F., Li, Z.-Y., Hirano, N., et al. 2018, ApJ, 863, 94
- Lee, C. W., Myers, P. C., & Tafalla, M. 2001, ApJs, 136, 703
- Li, C., Wang, H., Zhang, M., et al. 2018, ApJs, 238, 10
- Lin, Y., Spezzano, S., Sipilä, O., Vasyunin, A., & Caselli, P. 2022a, A&A, 665, A131
- Lin, Y., Wyrowski, F., Liu, H. B., et al. 2022b, A&A, 658, A128
- Linsky, J. L. 2003, , 106, 49
- Loison, J.-C., Wakelam, V., Gratier, P., & Hickson, K. M. 2020, MNRAS, 498, 4663
- Magalhães, V. S., Hily-Blant, P., Faure, A., Hernandez-Vera, M., & Lique, F. 2018, A&A, 615, A52
- Mangum, J. G., Emerson, D. T., & Greisen, E. W. 2007, A&A, 474, 679
- Mangum, J. G. & Shirley, Y. L. 2015, PASP, 127, 266
- Marchand, P., Coutens, A., Scigliuto, J., et al. 2024, A&A, 687, A195
- Markwick, A. J., Charnley, S. B., Butner, H. M., & Millar, T. J. 2005, ApJ, 627, L117
- Mathews, G. S., Klaassen, P. D., Juhász, A., et al. 2013, A&A, 557, A132

- Mccarthy, M. C., Gottlieb, C. A., & Thaddeus, P. 1995, *Journal of Molecular Spectroscopy*, 173, 303
- McCaughrean, M. J., Rayner, J. T., & Zinnecker, H. 1994, *ApJl*, 436, L189
- McGuire, B. A. 2022, *ApJS*, 259, 30
- McInnes, L., Healy, J., & Astels, S. 2017, *The Journal of Open Source Software*, 2
- McKee, C. F. & Ostriker, E. C. 2007, *ARA&A*, 45, 565
- McKee, C. F. & Ostriker, J. P. 1977, *ApJ*, 218, 148
- Milam, S. N., Savage, C., Brewster, M. A., Ziurys, L. M., & Wyckoff, S. 2005, *ApJ*, 634, 1126
- Millar, T. J., Bennett, A., Rawlings, J. M. C., Brown, P. D., & Charnley, S. B. 1991, , 87, 585
- Mladenović, M. & Roueff, E. 2014, *A&A*, 566, A144
- Morris, M., Turner, B. E., Palmer, P., & Zuckerman, B. 1976, *ApJ*, 205, 82
- Motte, F. & André, P. 2001, *A&A*, 365, 440
- Moulavi, D., Jaskowiak, P. A., Campello, R. J. G. B., Zimek, A., & Sander, J. 2014, *Density-Based Clustering Validation*, 839–847
- Müller, H. S. P., Thorwirth, S., Roth, D. A., & Winnewisser, G. 2001, *A&A*, 370, L49
- Nagy, Z., Spezzano, S., Caselli, P., et al. 2019, *A&A*, 630, A136
- Navarro-Almaida, D., Bop, C., Lique, F., et al. 2022, *arXiv preprint arXiv:2212.07675*
- Nomura, H., Furuya, K., Cordiner, M. A., et al. 2022, *arXiv e-prints*, arXiv:2203.10863
- Öberg, K. I. & Bergin, E. A. 2021, *Phys. Rep.*, 893, 1
- Öberg, K. I., Boamah, M. D., Fayolle, E. C., et al. 2013, *ApJ*, 771, 95
- Ohashi, N., Lee, S. W., Wilner, D. J., & Hayashi, M. 1999, *ApJl*, 518, L41
- Okoda, Y., Oya, Y., Abe, S., et al. 2021, *ApJ*, 923, 168
- Okoda, Y., Oya, Y., Sakai, N., Watanabe, Y., & Yamamoto, S. 2020, *ApJ*, 900, 40
- Ortiz-León, G. N., Loinard, L., Dzib, S. A., et al. 2018, *ApJ*, 865, 73
- Padovani, M., Walmsley, C. M., Tafalla, M., Galli, D., & Müller, H. S. P. 2009, *A&A*, 505, 1199

- Padovani, M., Walmsley, C. M., Tafalla, M., Hily-Blant, P., & Pineau Des Forêts, G. 2011, *A&A*, 534, A77
- Pagani, L., Bourgoïn, A., & Lique, F. 2012, *A&A*, 548, L4
- Pagani, L., Salez, M., & Wannier, P. G. 1992, *A&A*, 258, 479
- Parise, B., Castets, A., Herbst, E., et al. 2004, *A&A*, 416, 159
- Parker, D. S. N. & Kaiser, R. I. 2017, *Chem. Soc. Rev.*, 46, 452
- Paszke, A., Gross, S., Massa, F., et al. 2019, arXiv e-prints, arXiv:1912.01703
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, *Journal of Machine Learning Research*, 12, 2825
- Penston, M. V. 1969, *MNRAS*, 144, 425
- Persson, M. V. 2014
- Persson, M. V., Jørgensen, J. K., Müller, H. S. P., et al. 2018, *A&A*, 610, A54
- Pety, J. 2005, in *SF2A-2005: Semaine de l’Astrophysique Française*, ed. F. Casoli, T. Contin, J. M. Hameury, & L. Pagani, 721
- Pickett, H. M., Poynter, R. L., Cohen, E. A., et al. 1998, *J. Quant. Spectr. Rad. Transf.*, 60, 883
- Pineda, J. E., Arzoumanian, D., Andre, P., et al. 2023, in *Astronomical Society of the Pacific Conference Series*, Vol. 534, *Protostars and Planets VII*, ed. S. Inutsuka, Y. Aikawa, T. Muto, K. Tomida, & M. Tamura, 233
- Pineda, J. E., Arzoumanian, D., André, P., et al. 2022, arXiv e-prints, arXiv:2205.03935
- Pineda, J. E., Segura-Cox, D., Caselli, P., et al. 2020, *Nature Astronomy*, 4, 1158
- Pineda, J. E., Zhao, B., Schmiedeke, A., et al. 2019, *ApJ*, 882, 103
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, *A&A*, 586, A138
- Plummer, H. C. 1911, *MNRAS*, 71, 460
- Punanova, A., Caselli, P., Feng, S., et al. 2018, *ApJ*, 855, 112
- Puzzarini, C. 2022, *Frontiers in Astronomy and Space Sciences*, Volume 8 - 2021
- Quénard, D., Bottinelli, S., Caux, E., & Wakelam, V. 2018, *MNRAS*, 477, 5312
- Quénard, D., Vastel, C., Ceccarelli, C., et al. 2017, *MNRAS*, 470, 3194

- Redaelli, E., Bizzocchi, L., Caselli, P., & Pineda, J. E. 2023, *A&A*, 674, L8
- Redaelli, E., Bizzocchi, L., Caselli, P., et al. 2019, *A&A*, 629, A15
- Redaelli, E., Chacón-Tanarro, A., Caselli, P., et al. 2022, *ApJ*, 941, 168
- Redaelli, E., Sipilä, O., Padovani, M., et al. 2021, *A&A*, 656, A109
- Rivilla, V. M., Colzi, L., Fontani, F., et al. 2020, *MNRAS*, 496, 1990
- Robert, F. 2003, *SSR*, 106, 87
- Roberts, H., Fuller, G. A., Millar, T. J., Hatchell, J., & Buckle, J. V. 2002, *A&A*, 381, 1026
- Robitaille, T. P. & Whitney, B. A. 2010, *ApJL*, 710, L11
- Roueff, E., Loison, J. C., & Hickson, K. M. 2015, *A&A*, 576, A99
- Sakai, N., Sakai, T., Hirota, T., & Yamamoto, S. 2008, *ApJ*, 672, 371
- Sakai, N., Sakai, T., Hirota, T., & Yamamoto, S. 2009, *ApJ*, 702, 1025
- Sakai, N., Saruwatari, O., Sakai, T., Takano, S., & Yamamoto, S. 2010, *A&A*, 512, A31
- Saykally, R. J., Szanto, P. G., Anderson, T. G., & Woods, R. C. 1976, *ApJL*, 204, L143
- Schiff, H. I. & Bohme, D. K. 1979, *ApJ*, 232, 740
- Schmid-Burgk, J., Muters, D., Müller, H. S. P., & Brupbacher-Gatehouse, B. 2004, *A&A*, 419, 949
- Schöier, F. L., van der Tak, F. F. S., van Dishoeck, E. F., & Black, J. H. 2005, *A&A*, 432, 369
- Scibelli, S., Shirley, Y., Megías, A., & Jiménez-Serra, I. 2024, *MNRAS*, 533, 4104
- Segura-Cox, D. M., Harris, R. J., Tobin, J. J., et al. 2016, *ApJL*, 817, L14
- Segura-Cox, D. M., Schmiedeke, A., Pineda, J. E., et al. 2020, , 586, 228
- Semenov, D., Hersant, F., Wakelam, V., et al. 2010, *A&A*, 522, A42
- Sheehan, P. D., Tobin, J. J., Federman, S., Megeath, S. T., & Looney, L. W. 2020, *The Astrophysical Journal*, 902, 141
- Shimajiri, Y., Kitamura, Y., Saito, M., et al. 2014, *A&A*, 564, A68
- Shingledecker, C. N., Lee, K. L. K., Wandishin, J. T., et al. 2021, *A&A*, 652, L12

- Shu, F. H. 1977, *ApJ*, 214, 488
- Sipilä, O., Caselli, P., & Harju, J. 2015, *A&A*, 578, A55
- Sipilä, O., Caselli, P., Redaelli, E., Juvela, M., & Bizzocchi, L. 2019, *MNRAS*, 487, 1269
- Sipilä, O., Caselli, P., Redaelli, E., & Spezzano, S. 2022, *A&A*, 668, A131
- Sipilä, O., Colzi, L., Roueff, E., et al. 2023, *A&A*, 678, A120
- Soler, J. D., Hennebelle, P., Martin, P. G., et al. 2013, *ApJ*, 774, 128
- Spezzano, S., Bizzocchi, L., Caselli, P., Harju, J., & Brünken, S. 2016, *A&A*, 592, L11
- Spezzano, S., Brünken, S., Schilke, P., et al. 2013, *ApJl*, 769, L19
- Spezzano, S., Caselli, P., Bizzocchi, L., Giuliano, B. M., & Lattanzi, V. 2017, *A&A*, 606, A82
- Spezzano, S., Caselli, P., Pineda, J. E., et al. 2020, *A&A*, 643, A60
- Spezzano, S., Caselli, P., Sipilä, O., & Bizzocchi, L. 2022a, *A&A*, 664, L2
- Spezzano, S., Caselli, P., Sipilä, O., & Bizzocchi, L. 2022b, *A&A*, 664, L2
- Stahler, S. W. & Palla, F. 2004, *The Formation of Stars*
- Sun, K., Kramer, C., Ossenkopf, V., et al. 2006, *A&A*, 451, 539
- Tafalla, M., Mardones, D., Myers, P. C., et al. 1998, *ApJ*, 504, 900
- Tafalla, M. & Santiago, J. 2004, *A&A*, 414, L53
- Taniguchi, K., Herbst, E., Ozeki, H., & Saito, M. 2019, *ApJ*, 884, 167
- Terebey, S., Shu, F. H., & Cassen, P. 1984, *ApJ*, 286, 529
- Thaddeus, P., Vrtillek, J. M., & Gottlieb, C. A. 1985, *ApJl*, 299, L63
- Tielens, A. 2021, *Molecular Astrophysics*
- Tinti, F., Bizzocchi, L., Degli Esposti, C., & Dore, L. 2007, *ApJl*, 669, L113
- Turner, B. E. 1971, *ApJ*, 163, L35
- Turner, B. E. 2001, *ApJs*, 136, 579
- Turner, B. E., Terzieva, R., & Herbst, E. 1999, *ApJ*, 518, 699
- Valdivia-Mena, M. T., Pineda, J. E., Segura-Cox, D. M., et al. 2023, *A&A*, 677, A92

- van der Tak, F. F. S., Lique, F., Faure, A., Black, J. H., & van Dishoeck, E. F. 2020, *Atoms*, 8, 15
- van der Tak, F. F. S., Müller, H. S. P., Harding, M. E., & Gauss, J. 2009, *A&A*, 507, 347
- van Dishoeck, E. F., Blake, G. A., Jansen, D. J., & Groesbeck, T. D. 1995, *ApJ*, 447, 760
- Vastel, C., Caselli, P., Ceccarelli, C., et al. 2006, *ApJ*, 645, 1198
- Vastel, C., Ceccarelli, C., Lefloch, B., & Bachiller, R. 2014, *ApJ*, 795, L2
- Vasyunin, A. I., Caselli, P., Dulieu, F., & Jiménez-Serra, I. 2017, *ApJ*, 842, 33
- Wakelam, V., Herbst, E., Loison, J. C., et al. 2012, *ApJs*, 199, 21
- Wakelam, V., Loison, J. C., Herbst, E., et al. 2015, *ApJs*, 217, 20
- Walker, G. W., Kalinauskaitė, E., McCarthy, D. N., et al. 2016, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 9914, Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VIII, ed. W. S. Holland & J. Zmuidzinas, 99142V
- Walmsley, C. M., Flower, D. R., & Pineau des Forêts, G. 2004, *A&A*, 418, 1035
- Walmsley, C. M., Winnewisser, G., & Toelle, F. 1980, *A&A*, 81, 245
- Ward-Thompson, D., Motte, F., & Andre, P. 1999, *MNRAS*, 305, 143
- Watanabe, N. & Kouchi, A. 2008, *Progress In Surface Science*, 83, 439
- Watson, W. D. 1976, *Reviews of Modern Physics*, 48, 513
- Wes McKinney. 2010, in *Proceedings of the 9th Python in Science Conference*, ed. Stéfan van der Walt & Jarrod Millman, 56 – 61
- Whitworth, A. P. & Ward-Thompson, D. 2001, *ApJ*, 547, 317
- Williams, J. P., Myers, P. C., Wilner, D. J., & Di Francesco, J. 1999, *ApJL*, 513, L61
- Wilson, T. L. 1999, *Reports on Progress in Physics*, 62, 143
- Xu, L.-H. & Lovas, F. 1997, *J. Phys. Chem. Ref. Data*, 26, 17
- Yamagishi, M., Hara, C., Kawabe, R., et al. 2019, *ApJ*, 875, 62
- Yamamoto, S. 2017, *Introduction to Astrochemistry: Chemical Evolution from Interstellar Clouds to Star and Planet Formation* (Springer Japan)
- Yan, Q.-Z., Yang, J., Su, Y., et al. 2022, *AJ*, 164, 55

Yoshida, K., Sakai, N., Nishimura, Y., et al. 2019, PASJ, 71, S18

Young, A. K. 2023, *Frontiers in Astronomy and Space Sciences*, Volume 10 - 2023

Yun, H.-S. & Lee, J.-E. 2023, ApJ, 958, 113

Zhang, C., Wu, Y., Liu, X. C., et al. 2021, A&A, 648, A83

Zuckerman, B. & Palmer, P. 1974, ARA&A, 12, 279

Danksagung

In your acknowledgements, you are supposed to feel grateful, but mostly, I am just tired. My PhD journey has been a long one, and it feels unreal to finally stand (or in my case rather: sit) in front of the end. My past four years have been shaped by extraordinary circumstances: from the global catastrophe of a viral pandemic to my own personal catastrophe of becoming chronically ill. Pursuing a PhD is a tough and challenging adventure; suddenly being chronically ill is a devastating, life-changing experience. Going through both simultaneously makes your life a constant struggle, full of daily challenges, often failing in both. I tried to do my best, given the circumstances I had to deal with. But looking back on two and a half years filled with suffering, haze, and exhaustion makes me wonder what could have been, grieving the scientist I could have become. Nevertheless, my PhD gave me the opportunity to do amazing things, visit beautiful places, meet wonderful people, learn nerdy stuff, and follow my curiosity, and I have grown a lot, both on a professional and a personal level. I am, always have been, and always will be very proud to be a part of the Max Planck scientific community.

According to my parents, I was an astro-girl from very early on, and loved to read all the books with the pretty, colourful space pictures in them. So first and foremost, thank you to my wonderful parents, who were with me on every part of my journey, on every turn I took. Thank you for your trust, support (both emotional & financial), guidance, pride, and your unconditional love. Thank you for giving me the opportunity to find and follow my passions.

Thank you to my "official supervisor", the great Paola. Thank you for giving me the opportunity to pursue a PhD in your group, and for your guidance throughout those challenging years. Thank you for your never-ending fascination and joy in astrochemistry (or anything else even vaguely scientific), which is always a true inspiration and motivation for a struggling student. And thank you for your support and understanding for my unusual circumstances. Thank you also to my "daily supervisor", Silvia, for guiding me through my PhD and teaching me all about prestellar cores and deuteration. It has been quite a ride, with projects exploding in size and time, tricky data and weird molecules. Thank you for all the joint trouble-shooting and for answering all my naive or stupid or complicated or overwhelmed questions. Thank you for your patience and your trust in me – and for sharing L1544!

Thank you to all my collaborators during my different projects – Felipe Alves, my second master thesis supervisor and confidence bringer, Elena Redaelli, my go-to paper

author and expert on random science questions, Olli Sipilä, my personal abundance profile provider and trusted chemical modeller, Eva Wirström, my lovely committee member, Yuxin Lin, my machine learning expert, and Jaime Pineda, my python support center.

Thank you to all members of the Centre for Astrochemical Studies that have accompanied me throughout my years. Thank you for being such an open, welcoming, supporting and generally amazing group. You are the best! Thank you to all the lunch people (in particular to Birgitta and Pavol for taking care of the "lunch gathering"), for enjoying outdoor and indoor lunches, and for interesting discussions, particularly about language. A special thank you goes to my lovely office mates that made the daily struggles more bearable: Davide, Paula, Laura, and Hannah. And a very special thank you to Angie, the CAS angel. Thank you for being the backbone of CAS, always ready to help, answer questions and organise urgent things. But also thank you for being the heart of CAS, for always taking the time to chat, for always having an open ear.

Thank you to all my wonderful friends, spread across the world (but mostly Baden-Württemberg), for keeping me sane – love you all! Thank you Teresa, for the time we shared at CAS – for your emotional support, your coding advice, fun at conferences, our art teas, and of course the eurovision parties. Thank you Caro, for extensive shopping and lunching at the OEZ, for christmas cookies orgies and Harry Potter parties. Thank you to my sibling, Isabell, for countless Hogwarts battles, creative writing fun and music sessions. Thank you Klara, for all the visits, birthdays, trips, hikes, and deep talks. Thank you Lydia, for complaining office hours, dinner-breakfast gossips, and introducing me to Sushi – Tomatensalat! Thank you Jessy, for cooking adventures, crafting & Co, and extended video calls.

Thank you Matti. My love. My partner in crime, my partner in everything. You have been my rock in the surf during the past years – without you, I would not have made it this far. Thank you for your never-ending support, your trust in me, your love, and for choosing me every single day. You are my home, and I'm looking forward to sharing all the big and small moments of life with you that the future might bring us.

And last but not least: thank you to myself. Girl, you did it.