UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
MAGÍSTER EN ASTRONOMÍA

Observaciones de moléculas deuteradas

Observations of deuterated molecules

Profesor Guía: Dr. Stefano Bovino
Departamento de Astronomía
Facultad de Ciencias Físicas y Matemáticas
Universidad de Concepción

Tesis para ser presentada a la Dirección de Postgrado de la Universidad de Concepción

MORIN ANTONIA ÓRDENES ESCOBAR
CONCEPCION - CHILE 2021

UNIVERSIDAD DE CONCEPCIÓN

Abstract

Facultad de Ciencias Físicas y Matemáticas
Departmento de Astronomía

MSc.

Observations of deuterated molecules

by Morin ÓRDENES ESCOBAR

The deuterium enrichment of key gas-phase molecules like H_3^+ is thought to be boosted in massive star forming regions due to the high densities and low temperatures inside these regions. It has been suggested that ortho-H_2D^+, the main deuterated ion in these regions, might work as a chemical clock, but there are still a lot of unanswered questions about its effectiveness in probing the evolutionary stages of the high-mass star formation process.
In this thesis we study a large sample of o-H_2D^+ observations in different high-mass star-forming regions, presenting possible empirical correlations with physical characteristics with the aim of clarifying the relation between the presence of this deuterated molecule and the evolutionary stage of star-forming regions. Our sample consists of APEX observations of the o-H_2D^+ ground-state transition observed in the ATLASGAL survey of high-mass clumps. We obtained column densities and abundances of o-H_2D^+ by performing a Bayesian fitting of the data. The o-H_2D^+ emission line was detected in 16 sources and we found some clear correlations between the N(o-H_2D^+) column density and the bolometric luminosity, dust temperature and a slight correlation with the CO-depletion factor f_d. We also found a clear correlation with the luminosity-to-mass ratio L/M, suggested to be a good tracer of the evolution of high-mass star forming regions. We conclude that the H_2D^+ tracer is very abundant in the early stages of star formation, and suggest that the deuteration process is strongly dependent on the time evolution of the clumps. Together with N_2D^+ this can be used to trace the different evolutionary stages of the high-mass star-formation process and provide relevant information about their timescales.

Acknowledgements

First, I would like to thank my family for their infinite support during all these years of study, without their trust in me I would not have been able to accomplish anything.
I want to thank all the friends I made in these years at Universidad de Concepción, specially the ones that gave me support in the hardest weeks of the semester and who were always available for a study night at the library. I want to thank all the professors that were always willing to answer questions and give support and to all my friends in the thesis room, who always knew when it was a perfect moment for taking a break.
I would like to thank all the people in the Astro-UdeC Theory Group, including professors and students. Specially to the members of the Astrochemistry Group, the place where I could always find encouraging words, great suggestions on the presentations and also very good friends.
I would like to express my full gratitude to Stefano Bovino and Giovanni Sabatini for their support, commitment and patience in every step of this project. Your guidance was very important to me through these years.
And finally I would like to thank the Astronomy department and the Dirección de Postgrado for the financial help in this process.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The study of stars has been a relevant field for the scientific community for several decades, and these days is still an active field of study because of the important effect of stars into their galactic ecosystem. The contribution of stars to their surroundings consists mostly on the stellar feedback that is released during their lives, including the huge amount of energy released when these objects die. In the specific case of massive stars, they die in a violent supernovae explosion where a great amount of material and energy is released in a short period of time, instantly supplying the interstellar gas with a wide variety of chemical species and producing perturbations even on distant objects. Understanding how these stars form and evolve is then fundamental to understand the lifecycle of the interstellar medium.

Stars form by gravitational contraction of molecular cloud cores. The evolution of these cores and the subsequent formation of a star depends on the mass of these regions. There are different formation scenarios depending on the final mass of the stars, from stars with masses of a few solar masses to stars with more than 8 $M_{\odot}$, called massive stars. During the early phases of star formation, the volume density of molecular hydrogen, $n(H_2) > 10^4$ cm^{-3}, and the temperature of the clouds are $T < 20$ K, reaching the ideal conditions to favour processes like freeze-out or depletion. Several C-, N- and O- bearing species are efficiently removed from the gas phase and frozen onto the surface of dust grains, especially CO. This segregation of CO enhances the enrichment of the deuterium atoms, compared to the hydrogen in the molecules, boosting the abundance of o-H_2D^+ and other deuterated species.

The measurements of timescales in processes like star formation are crucial to distinguish between the competing star formation theories. Chemistry is a powerful tool to infer how long the star forming gas remains cold and especially deuteration has been suggested to be a reliable chemical clock of star-forming regions due to its strong dependence on density and temperature changes during cloud contraction.

In this work we will present o-H_2D^+ APEX observations in different sources from a subsample of the ATLASGAL survey. This represents the largest sample of o-H_2D^+ in high-mass star-forming regions to date. We have compared the N(o-H_2D^+) obtained from the detected spectra with different physical parameters of the clumps to look for possible correlations, which can confirm the reliability of o-H_2D^+ as a chemical clock.

This thesis is structured as follows: The first chapter is an introduction containing a summary of the most important properties of the interstellar medium (ISM), introduce the TOP100 sample we used in this study and its properties. The chapter 2 reviews the chemical behaviour of the ISM, focusing mainly on the molecular clouds chemistry and their most relevant reactions. In the following chapters we review the sample and the sources we studied, explain all the data reduction and the code operation. In chapter 4 we check the different parameters that we compared with the derived o-H_2D^+ column density and show the correlations we found. In the last chapter there are

the summary and conclusions from this work.

1.1 Interstellar medium

The space between the stars that once was thought to be empty, is actually filled with the interstellar medium. This interstellar matter can be either found in a gaseous state, like atoms, ions and molecules, or in a solid state, i.e. micro-sized particles called dust grains. From this mix of dust and gas in the interstellar medium stars, the dominant source of energy, form. How the ISM is organized in the Milky Way and in other galaxies depend on different conditions like temperature, density, ionization, etc. At early times, the baryonic mass of the galaxies is mainly the gas of the interstellar medium. But as galaxies evolve, the interstellar medium is gradually converted to stars, and some of this gas is ejected in form of galactic winds. Nowadays, roughly 10% of the baryonic mass of the Milky Way is found in the interstellar medium. The ISM contain a number of phases characterized by different temperatures, densities, and ionization fractions. Those phases are the Cold Neutral Medium (CNM), the Warm Neutral Medium (WNM), the Warm Ionized Medium (WIM), the HII regions, the Coronal gas and the Dense Molecular gas.

1.1.1 Phases of the ISM

The ISM is dynamic, the particles can undergo changes of phases that result from the interplay of radiation, large scale flows, magnetic fields and turbulence in the gas. All the phases are close to pressure equilibrium, and the hottest phases have low average particle densities. The different phases will be described in the following paragraphs.

- Coronal gas: Gas that has been collisionally ionized, by the incoming shocks originated by the stellar winds that come from early type stars and supernova explosions. This gas phase has temperatures around $10^{5.5}$-10^{7} K. Most of the coronal gas has low density, filling around half of the volume of the galactic disk. These regions may have characteristic dimensions of 20 pc and may be connected to other coronal gas volumes.

- HII gas: In this phase, the hydrogen has been photoionized by ultraviolet photons from hot stars. Most of this gas keeps photoionized by the radiation from recently formed hot massive O-type stars. If the photoionized gas is dense material from a nearby cloud, this phase is called HII region. If it has lower densities, intercloud medium, it is referred as diffuse HII.
  Bright HII regions, such as the Orion Nebula, have dimensions of a few pc and their lifetimes are the same of their ionizing stars, around 3 to 10 Myr. When this gas is extended and has low densities, it is called warm ionized medium WIM, and contains much more total mass than the more visually notorious high density HII regions. The galaxy contains approx 1.1×10^{9} $M_{\odot}$ ionized hydrogen, and about half of it is contained within 500 pc of the disk midplane. This kind of photoionized gas is also found in planetary nebulae, formed from the rapid mass loss during the late stages of evolution of stars with initial mass 0.8 $M_{\odot}$<M<6 $M_{\odot}$ that expose their hot stellar cores, the radiation from this core photoionizes the outflowing gas, creating a luminous planetary nebula.

- Warm and cold neutral medium: The warm neutral medium is mostly atomic gas heated to temperatures around $10^{3.7}$. This gas is found at densities around 0.6 cm^{-3}. It fills a significant fraction of the volume of the disk, like 40% of it. The cold neutral medium has a temperature around 10^{2} K, with densities of $\approx$ 30 cm^{-3} and fills around 1% of the ISM. This gas is also referred to as the cold neutral medium, or CNM.

Gas phase	Temperature [K]	Density (cm^{-3})	Filling factor (%)	Total mass ($\times 10^9$ $M_\odot$)
Coronal gas	$10^{5.5}$-10^7	0.004	50	-
Warm ionized medium	10^4	0.3	25	1.0
Warm neutral medium	5000	0.6	30	2.8
Cold neutral medium	40-100	30	1.0	2.2
Dense molecular gas	6-50	10^3-10^6	0.05	1.3
HII regions	10^4	10^2-10^4	0.01	0.05

TABLE 1.1: Summary of the different gas phases of the ISM with their characteristics.

- Diffuse molecular gas: Despite the cool HI clouds, this gas has large enough density and column density to produce H_2 self-shielding, which allows this molecule to be more abundant in the cloud interior.

- Dense molecular gas: This kind of gas consist of gravitationally bound clouds that have achieved densities greater than $10^3 cm^{-3}$. This clouds have a visual extinction of $A_v \gtrsim 3$ mag through their central regions. In this dark clouds the dust grains are covered with mantles made of different molecular ices. In this region star formation takes place.

- Stellar outflows: This gas is produced because of cool stars losing mass at high rates, close to 10^{-4} $M_\odot$ yr^{-1} and low outflow velocities below 30 km s^{-1}, which leads to more or less high density outflows. With hot stars the winds are much faster, but less dense.

- Interstellar dust: In addition to the already mentioned gas phases of the ISM, there is also an important part of it in a solid state. This solid particles are present in the ISM, and in the circumstellar disks, protoplanetary and debris disks and evolved star envelopes. Are submicronic solid particles that are formed mainly from the last evolutionary phases of stars, and possibly in other energetic events. Due the different sources of dust grains, there is a vast diversity of dust grains materials. The main two type of dust grain composition are silicate grains and carbonaceous particles. This particles play an important role in the physics of dense interstellar gas. They can also contribute to the thermal balance as they efficiently cool through their thermal emission. They also act like chemical reactors, indeed one of the most important roles of dust grains is the production of complex organic molecules on grain mantles.

1.1.2 Energy sources

The energy balance of the ISM is regulated by contributions from the external medium, e.g. extragalactic sources and the Cosmic Microwave Background (CMB) radiation and also the internal energy sources, which determine the evolution of the ISM.

The interstellar radiation field (ISRF) is a frequency-dependent average of the intensity contribution from different objects into the ISM. The ISM is permeated by various photon fields, that influence the chemical and physical state of the gas and dust.
The stellar radiation field contains contributions from different sources, as shown in Fig. 1.1. At the highest wavelengths the main contribution is provided by the CMB, followed by dust grains, which absorb the stellar photons and re-radiate them at longer wavelengths. Moving towards lower wavelengths comes the Polycyclic Aromatic Hydrocarbons (PAHs), large organic molecules composed of hydrogen and carbon. They both emit in the infrared. Then, the contribution of stars is classified between late-type stars, important between far-infrared (FIR) to mid-infrared (MIR) wavelengths and early-type stars, which dominate the far-ultraviolet (FUV). A-type stars control the visible region of

the ISRF, and finally the hot coronal gas contributes at the extreme ultraviolet (EUV) and even at X-ray wavelengths (between 10^{-5}-10^{-7} cm).

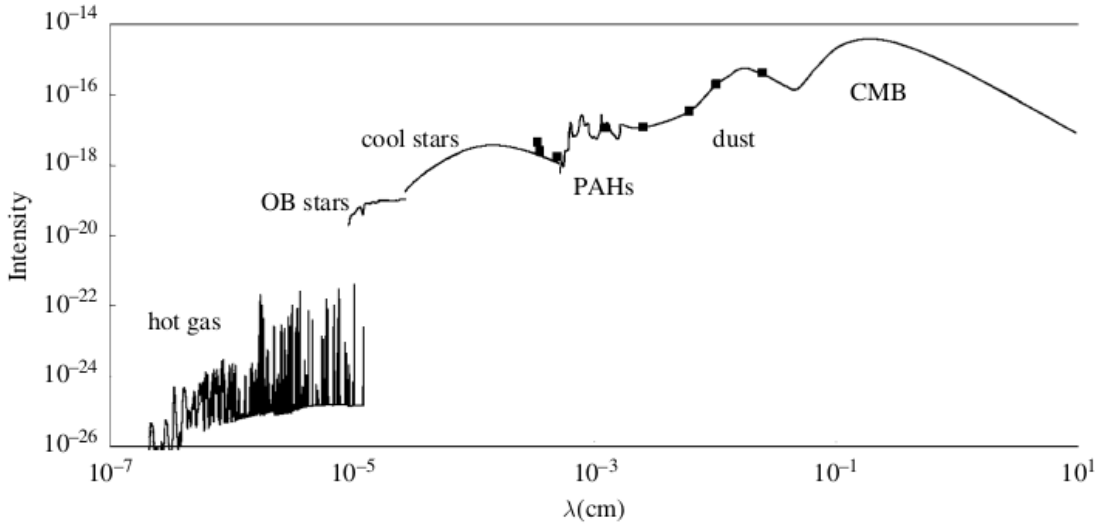

FIGURE 1.1: Spectral energy distribution of the interstellar radiation field. Figure taken from Tielens (2005).

Magnetic fields are an important energy and pressure source in the ISM, maintaining the energetic balance and controlling the dynamics of the gas. The magnetic fields strength increases inside dense clouds, having a typical strength of $\lesssim 10\mu$G.
Measuring the total magnetic field strength $|\mathbf{B}|$ and the column density N allows to determine the M/Φ (mass-to-magnetic flux ratio), a parameter that indicates the ability of a cloud to counteract the self-gravitational effect, meaning that the cloud can be classified as subcritical or supercritical by knowing this ratio. A cloud is considered with subcritical mass when the cloud mass is lower than $M_\Phi = \Phi/2\pi\sqrt{G}$, where M_Φ is the critical mass, Φ is the magnetic flux, and G is the gravitational constant. Here, the magnetic pressure is strong enough to contrast the gravitational effect, preventing gravitational collapse. When the clouds mass increases to the point that gets greater than M_Φ, then the cloud is called supercritical. At this point the cloud collapses and start with the star formation processes (Nakano and Nakamura, 1978).
The only available technique for directly measuring magnetic field strengths in interstellar clouds is measuring the Zeeman splitting observations. Fig. 1.2 shows $B_{\rm LOS}$ vs. $N_{\rm H}$ using data from Zeeman surveys of HI, OH, and CN. The blue dashed line represents a critical M/Φ, where M is the mass of the clouds and Φ is the flux passing through the cloud. On the left side of the blue line are located clouds with $N_{\rm H} \gtrsim 10^{21}$ cm^{-2}, mass-to-flux ratios M/Φ are subcritical. These clouds on the left side correspond to low density HI clouds. Meanwhile, on the right side of the blue line the M/Φ are supercritical clouds, with $N_{\rm H} \lesssim 10^{21}$ cm^{-2}. Also, the clouds on the left side are mainly the higher density molecular clouds and cores. Here, $B_{\rm LOS}$ is the magnetic field measured in the line of sight, and its defined as $B_{\rm LOS} = B_{\rm TOT} cos\theta \approx B_{\rm TOT}$, for $\theta \approx 0$. In Fig. 1.2 it is possible to see that there are no subcritical Molecular Clouds, which means that this figure does not show an evolutionary sequence driven by ambipolar diffusion. From the Crutcher (2012) there is no conclusive evidence for any ambipolar diffusion driven star formation.

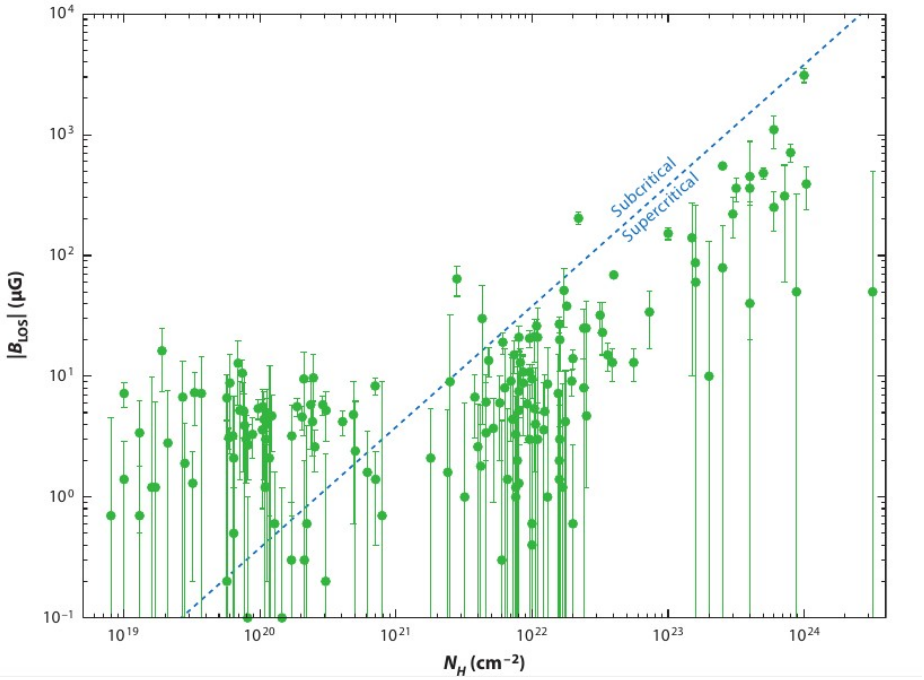

FIGURE 1.2: HI, OH and CN Zeeman measurements of B_{LOS} versus $N_H = N_{HI} + 2N_{H_2}$, where the dashed blue line represents the critical $M/\Phi = 3.8 \times 10^{-21} N_H / B$. Figure taken from Crutcher (2012).

Cosmic rays consist mainly of relativistic protons with energies between 1 and 10 GeV, 10% helium and heavier elements and electrons at about the 1% level. CRs are essential for the chemical reactions inside molecular clouds. The pressure due these CRs provides support against gravity for the gas in the ISM. When high-energy CR protons with 1-10 GeV reacts with interstellar gas gives raise to gamma-rays. Low-energy ($\simeq$ 100 MeV) CRs play a main role in heating and ionizing the interstellar gas, which activates the chemistry reactions inside the densest clouds.

1.2 Molecular clouds

Molecular clouds are regions where the gas is mainly molecular. They are composed by filaments, sub-filaments, cores and clumps, and since most of the stars are born inside molecular clouds, they are thought to be stellar nurseries. This is why their dynamic behavior and physical and chemical properties affect directly the formation of stars. They are characterized by having high levels of turbulence, which is indicated by the large line width of emission lines. These clouds are in self-gravitating equilibrium instead of pressure equilibrium with other phases of the ISM, and when there is a balance between the magnetic and turbulent pressure and gravity, and the molecular clouds are stable over time scales of $\simeq 3{\times}10^7$ years, these regions are forming stars. From observations of rotational transitions of a number of species it has been possible to study their physical and chemical properties.

The molecular gas can be part of three different types of clouds: The diffuse clouds, giant molecular clouds and dense cores. These clouds distinguish from each other because of their physical properties, like the density, mass and size. In table 1.2 we report a brief summary of the different cloud characteristics. Molecular clouds and their structure can be mapped using radio spectroscopy of molecular lines, continuum emission from dust or via stellar absorption by dust. The most common tracer of molecular gas is the CO J = 1-0 transition at 2.6 mm, and almost all the molecular clouds in

the Galaxy are detectable with CO (Williams et al., 2000). It is known that most of the molecular gas in the Milky Way is contained inside Giant Molecular Clouds (GMCs), with masses around 4×10^5 $M_\odot$, temperatures of 10 K and densities of $\simeq 200$ cm^{-3}. Giant molecular clouds are the most massive structures in the Galaxy, with masses even greater than 10^7 $M_\odot$. They are made of molecular gas with diameters of 5 to 200 parsec. Only ∼2% of the total gas from a GMC is converted in stars before the cloud is disrupted by the radiation pressure of the star clusters they produce, according to Evans (1991), but surprisingly ∼30% of the star formation in the Milky Way occurs in only 32 GMCs (Murray, 2011).

GMCs are located first in the hierarchical organization of molecular clouds, and are organized in the same way, following the same scaling laws. This means that the mass (M) of the molecular clouds varies as the square of their radius (R) and their internal velocity dispersion (σ) as the square root of their radius (Roman-Duval et al., 2010). For a system in virial equilibrium it is expected that the gravitational energy, that scales as GM^2/R, is balanced by the kinetic energy, that scales as $M\sigma^2$, which leads to a constancy near to a factor of $M\sigma^2/R$. The scaling laws are reminiscent of the scaling laws for turbulence, that are not really compressible yet. There is a relation between the scale size l and the local velocity dispersion at this scale δv as $l \sim \delta v^\alpha$, where α depends on the properties of turbulence and is expected to be 1/3 for incompressible turbulence, as predicted by Kolmogorov (1941).

Molecular clouds are observed as dark spots in the sky (as shown in Fig. 1.3), because of the effect

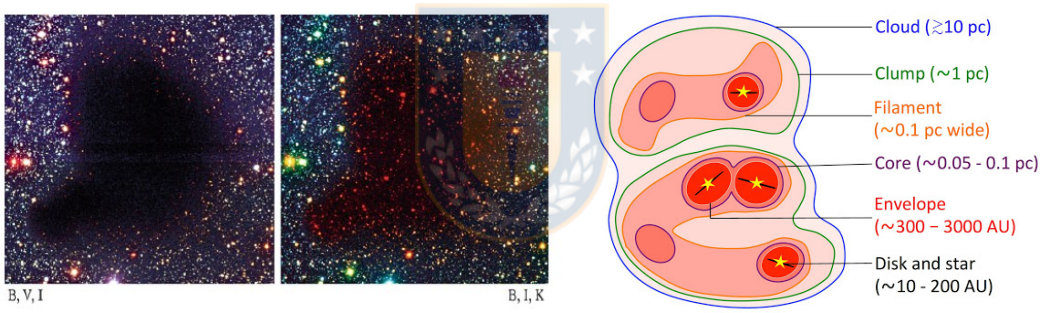

FIGURE 1.3: The first panel shows a picture of the Barnard 68 dark cloud at visible wavelengths. The second panel is the same cloud combined with near infrared. Both were taken from Smith et al. (2013). The third panel shows the hierarchical structure of molecular clouds, with their typical lengths. Figure taken from Williams et al. (2000).

of dust grains contained in them, that absorb and scatter the light of the nearby stars. Molecular clouds follow a hierarchical configuration, as can be seen in Fig. 1.3, in the third panel. They tend to have smaller subunits appearing inside larger ones, a categorization raised by Williams et al. (2000). In general, the nomenclature of molecular complexes and their substructure is chaotic, the general characteristics of every type of cloud is summarized in Table 1.2. Most of the authors adopt that the clouds are composed of gravitationally-bound clumps, that at the same time can form clusters or associations of stars, and those clumps are composed also of smaller cores, which are the places where single or multiple stars are born. Clumps are generally identified from spectral line maps of molecular emission, and star-forming clumps are the massive clumps out of which stellar clusters form. Most of the clusters are unbound from each other, but the gas from where they were formed is bounded. On the other side, cores are the regions where stars are formed, and are necessarily gravitationally bound (Williams et al., 2000).

Property	Clouds	Clumps	Cores
Mass ($M_\odot$)	10^3-10^4	50-500	0.5-5
Size (pc)	2-15	0.3-3	0.03-2
Density cm^{-3}	50-500	10^3-10^4	10^4-10^5
Velocity extent (km s^{-1})	2-5	0.3-3	0.1-0.3
Crossing time (Myr)	2-4	1	0.5-1
Gas Temperature (K)	10	10-20	8-12

TABLE 1.2: Summary of the properties on different types of clouds, adapted from Tielens (2005).

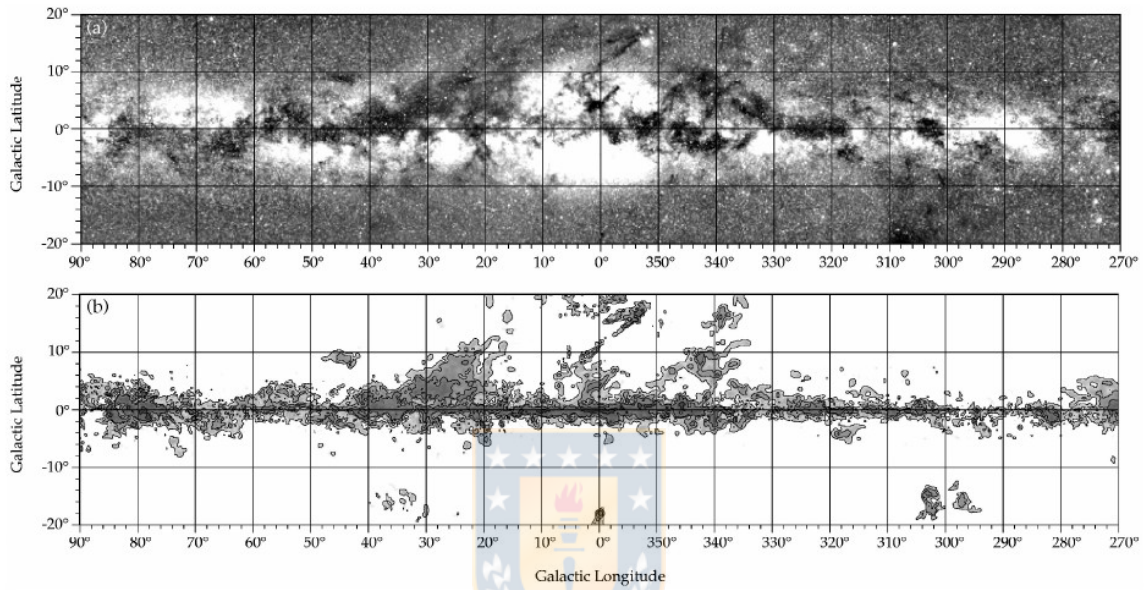

FIGURE 1.4: (a) A Milky Way optical map that covers the first and fourth Galactic quadrants, where the most intense regions toward the Galactic bulge have been saturated to better display the dark clouds. (b) This is a spatial map of molecular clouds with near kinematic distances less than 2.5 kpc according to the rotation curve performed by Brand and Blitz (1993), obtained by integrating over a longitude-dependent velocity window. The window always included the range v = -10 to 10 km s^{-1} in order to include the random motions of local clouds, but changed in the center portions as high as $\sim$ 35 km s^{-1}. The two panels were taken from Dame et al. (2001).

1.2.1 Infrared dark clouds

Infrared dark clouds (IRDCs) are far, rare, and dense regions observed in the mid-infrared (MIR), where massive star formation is thought to occur. These type of clouds are promising hunting grounds for the initial stages of high-mass star formation. The temperatures of these objects are usually below 20 K and their densities are around 10^5 - 10^7 cm^{-3} (Carey et al., 1998; Teyssier et al., 2002). Typical clump masses are around 100-1000 $M_\odot$, the sizes range from 0.3 to 0.7 pc (Sridharan et al., 2005; Ragan et al., 2006; Rathborne et al., 2006) and even when they have different shapes and sizes, they tend to have irregular edges and their appearance is filamentary and wind-blown. An IRDC has often two or three filaments that are considered long compared with its total length, and they are either parallel or converge at a low angle in a massive condensation that can contain an active cluster-forming site (Tachihara et al., 2002; Burkert and Hartmann, 2004). IRDCs are seem to be born with a filamentary distribution of material that extends over a large percent of their whole size. As this clouds evolve and begin with star formation, the processes that occur over time continue with the filamentary structure of the parental gas (Hartmann, 2002).

1.3 Star formation

The equilibrium state of dark clouds and their collapse under gravity in the process to form stars is very controversial because of their uncertainties. From studies related to the magnetic fields and the turbulent motions two different scenarios of the global state of the clouds have been proposed. The first one states that clouds are close to equilibrium and that their evolution towards star formation is approximately quasi-static, and the other one defends that clouds are dynamic objects that evolve and form stars within a crossing time. The quasi-static scenario stands that clouds are close to equilibrium because of their relatively long lives ($\sim$ 10 Myr), they are close to be virialized. Here equilibrium against self-gravity is provided by the magnetic field which has a double contribution toward stability.

In the ISM through some hierarchical fragmentation processes, it produces a spanning between many orders of magnitude in mass scale, and it occurs preferentially in objects with masses from 10^1 to $10^2 M_\odot$ (Shu et al., 1987). One of the first studies that makes a difference between low-mass and high-mass star formation classified molecular clouds into a cold population ($<$10 K) and a warm one (clouds with cores $>$ 20 K). Inside cold clouds the most late in spectral type of star that can be found are the late B stars.

Warm clouds tend to be among the largest and most massive, are associated with radio HII regions and appear to be a spiral arm population.

Most of the stars in the Galaxy are found in binary or multiple systems according to Batten (1973). Cloud cores start losing their magnetic support and turbulence over time, becoming gravitational collapse.

Low and high-mass stars share the very processes in the early stages of their formation, but at certain point of evolution their formation take different paths according to the amount of available mass. Nevertheless, none of these parts can be completely understood without the study of the other, since they are both connected by a number of competing feedback loops (Klessen and Glover, 2016). The spatial distribution of stars in the forming cluster shows a dependence on stellar mass. As more massive stars are located closer to the central part of the group, the stars with lower masses tend to be located at the outer regions of the clusters (Girichidis et al., 2012).

1.3.1 Stellar initial mass function

The stellar initial mass function (IMF) is the distribution of masses with which stars are formed, and has a great importance to understand how stars evolve with time. It is also useful for the understanding of the galaxy formation process, in early stages. In the past, evidence suggested that the IMF should vary according to the pressure and temperature of the star forming cloud. The galaxy luminosity depends meanly on the number of stars with masses near to 1 $M_\odot$, while the heavily-element contents and the energy feedback effects that control the evolution of their gas contents depend on the number of stars that formed in them with masses greater than 10 $M_\odot$. Considering that the low and high-mass stars number is not the same at early stages of the galaxies, all the properties of galaxies as luminosity, color or chemical compositions and the evolution with time would be altered over time. The strongest evidence for variability of the IMF according to several authors is the large total mass of heavy elements observed in cluster of galaxies. Even with the existence of arguments suggesting that IMF should vary with the pressure and temperature of the star forming cloud, with higher temperature regions expected to produce higher average stellar masses (Larson, 1998), there is still no strong evidence confirming a variable IMF. The IMF inferred

from Galactic-field star-counts can be conveniently described by this multiple-part power-law

$$\xi(m) \propto \mathrm{m}^{-\alpha_i} = \mathrm{m}^{\gamma_i} \tag{1.1}$$

Where

$$\begin{aligned}
\alpha_0 &= +0.3 \pm 0.7, 0.01 \leq \mathrm{m}/\mathrm{M}_\odot \leq 0.08, \\
\alpha_1 &= +1.3 \pm 0.5, 0.08 \leq \mathrm{m}/\mathrm{M}_\odot \leq 0.50, \\
\alpha_2 &= +2.3 \pm 0.3, 0.50 \leq \mathrm{m}/\mathrm{M}_\odot \leq 1.00, \\
\alpha_3 &= +2.3 \pm 0.7, 1.00 \leq \mathrm{m}/\mathrm{M}_\odot,
\end{aligned}$$

and ξ(m)dm is the number of single stars in the mass interval m to m+dm. Since the Galactic field is populated by many different star formation events, there is a big diversity of objects in the solar neighborhood, and together with the nowadays star formation, this leads to variations of the Galactic-field IMF. There is a strong difference between the low density environments, compared to high-density regions (Kroupa, 2001). When a star-forming clump reaches a critical density, it starts to interact with other clumps until their collapse is complete. They also proposed that the dynamical interactions (collisions) of the accreting protostar with other young stellar objects is what determine the end of the main accreting phase. If the characteristic interaction time is $\mathrm{t}_{\mathrm{int}}$, then the final stellar mass will be determined by the condition $t_{\mathrm{acc}} \sim t_{\mathrm{int}}$ or $\mathrm{M} \sim \mathrm{M}t_{\mathrm{int}}$. Considering this, stars with masses much larger than $\mathrm{M}t_{int}$ will not be able to form, introducing a cut-off in the mass distribution directly depending on the star formation environment (Price and Podsiadlowski, 1995). Many studies have compared low-density environments, e.g. Taurus-Auriga and Oph with high-density ones e.g., Orion nebula cluster, but no conclusive difference has been found between each of their IMFs. At this point most authors are accepting the power-law distribution described by Kroupa (2001) as

$$\frac{\mathrm{dN}}{\mathrm{dM}} \sim \mathrm{M}^{-\eta} \tag{1.2}$$

where N is the number of stars and M is the total mass of the object and

$$\begin{aligned}
\eta &= 0.3, \mathrm{m} < 0.08, \\
\eta &= 1.3, 0.08 < \mathrm{m} < 0.5, \\
\eta &= 2.3, \mathrm{m} > 0.5
\end{aligned}$$

this power-law distribution is shown in Fig.1.5. Bergin and Tafalla (2007) suggest that the observable similarity between the mass distribution of cores and the stellar IMF are determined by the same process that fragments the molecular gas at smaller scales.

1.3.2 Low-mass star formation

Low-mass stars (with $\mathrm{M}_\odot < 8\ \mathrm{M}_\odot$) formation follows six steps summarized in Fig. 1.6, explained in detail by Klessen and Glover (2016). It begins with the formation of molecular cloud complexes in the turbulent phases of the ISM, followed by (see panel a) the supersonic turbulence generating high density clumps with a wide range of densities and sizes. Some of the said density fluctuations will become gravitationally unstable (see panel b), leading to a collapse. When this process starts, the central density increases until the compressional heat generated by the contraction becomes difficult to be completely irradiated. After that, a quasi-hydrostatic object begins to form at the central part

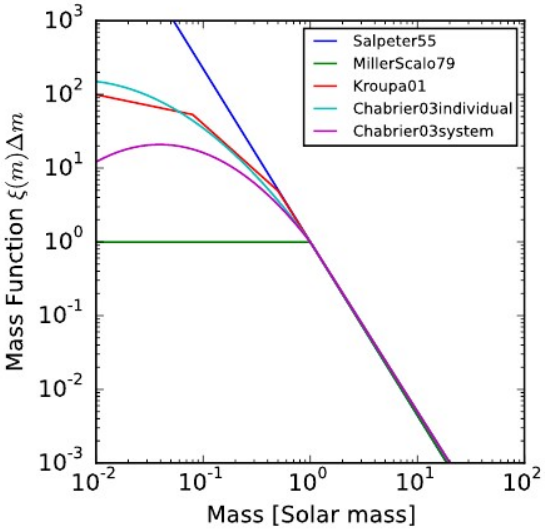

FIGURE 1.5: Different proposed iterations for the shape of the IMF. In this figure the red line is the power-law, proposed by Kroupa (2001).

of the core, leading to a protostar. This starts to get more massive over time, accreting more material from the infalling envelope. When this envelope is still in the Class 0 phase of the protostellar evolution, it is called *enclosing envelope*. The envelope, at this stage is optically thick and absorbs the accretion luminosity generated as the infalling material stops at the surface of the protostar. At this point, the spectral energy distribution (SED) of the system is dominated by the reprocessed emission from the envelope, that radiates mostly at sub-mm wavelengths.

Most of the infalling material will not be accreted by the protostar, due to angular momentum (see panel c), so it ends up building an accretion disk around the protostar. The mass of such disk can also reach a critic mass, at which stops its capacity to incorporate material with gravitational processes. With this, the disk starts to fragment into a multiple stellar systems. The magnetic field of the molecular clouds is compressed and amplified by dynamo processes during the contraction, leading the accretion disk to launch magnetically driven outflows along the rotational axis of the system. These outflows begin to disperse the envelope material, in a higher rate over time, causing a visible effect on the observed SED. As more of the inner regions of the disk become visible, the emission peak also moves closer to the infrared wavelengths. Next, in the panel (d) the protostar starts cleaning the envelope and its evolution from class 1 to class 2. Here its showed that the protostar is in the T Tauri phase of its evolution, where most of the mass is assembled in the central star with the remaining accretion disk contributing with only a few percent to the overall matter budget, and its at this point where planet formation begins. In panel (e) it is showed how the envelope is completely removed or accreted into the disk, allowing the central part of the star to be visible. The planet formation takes places in the residuals of the disk. The protostar at this point is fully convective, and all the energy it loses in emission of radiation at its surface is compensated releasing gravitational energy. The protostar will slowly contract, becoming hotter and denser over time. Now the conditions of the central part of the protostar are optimal to begin with the nuclear fusion. At this point, the stars enters the main sequence and settles to a quasi-equilibrium state. At the final phase (panel f) the remaining disk is cleared away and the central star is clear and likely surrounded by a planetary system. The successive evolution of the protostar can be classified into

four distinct phases, based on the spectral index at IR wavelengths

$$\alpha_{K,N} = \frac{log(\lambda F_\lambda)_{10\mu m} - log(\lambda F_\lambda)_{2.2\mu m}}{log(10/2.2)} \tag{1.3}$$

Where K the 2.2 μm K -band, and N the N-band. From this, the evolution sequence follows as:

- **Class 0:** These objects are highly obscured protostars, with a blackbody spectrum peaking at 100 μm.
- **Class I:** These correspond to blackbody objects with T$\lesssim$850 K. They exhibit an intense component at $\sim$ 10 μm emitted by its protostellar disk and a weaker emission around 2 μm. The spectral index in these objects is positive.
- **Class II:** For this objects the spectral index is -1.5 $< \alpha_{K,N} <$ 0. The protostar is blackbody with 870 $<$ T $<$ 1540 K, and a peak at $\sim$ 2 μm and a less intense disk component that increases the flux at $\lambda \gtrsim 2\mu$m. This kind of stars are also known as T-Tauri stars, and are located in the pre-main sequence.
- **Class III:** These last objects are blackbodies with T $>$ 1540 K, where the accretion disk component of the flux is extremely weak, sometimes even completely absent. This class of objects are close to their final mass.

All the steps previously described are valid for a isolated low-mass star, the situation for high-mass stars follows a different process that will be explained in the next section.

1.3.3 Massive star formation

The formation of massive stars is a complex process, still not well understood, and there is no observational evolutionary sequence that is firmly established, mainly because timescales are much faster in comparison with the formation of low-mass stars. In addition they are formed in distant clouds within stellar clusters, making this process a highly complex matter of study.
High-mass star forming regions have more extreme conditions compared to the low-mass case. Their typical sizes are $\sim$ 1 pc, their mean densities are $n \approx 10^5 \mathrm{cm}^{-3}$, masses from $\sim 10^3$ $M_\odot$ and greater, and velocity dispersions between 1.5 and 4 km s^{-1}. It is known that massive stars are formed by multiple stellar systems (Ho and Haschick, 1981; Lada, 2006), that at the same time are part of bigger clusters, according to Lada and Lada (2003); de Wit et al. (2004) and other authors. Because of this, it is much difficult to differentiate the dynamic interactions from the influence of individual stars with observational data. The Kelvin-Helmholtz pre-main sequence contraction of massive stars is considerably shorter compared to their accretion time. At the moment when the stars reach a mass of 10 $M_\odot$, they begin to ionize the environment, while the spectrum starts to be UV-dominated. In spite of this, different theoretical models have been proposed to better understand the high-mass star formation process. The two models are the turbulent core accretion, also known as monolithic collapse model, and the competitive accretion model. In the turbulent core model, that is an extrapolation of low-mass star formation to the higher mass regime, prestellar cores on the verge of their collapse should be static and isolated from their surroundings. As soon as the cores loose their supplementary turbulent and magnetic support, high-mass prestellar cores start to collapse and enter the protostellar phase. It relies on the apparent similarity of the core mass function and the stellar IMF. In this model the turbulent energy in very massive cores is large enough to cause them to fragment in multiple objects. Only density profiles which are strongly concentrated

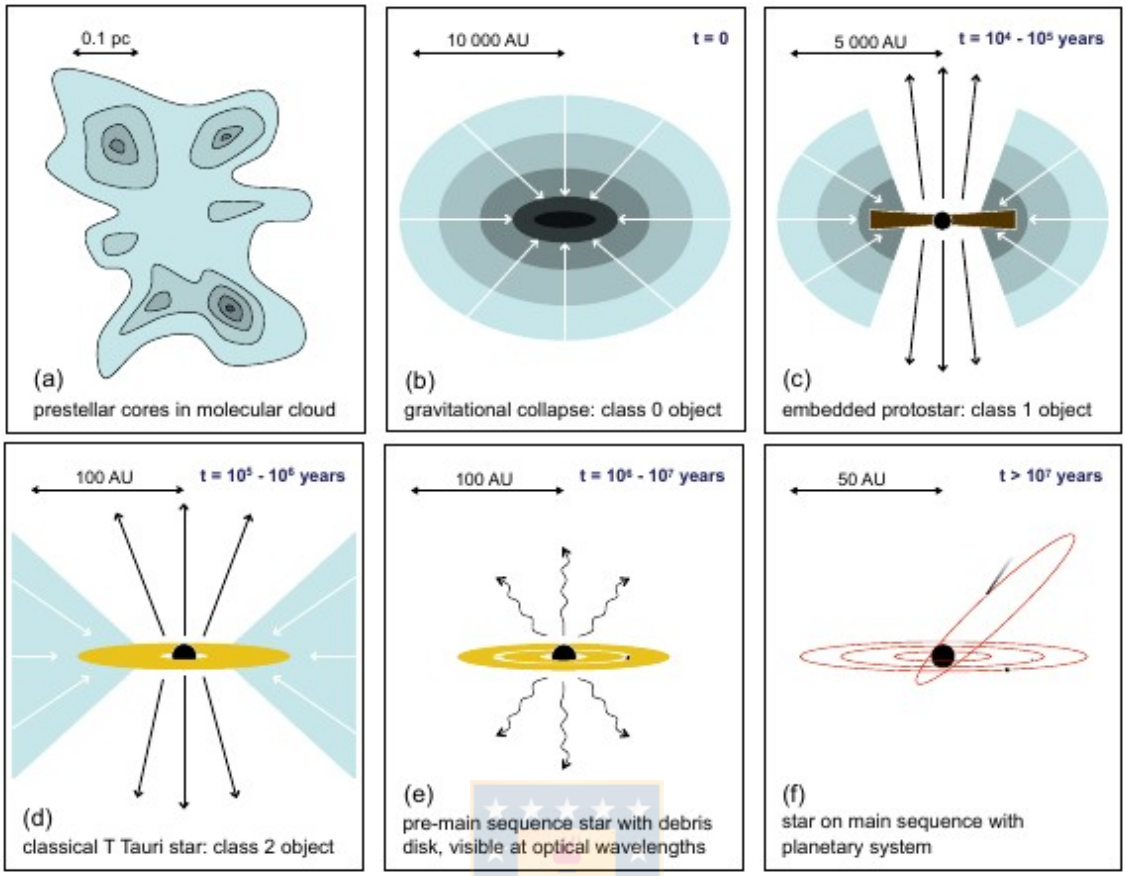

FIGURE 1.6: Low-mass star formation process summarized by Klessen and Glover (2016). In the panel (a) the prestellar core form due turbulent compression inside larger molecular clouds.. (b) It becomes gravitationally unstable and starts to collapse. At this phase the protostar gains mass rapidly, and corresponds to a class 0 stage. (c) The infalling material accretes to the protostar, conserving the angular momentum and starting with this the protoplanetary disk. The outflows, magnetically launched, disperse the infalling envelope. The protostar is now a class 1. (d) The protostar now becomes visible as the envelope keeps being removed. Now it is in the class 2 phase. (e) Once the envelope is completely removed, there is only a remnant disk, where planet formation takes place. The young star is on the Kevin-Helmholtz pre-main sequence contraction phase. (f) In the final phase all the gas and dust, leaving the star with a planetary system alone.

and weak initial turbulent motions can prevent this fragmentation, meanwhile the amount of radiation feedback can reduce the amount of fragmentation, but not prevent the core from fragmenting into multiple protostars (Girichidis et al., 2012).

The competitive accretion model proposes that the collapse starts within low-mass fragments, close to the opacity limit that start to gain and accrete mass, building up a greater IMF. This accretion rate is determined by the position of the forming star in the cluster. The gas is concentrated at the center of the cluster due to the gravitational potential, which leads to having greater accretion rates in the objects located at the central region. This central protostars can grow to be the most massive stars. Also, the more massive is a star, the faster it migrates to the center of mass of the core, being located continuously at a gravitationally privileged position. The most massive stars also continue to accrete at higher rates. In the competitive accretion model, there is a relation proposed between the total stellar mass and the most massive stars, as $M_{mm} \propto M_{tot}^{2/3}$ (Bonnell et al., 2004). This mass to mass relation was proposed as an indicator for this accretion model, by Krumholz and Bonnell

(2009), together with the fragmentation induced starvation, although it is also known that it does not represent an unique signpost of competitive accretion. In this scenario high-mass protostars would be fed from the gas of their surrounding (Motte et al., 2018).
The main difference between the two presented models is the speed of collapse, which is highly determined by the initial conditions of the cloud. It is important to mention that there are a number of different scenarios between this two most important already discussed, for example in Peretto et al. (2006) via observations of the NGC 2264 clumps, they support the idea of a global dynamical collapse cluster of protoclusters, and that most of the stars in a cluster acquire their final masses while they are still inside a gas-dominated environment. Here they propose a mixed scenario of protocluster formation where a massive turbulent clump is compressed and induced to collapse from the outside (Girichidis et al., 2011).
In order to clarify some of the uncertainties about the formation of the high-mass stars process, Motte et al. (2018) proposed a new evolutionary scenario, constrained by observations, that recalls the global hierarchical collapse and clump-fed accretion scenarios (Vázquez-Semadeni et al., 2009), confirming that high-mass stars are not formed in a scaled-up version of the low-mass star formation process. They proposed a third scenario, summarized in Fig. 1.7. This new scenario follow the next steps

1. In the first phase high-mass stars form in molecular complexes that host massive clouds and often OB clusters. Massive clumps/clouds called ridges and hubs are the most likely or maybe the only ones harvesting high-mass star formation.

2. At this point, IR-quiet massive dense cores (MDCs) are 0.1 pc massive cloud fragments that host low-mass prestellar cores. This phase is the starless MDC that prolongs for around one free-fall time ($\sim 10^5$ yr).

3. Inside the MDC, low-mass prestellar cores become protostars with growing mass and stars that do not have enough mass to be considered high-mass ones. The global collapse of ridges or hubs generates gas flow streams, that cause an important increase in the mass of the MDCs.

4. At this point the inflowing gas streams are efficient enough to reach the low-mass protostellar cores, and help to feed them, which with time leads them to become IR-quiet high-mass protostars. They are only about 0.02 pc and are capable of harbor low-mass under 8 $M_\odot$ stellar embryos. Also, the accretion rate is very high and can drive outflows and power hot cores.

5. Reaching the 8 $M_\odot$, the protostar luminosity increases, becoming an IR-bright object. In this phase the hot cores of the protostar will soon start producing hyper-compact HII regions (HCHII) (Hoare et al., 2007), that would only be stopped by the infalling gas or localized toward photo-evaporating disks.

6. Finally, the stellar embryos count with a substantial and increasing UV field, that leads to a H_{II} region. This, together with the winds and outflows will stop gas accretion toward the newborn star.

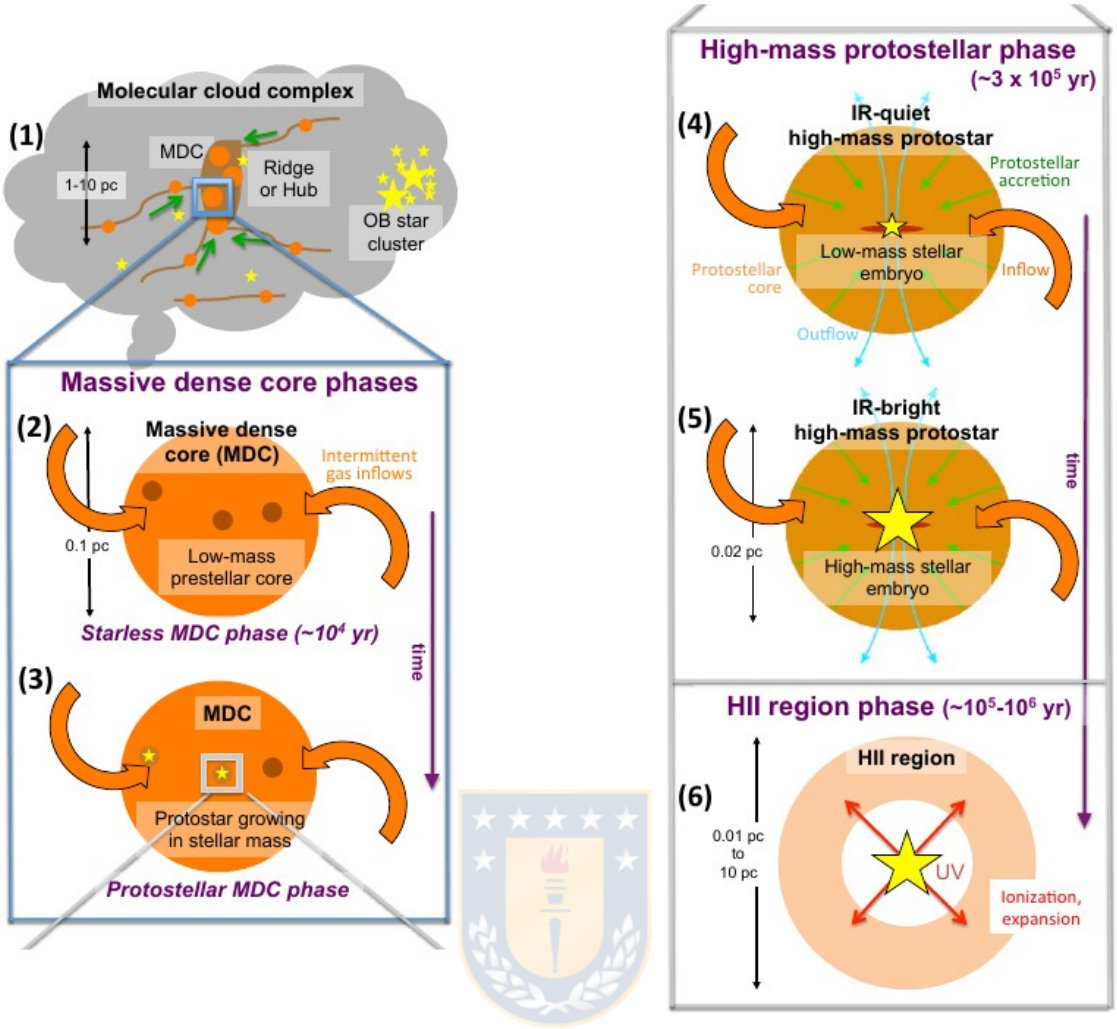

FIGURE 1.7: Schematic evolutionary diagram proposed for the formation of high-mass stars, taking from Motte et al. (2018). (1) It is shown the formation of filaments and spherical clumps, called ridged and hubs, host massive dense cores, starting with the formation of high-mass stars. (2) At this starless phase, MDCs can only harbor low-mass prestellar cores. (3) The IR-quiet MDCs evolve intro protostellar objects when they are able to host a low-mass stellar embryo. Also, at this point ridges and hubs are collapsing. (4) Now the protostellar envelope starts to feed from inflows, starting the process of the high-mass star formation. (5) High-mass protostars transform into IR-bright starting from stellar embryos with mass larger that 8 $M_{\odot}$. The accretion finishes when the stellar UV field is ionizing the protostellar envelope and an HII region is formed.

1.3.4 The APEX Telescope Large Area Survey of the Galaxy (ATLASGAL)

The APEX Telescope Large Area Survey of the Galaxy (ATLASGAL)[1] is the first complete survey at 870 μm of the inner Galactic plane, providing the most complete inventory that includes all massive dense clumps of the Galaxy, with masses $\geq$ 1000 $M_{\odot}$, at a heliocentric distance of 20 kpc. The regions covered by ATLASGAL are shown in Fig. 1.8.

The Atacama Pathfinder EXperiment (APEX) [2] telescope is a single-dish 12-m sub-mm reflector, with a 0.75-m secondary reflector, located in the Chajnantor plateau in Chile, at 5.000 m above the sea level, being the ideal place to observe the inner Milky Way. It is a Cassegrain type telescope, with

[1] https://www3.mpifr-bonn.mpg.de/div/atlasgal/
[2] http://www.apex-telescope.org/ns/

an alt-azimuthal mounting, specially designed to work at millimeter and sub-mm wavelengths from 0.2 mm to 1.5 mm. APEX was designed as a prototype antenna for ALMA. [3]
One of the APEX instruments is the Large APEX BOlometer CAmera, also called LABOCA, a multi-channel bolometer array for observations in the continuum at 345 GHz, i.e. 870 μm. It consists of 9 hexagons that together add up 295 pixels. The HPBW angular resolution at this specific frequency is of 18".6. The total Field of View (FoV), that corrresponds to the region of the sky that can be measured by the antenna, in the case of the APEX telescope has a total FoV of 11'.4.
Observations and studies focused on the spiral arms dynamics have proved that there is a direct correlation between the presence of molecular gas and star formation, and that also trace an spiral pattern, as seen in Colombo et al. (2014). Moore et al. (2012); Urquhart et al. (2014) performed an analysis of the molecular gas distribution, showing peaks at certain specific Galactocentric radii. From these studies it is confirmed that there is a strong relation between spiral arms and the star formation process, but still is not specifically clear how these two are connected. The purpose of ATLASGAL Survey is to provide the largest inventory of dense molecular clumps, helping through observational the data to answer some of the questions about massive star formation.
There is a total of $\sim$ 8000 dense clumps in the ATLASGAL sample, including their different properties, i.e., velocities, distances, masses, luminosities and spatial distribution. In the next paragraphs we will give a full revision of how all these parameters were obtained.

- Spectral energy distribution: Calculating the spectral energy distribution (SED) for every source is the first step for obtaining other parameters like the bolometric luminosity and temperature for the sources. The method used to reconstruct the SED was extracting the multi-wavelength continuum data (obtained from mid-IR and sub-mm wavelengths maps) and then perform aperture photometry to the sources. These obtained fluxes were fitted with a simple model and from this it is possible to derive the dust temperature and integrated flux. From these two parameters it is possible to obtain the masses and bolometric luminosities.
  For the sources in which there is no MIR emission detected, or where the flux density is available only for one MIR band. The method consists of fitting a single graybody model to the flux densities measured for the cold dust envelopes. In the case where two flux measurements are available at the shorter and different wavelengths a two-component, made of a blackbody and a graybody fitted to the SED. All the fitting procedure is explained in detail by König et al. (2017) and references therein. For 38 sources the of the MIR dark sources was used the single graybody component fit and 73 sources were fitted with two-component model. The uncertainties considered for the SED is about 42 percent.

- Heliocentric distance: The determination of the heliocentric distances of the clumps starts with the radial velocities associated for each clump. Once obtained the radial velocities, the distance determination is approached from a maser parallax and spectroscopic distances. H_I data can be used to make a more accurate approach of the distance.
  The distance distribution of the clumps shows that a great part of them are located no further away than 5 kpc, with the highest density of them between 2 and 4 kpc. with an uncertainty of $\sim$6 percent, that corresponds to the mean uncertainty, at 5 kpc which is the mean distance of the total clumps.
  The kinematic distance of around 7000 sources of the ATLASGAL survey was determined, from which $\sim$ 80% are considered as reliable.

[3] https://www.almaobservatory.org

- Dust temperature: This parameter was derived from the greybody and blackbody fit previously described. The range of the clumps dust temperatures is from 10 to 40 K, with most of the clumps near to ∼16 K. There is also a tendency of as more distant the clumps are, they have slightly higher temperatures. The mean uncertainty associated to the dust temperature measurement is ∼ 10 percent.

- Sizes: The clump sizes were calculated using the effective angular radii and the distances already obtained. A great part of the sample consist in unresolved sources, so the amount of clump sizes successfully determined was about 60% of the total. The mean size of the resolved clumps is 0.72±0.01 pc. The estimated error for the sizes of the clumps is ∼6 percent.

- Mass: To obtain the mass clumps the first step is to estimate the isothermal clump masses, assuming that the total mass of the clump is proportional to the integrated sub-mm flux density measured over the source and assuming that the emission is optically thin. The definition of flux is given by

$$S_\nu = \int_{\Omega_s} I_\nu(\theta,\phi) cos\theta d\Omega \tag{1.4}$$

  where I_ν is the *specific intensity* or *brightness* and Ω_s is the solid angle subtended by the source. In the Appendix A there is more detailed information about basic radioastronomy concepts. Based on the flux, the mass of the clump can be obtained from

$$M_{\rm clump} = \frac{D^2 S_\nu \gamma}{B_\nu(T_{\rm dust})\kappa_\nu} \tag{1.5}$$

  where S_ν is the integrated 870 μm flux density, D is the heliocentric distance, γ is the gas-to-dust ratio, here assumed ($\gamma = 100$), B_ν is the Planck function for a given dust temperature T_{dust} (and then used the previously obtained dust temperatures and assumed to be an estimation of the whole clump temperature) and κ_ν is the dust absorption coefficient, with a value of 1.85 $\rm cm^2 g^{-1}$, according to Schuller et al. (2009).
  The clump masses vary from a few solar masses to 10^4 $M_\odot$, where most of the clumps have masses close to the 400 $M_\odot$. The estimated uncertainty for the mass is 20 percent.

- Column density: The H_2 column density were obtained using the following equation

$$N_{\rm H_2} = \frac{S_{\nu,peak}\gamma}{B_\nu(T_{\rm dust})\Omega\kappa_\nu\mu m_{\rm H}} \tag{1.6}$$

  with $S_{\nu,peak}$ is the peak flux of the source, Ω is the solid angle, μ is the mean molecular weight of the gas (assumed to be equal to 2.8), m_H is the mass of the hydrogen atom and κ_ν and γ are previously defined. The column densities uncertainties is, similarly to the mass uncertainties, around 20 percent.

- Bolometric luminosity: For the estimation of the bolometric luminosities of the sources, it was used the following equation:

$$L_{\rm bol} = 4\pi d^2 \int S_\lambda d\lambda \tag{1.7}$$

  where d is the distance to the source and $\int S_\lambda d\lambda$ the integrated flux. The bolometric luminosities in the sample fluctuate from 57 $L_\odot$ to 3.8 $\times 10^6$ $L_\odot$. The uncertainties associated with the luminosity approach is around 42 percent.

Most of the clumps studied in the sample (∼ 66%) are actively forming stars, corresponding to 5275 clumps. About 16% of these clumps are associated with massive star formation (MSF) and the

remaining $\sim$ 52% corresponds to YSOs. $\sim$21% of the sample (1640 clumps) correspond to protostellar clumps. The last $\sim$12% (946 clumps) of the sample correspond to the quiescent class. Thus, it is inferred that around 88% of the ATLASGAL clumps are involved in star formation processes.
Several trends were identified in the variation of the physical parameters of the ATLASGAL survey, and a lot of them can be related to the different evolutionary stage of each clump. There is a tendency for the clumps that are more evolved, associated with massive star formation, to have bigger radius and greater densities compared with the younger clumps, not yet related with star formation processes. The mass remains more or less constant through the different evolutionary stages of the clump, and it is possible that the clumps can gain mass from infall material, which can be estimated from the surrounding lower density envelope. Also a clump can gain material because of the infall, and this can be estimated. The youngest clumps are the ones with the highest surface densities of all, and it decreases as the clump evolves. Finally, there is a noticeable trend on the temperatures and luminosities increasing as the evolutionary class advances.
From the ATLASGAL compact source catalog, a subsample of 110 sources with different characteristics was selected with the goal of constitute a representative sample of all diversity of clumps. This subsample is called the "TOP100" subsample from ATLASGAL, that will be described with more details in the following subsection.

TOP100 sample: an observational evolutionary sequence

This subsample consists in a total of 110 selected sources from the ATLASGAL compact source catalog, with diverse evolutionary stages.
For the evolution of the TOP100 subsample of ATLASGAL clumps, König et al. (2017) defined a classification for the different clumps according to their evolutionary stage. All the evolutionary stages are represented in Fig. 1.9. The classification is the following:

- Starless/pre-stellar stage: Represents the earliest stage of massive-star formation, a quiescent phase. These clumps tend to be the coldest and least luminous sources of the sample. This class is called "70 μm weak" or simply "Quiescent".

- Protostellar stage: The clumps are seen as compact point sources with protostellar object present. The embedded 70μm sources are either not associated with any mid-infrared counterparts within 10" of the peak emission or the associated compact emission is below the threshold of 2.6 Jy, which is an indicator that star formation is at an early stage and the clumps are likely to be dominated by cold gas. This class is known as "mid-IR weak" and in this work it will be referred as "protostellar".

- High-mass protostellar stage: In this stage there is strong compact mid-infrared emission, it is one of the most active stages of massive-star formation. These clumps are likely undergoing collapse in the absence of a strong magnetic field (Urquhart et al., 2014, 2015), shows signs of infall (Wyrowski et al., 2016) and are likely to drive strong outflows. This class of clumps tend to be significantly hotter than the sources in the quiescent phase and are called "mid-IR bright" or, as in this work "Young Stellar Object" (YSO).

- Compact HII region stage: The last evolutionary phase of massive star formation. Here, the sources begun to disperse their natal envelope and are ionizing their local environment, creating then, compact HII regions. They are called "HII regions" or, as we will call them from now on "Massive Star Forming" regions (MSF).

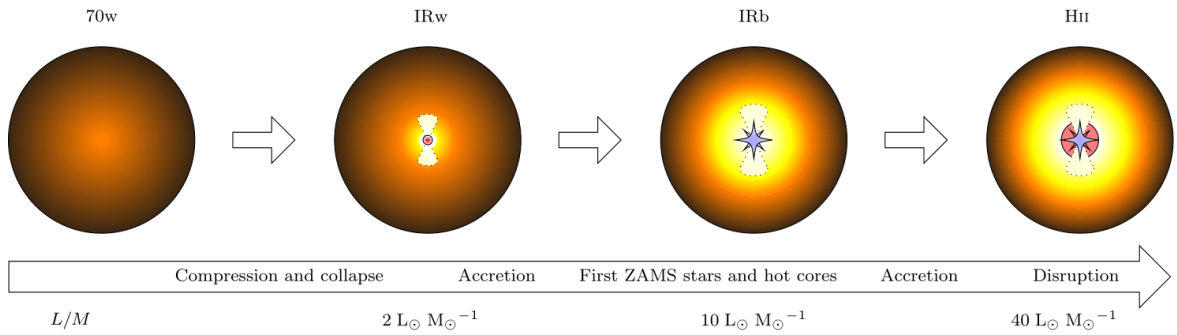

FIGURE 1.9: Representation of the different phases identified for the TOP100 sources, taken from Giannetti et al. (2017)

The TOP100 subsample of ATLASGAL was classified following the same schematic representation shown in Fig. 1.9. Using the Coordinated Radio and Infrared Survey for High-Mass Star Formation Survey (CORNISH) (Hoare et al., 2012; Purcell et al., 2013) and Red MSX Source Survey (RMS) (Urquhart et al., 2007, 2009) for radio continuum emissions and also with observations towards methanol masers, to classify all the ATLASGAL TOP100 sources. For the clumps with radio continuum emissions at 4 or 8 GHz within 10", then the source is classified as a compact HII region. In the TOP100 subsample there were 25 classified as HII. For the mid-infrared weak and bright, observations at 21μm and 24μm were used. A source is classified as a mid-infrared bright when there is a compact mid-infrared source associated with the sub-mm emission peak and the flux reported in the compact source catalogs above 2.6 Jy. A total of 36 sources were classified as mid-IR bright. On the other hand, a source is classified as mid-infrared weak when it presents emissions in the 21 μm and 24 μm below 2.6 Jy or there is no compact source related with this peak. A total of 33 sources were found to be in this evolutionary stage. The sources classified as prestellar were identified from a visual inspection at 70 μm. From the TOP100 subsample only 16 clumps were in the starless-prestellar phase.

Fig. 1.10 shows one per each of the four different evolutionary classes. Some of the trends observed in the full ATLASGAL sample study were also observed in the TOP100 sample.

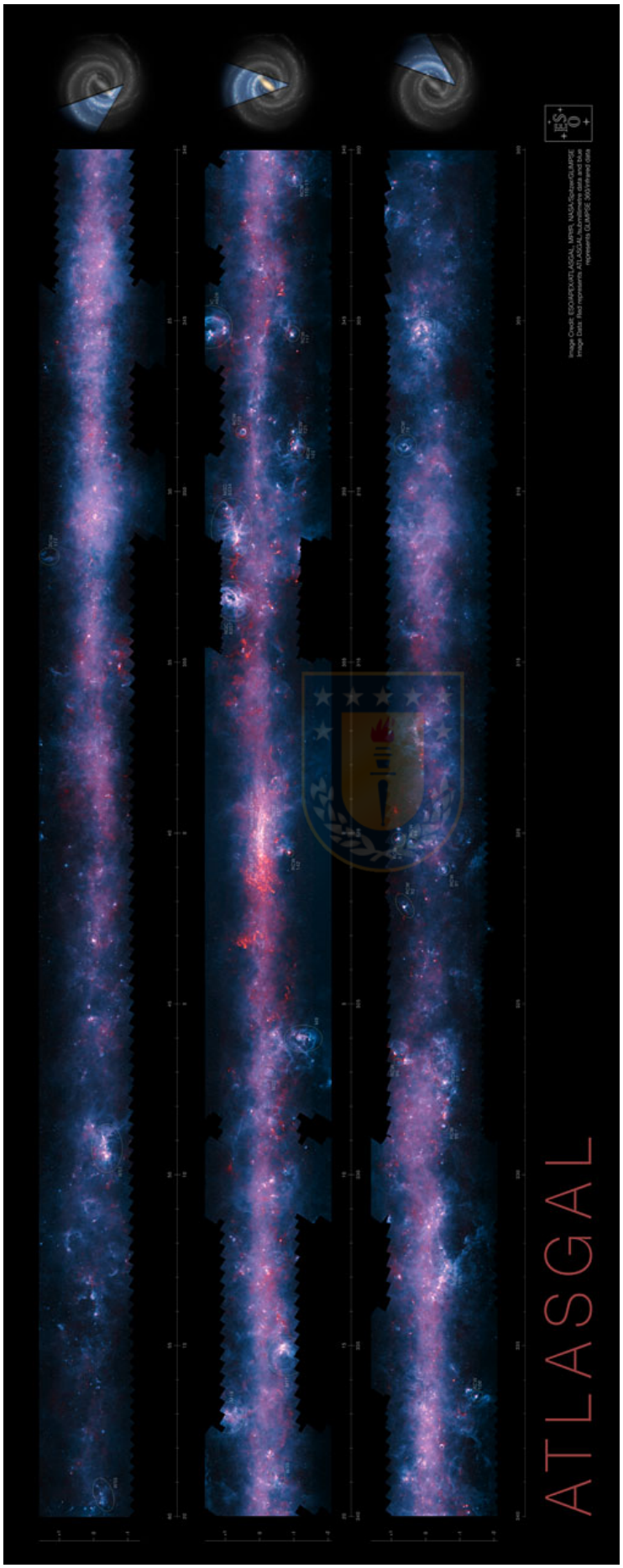

FIGURE 1.8: Coverage of the ATLASGAL sample, showing also in the left sides of very picture the regions of the galaxy that was covered. Image taken from the ESO website [a]

[a] www.eso.org

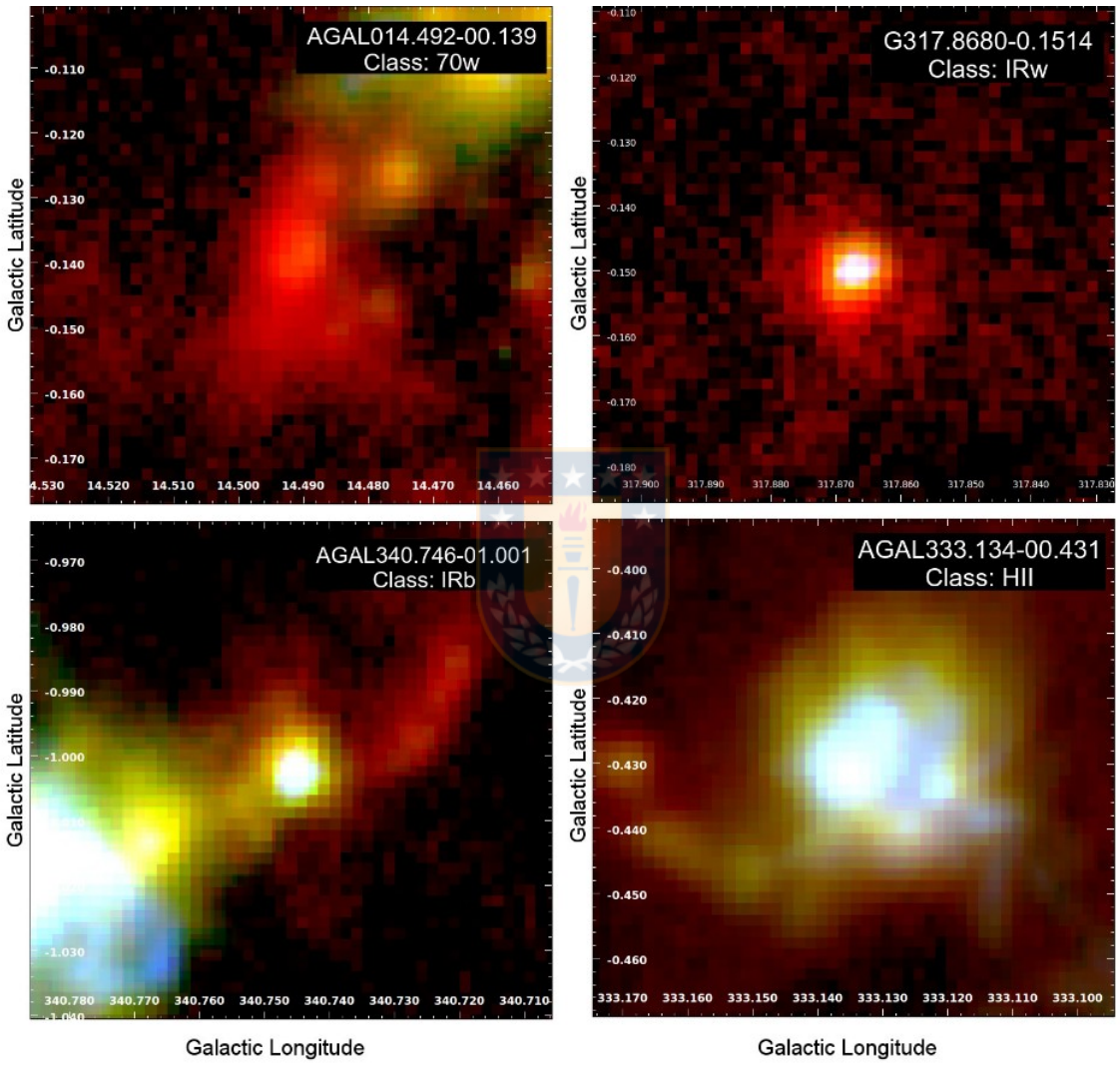

FIGURE 1.10: Three color images for each kind of sources classified from the TOP100 subsample of ATLASGAL. Figure adapted from König et al. (2017).

Chapter 2

Chemistry in the pre-stellar phase

Pre-stellar cores are rotating and collapsing magnetized dense cloud cores. Dense cores without any star already formed are called "starless cores", with average volume densities at least one order of magnitude larger that their surrounding medium and temperatures around 10 K. These starless cores are dominated by thermal motions in their insides, and not all of these clouds produce stars. Some of the clouds can reach hydrostatic equilibrium and display kinematic features similar to oscillations, or can produce expanding motions. Inside pre-stellar cores there can be a multiple densities that can go from $10^4 \mathrm{cm}^{-3}$ to $10^7 \mathrm{cm}^{-3}$, corresponding to the edges of the core and to the center of it, respectively. The chemical reactions also change strongly around the cloud. An example is reported in Fig. 2.1 for the pre-stellar core L1544, embedded in the Taurus Molecular Cloud Complex. In this figure the different chemical zones are shown in cyan contours. In the right side of the figure there are the corresponding more important chemical reactions, in blue are indicated the reaction partners. In particular, in the outer part of the core, the gas density is $\simeq 10^4$ cm^{-3} and temperature around $\simeq$ 10 K. In this region of the cloud, the chemistry consists of ion-molecule reactions dominating the carbon chemistry and neutral-neutral reactions. The most important reactions that take place in the central part of the cloud, will be reviewed in the next subsections.

2.1 Gas-phase chemistry

As it was already stated, molecular clouds are dominated by molecular gas, while the gas in the diffuse neutral phases of the ISM is almost entirely atomic. The transition that leads to molecular gas is regulated by the chemical state of hydrogen and carbon. H_2 is the simplest and most abundant molecule in the ISM, with a crucial role in many environments. However, this molecule is not very easy to observe, because being an homonuclear symmetric molecule does not possess a permanent dipole moment, and the rovibrational transitions in its ground state occur through quadrupole radiation (Bovino and Galli, 2019). CO is the second most abundant molecule (10^{-4} the H_2 abundance) and is instead used as a probe of the molecular gas.

Under typical ISM conditions, the formation of H_2 in gas-phase is extremely inefficient. Although, there are huge amounts of H_2 observed in molecular clouds. The most efficient reaction to form this molecule is through catalysis on the surface of dust grains. The formation of H_2 on the surface of dust grains will be explained with more details in the following sections.

On the other hand, CO forms through two sets of chemical intermediates. One path is obtaining hydroxyl (OH), then OH^+ and their products, while the other path involves CH and CH_2 and their positive ions. The neutral-neutral reaction that converts OH into CO does not need activation energy, something very rare between gas-phase reactions which also allows them to occur in very low temperature environments.

$$\mathrm{C} + \mathrm{OH} \rightarrow \mathrm{CO} + \mathrm{H} \tag{2.1}$$

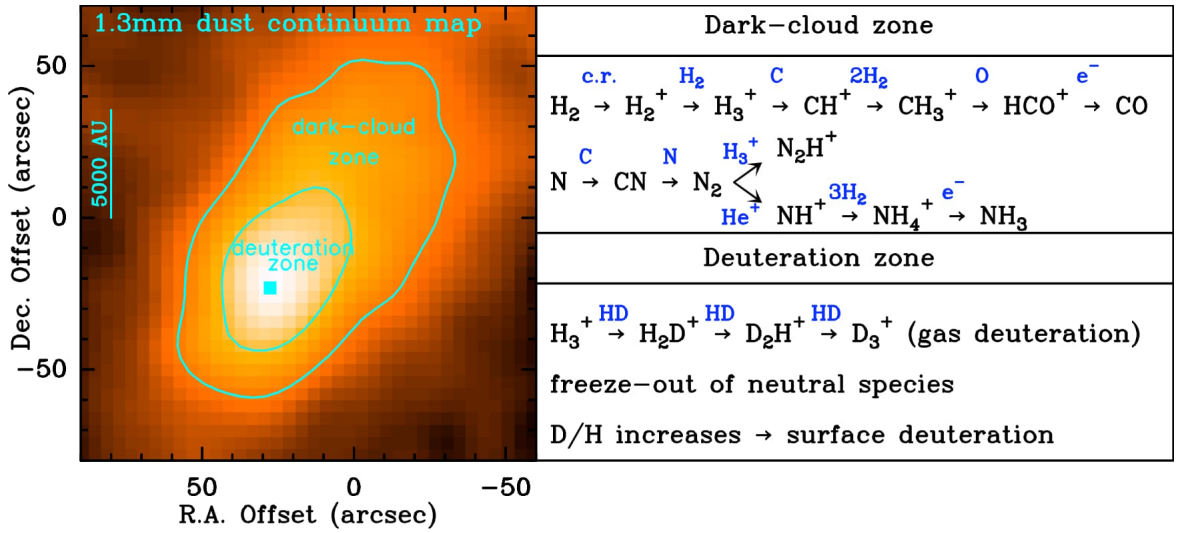

FIGURE 2.1: Pre-stellar core L1544, embedded in the Taurus Molecular Cloud Complex. In this figure the different chemical zones are shown in cyan contours. In the right side of the figure there are the corresponding more important chemical reactions, in blue are indicated the reaction partners. Taken from Caselli and Ceccarelli (2012)

while, the formation of CO^+ is caused by the following reaction

$$C^+ + OH \rightarrow CO^+ + H \tag{2.2}$$

in places where the C^+/C ratio is high, and then followed by

$$CO^+ + H \rightarrow CO + H^+ \tag{2.3}$$

or firstly forming HCO^+ in the reactions

$$CO^+ + H_2 \rightarrow HCO^+ + H, \tag{2.4}$$

$$HCO^+ + e^- \rightarrow CO + H \tag{2.5}$$

easily forming CO, although the formation of OH is not very simple for these environments. It is also important to notice that the previous reactions are endothermic, therefore they are only common in hot gas environments.

Another path that leads to the formation of CO is the following

$$C^+ + H_2 \rightarrow CH^+ + H, \tag{2.6}$$

or by radiative association

$$C^+ + H \rightarrow CH^+ + \gamma \tag{2.7}$$

From these two reactions, the first one is dominant over the second one, but is also an endothermic reaction, that needs 0.4 eV to occur, which means that is not a reaction that can occur inside molecular clouds.

CO can also be formed inside gas with considerable amounts of C^+ and H_2

$$C^+ + H_2 \rightarrow CH_2^+ + \gamma \tag{2.8}$$

The previous reaction has a large rate coefficient, compared to the C^+ and H_2 reaction, therefore in regions that present $n_{H_2} \gtrsim n_H$ is the main starting point to form CO.
Once CH^+ or CH_2^+ has been formed, they can react with H_2

$$CH^+ + H_2 \rightarrow CH_2^+ + H, \tag{2.9}$$
$$CH_2^+ + H_2 \rightarrow CH_3^+ + H. \tag{2.10}$$

Most of the CH_3^+ is destroyed by dissociative recombination, as follows

$$CH_3^+ + e^- \rightarrow CH + H_2,$$
$$CH_3^+ + e^- \rightarrow CH + H + H,$$
$$CH_3^+ + e^- \rightarrow CH_2 + H.$$

And reacting with atomic hydrogen, finally produce CO,

$$CH + O \rightarrow CO + H,$$
$$CH_2 + O \rightarrow CO + H_2,$$
$$CH_2 + O \rightarrow CO + H + H.$$

All the different paths to obtain CO rely on the abundance of molecular hydrogen. The characteristic timescales of the chemical reactions that lead to CO formation tend to be shorter than that of the H_2 formation. Most of this section was taken from Klessen and Glover (2016).

2.2 Gas-grain chemistry

Inside the central part of the cloud, the density reaches values around $10^5 cm^{-3}$, hence the only radiative source that is able to penetrate the cloud are cosmic rays and the temperature drops below 10 K. Under these conditions species heavier than He are frozen onto the ice mantles of grain surfaces, disappearing from the gas-phase. This process is called "freeze out". The freeze-out is a natural consequence of the quiescent nature of pre-stellar cores: when species land on grain surfaces, they cannot thermally evaporate neither photodesorb as interstellar photons cannot penetrate pre-stellar cores (Caselli and Ceccarelli, 2012). CO freeze-out can reach a level of 80-90%, a process called CO depletion. The depletion of CO and other neutrals cause variation in the deuterium fractionation (see section 2.3) and the degree of ionization. In dense cores of molecular clouds, the degree of ionization is thought to be an essential parameter in regulating the rate of star formation. We know already that the best tracers of the central regions of pre-stellar cores are the deuterated species. With the observation of these species and their non-deuterated isotopologues, it is possible to measure the elusive electron number density $n(e^-)$. This quantity plays a crucial role on the dynamical evolution of the cloud, because during gravitational collapse neutral species are strongly affected by magnetic fields being attracted to it and colliding with molecular ions, in a process called ambipolar diffusion (Mouschovias, 1979). Neutral-ion collisions can slow down the collapse depending on the fraction of ions present in the gas-phase.

2.2.1 Dust grains and ice formation on the ISM

Dust grains are essential in the physical and chemical behavior of the clouds, because they drive the interstellar atomic gas to molecular gas, beginning with the chain that leads to chemical complexity.

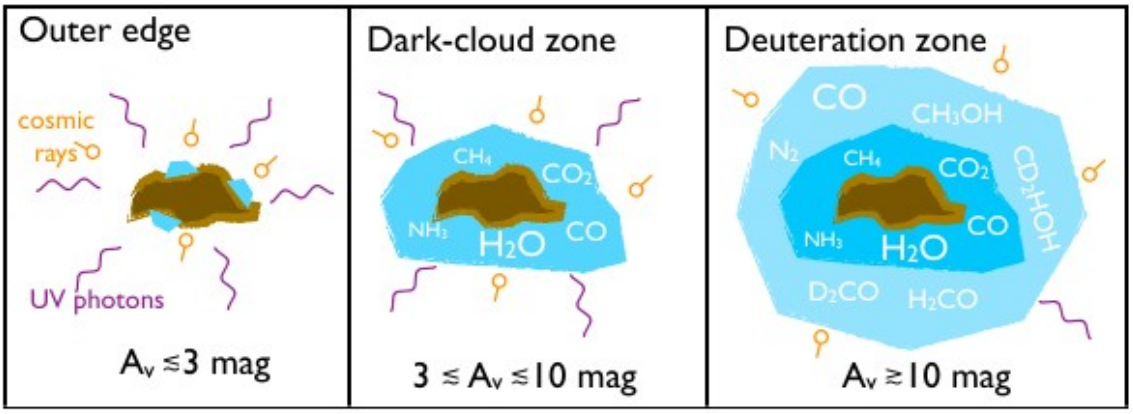

FIGURE 2.2: Illustration taken from Caselli and Ceccarelli (2012). This figure shows the evolution of the ice mantle growing on the surface of the grain. The ice mantle gets bigger and earns different species, more complex organic molecules, when the dust grains are located closer to the center of the pre-stellar core.

Dust grains also behave as filters of FUV photons, absorbing them and protecting the cloud and the molecules inside it from the UV destruction. Also, dust grains help to catalyze the formation of species, by capturing some simple molecules on its surface, where they can easily interact with other trapped molecules, making the reaction between these two much easier than it would be in a gas phase. They also act as gas coolants in the center of the clouds, making the temperatures drop even below 10 K. Ice mantles begin to grow as we get inside the clouds, as it is shown in figure 2.2 and start to comprise a bigger diversity of species and more complex molecules.

The thickness of the ice mantle will also depend on the density of molecules present in the cloud and on their freeze-out rate (e.g. we know that CO freeze-out rate is about 90% in the pre-stellar core).

2.3 Deuterium fractionation

As we get inside the molecular cloud, where star formation occurs, temperature decreases, while density increases. Under these conditions, heavy molecular species like CO freeze-out onto dust grain surfaces. This process is called "depletion" and proceeds progressively faster with increasing densities (typical timescale goes as $\propto 10^9/n_{H_2}$ yr). When CO depletion occurs, the following reactions

$$H_3^+ + CO \rightarrow HCO^+ + H_2 \tag{2.11}$$

$$H_2D^+ + CO \rightarrow DCO^+ + H_2 \tag{2.12}$$

become less important, boosting the abundance of H_2D^+ through

$$H_3^+ + HD \rightleftharpoons H_2D^+ + H_2 + 232K \tag{2.13}$$

and with it, the abundance of other deuterated species (e.g. N_2D^+), a process called deuterium fractionation (Caselli et al., 1999, 2003). In the above reactions, H_3^+ is produced by the cosmic ray ionization of H_2 that initiates the ion-neutral reactions in dense clouds

$$H_2 + \zeta_{C.R.} \rightarrow H_2^+ + e^- \tag{2.14}$$

$$H_2^+ + H_2 \rightarrow H_3^+ + H \tag{2.15}$$

The exothermic reaction chain $H_3^+ \rightarrow H_2D^+ \rightarrow D_2H^+ \rightarrow D_3^+$ is essential for deuteration process, and also the deuteration of H_3^+ depends on the ortho-to-para ratio of H_2. When this ratio decreases, there are more exothermic reactions, which leads to more deuteration (Sipilä et al., 2010). The molecules H_2D^+, D_2H^+ and D_3^+, together with H_3^+ are likely the most abundant ions in the nuclei of prestellar cores, and they react with other species increasing the amount of deuterated species at the core of the ion-neutral chemistry (Giannetti et al., 2019).
In molecular clouds, deuterium atoms exist mainly as HD. The ratio of deuterated species relative to their non-deuterated molecules of the same species is around 0.01-0.1. This ratio is much higher than the cosmic abundance of D/H, which is around $1.52 \pm 0.08 \times 10^{-5}$. The ratio between deuterated and non-deuterated species is called deuterium fractionation, e.g.,

$$D_{\text{frac}}^{N_2H^+} = \frac{N[N_2D^+]}{N[N_2H^+]} \tag{2.16}$$

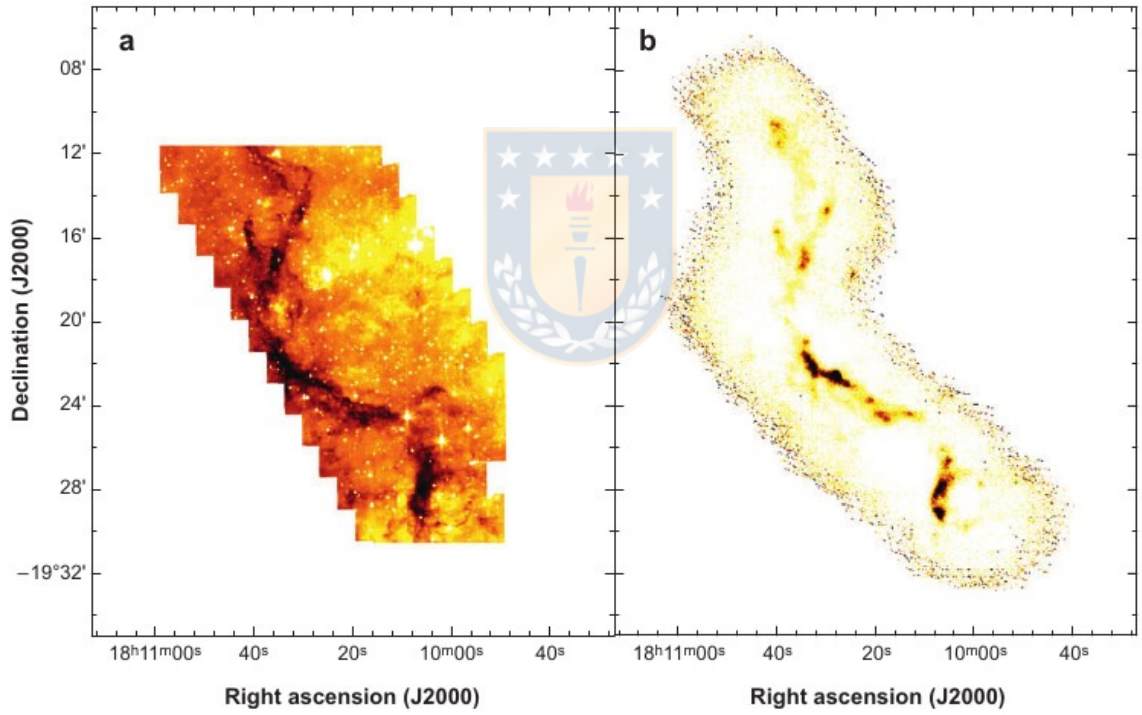

FIGURE 2.3: G11.11-0.12 IR dark cloud as seen by Spitzer IRAC band 4 in the left panel, and seen with an 850-μm sub-mm continuum emission, in the right panel. Taken from Bergin and Tafalla (2007)

Nitrogenated species such as N_2H^+ and N_2D^+ have been used as tracers for the physical properties and kinematics of very dense regions (Punanova et al., 2018; Redaelli et al., 2019). In fact, H_2D^+ and N_2D^+ are linked by the following reaction:

$$N_2 + H_2D^+ \rightarrow N_2D^+ + H_2 \tag{2.17}$$

For this reason, it has been proposed that these two species can be used in conjunction to determine the evolutionary stage of a source in the early phases of the star formation process. In a recent work

Giannetti et al. (2019) observed H_2D^+ and N_2D^+ in three clumps in different evolutionary stages, reporting an increase in the abundance of N_2D^+ with the clumps evolutionary stage, and a decrease of the abundance of H_2D^+.

2.3.1 Spin state chemistry

ortho-H_2D^+

The H_2D^+ molecule consists of two electrons bound to three nuclei (two protons and one deuteron), with an asymmetric prolate top with three vibrational modes: breathing (ν_1), bending (ν_2) and asymmetric stretch (ν_3), all of those are infrared active modes. Together with D_2H^+ are good tracers for H_3^+, the most important and the most abundant molecule in the interstellar medium (ISM), which has no dipole moment and cannot be directly observed. Between H_2D^+ and D_2H^+, the first is the most abundant and easier to observe. (Sochi and Tennyson, 2010). In this work we are studying the submillimeter-wave transition of the rotational 1_{10}-1_{11} transition line of ortho-H_2D^+, which is considered to be the most abundant ion in pre-stellar cores. In Fig. 2.4 we show an example of a detection of this molecule at the frequency of 372421.385 MHz.

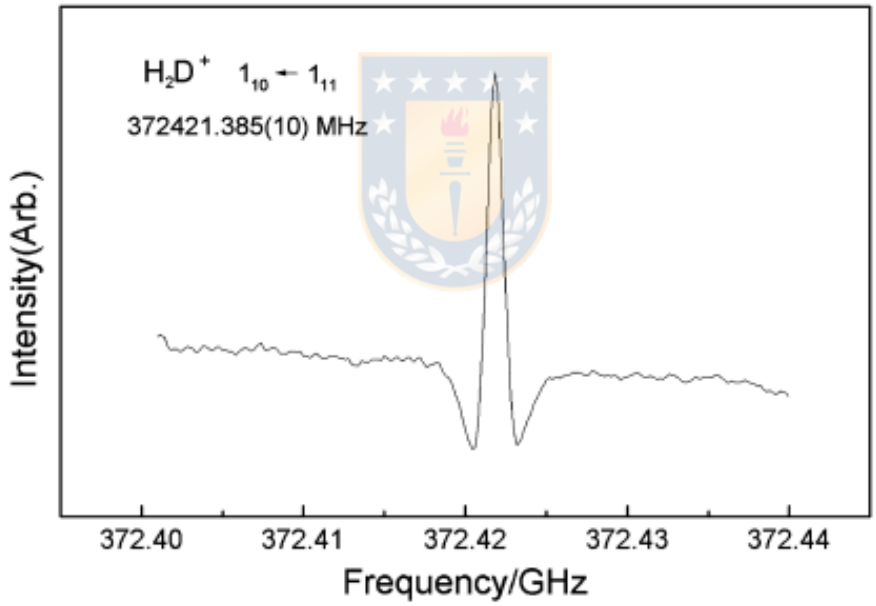

FIGURE 2.4: Example of an observed spectra of the $J = 1_{10}$-1_{11} transition line of H_2D^+. The y axis indicates the intensity and the x axis is the frequency in GHz. Figure taken from Amano and Hirao (2005)

The energy levels are designed by J_K where J is the total rotational angular momentum and K the projection of the angular momentum onto the symmetry axis of the molecule. H_2D^+ has a very large dipole moment, directed along its axis of symmetry (Khersonkii et al., 1987). Electric dipole transition do not allow the projected angular momentum to change, which means that $\Delta K = 0$. Because of this the energy level structure is usually organized in K-ladders, designated as 'ortho' (for K = 0, 3, 6, etc.) or 'para' (for K = 1, 2, 4, etc.) according to the spin symmetry. The ortho-H_2D^+ has its nuclear spins aligned parallel, while the para-H_2D^+ has its nuclear spins aligned anti parallel. In Fig 2.5 it is illustrated the low-lying energy levels of H_2D^+. The bold arrows indicate the transitions measured experimentally.

Since J and K do not change in a pure inversion transition, the notation for these transitions is (J, K) (Shirley, 2015).

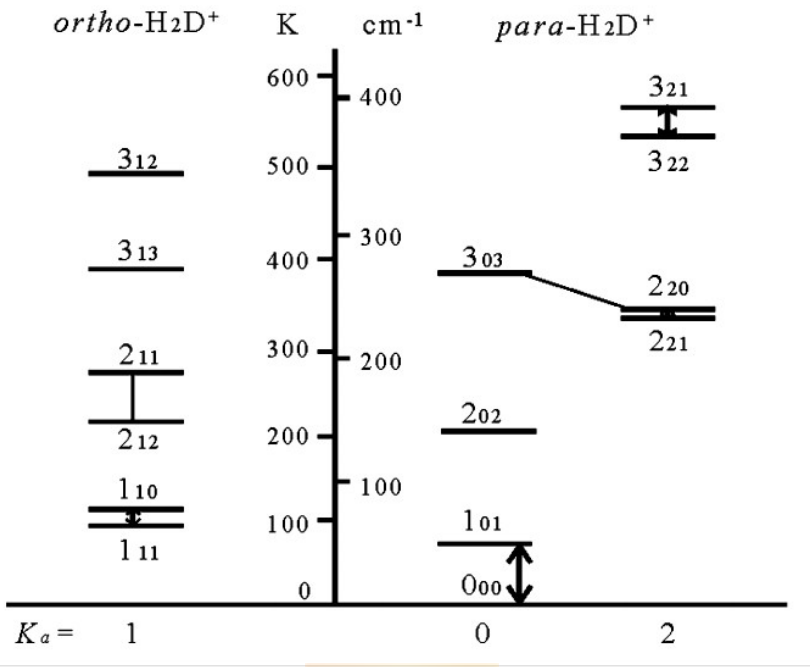

FIGURE 2.5: Low-lying rotational levels of H_2D^+. Bold arrows indicate the observed transitions. Thin lines show transitions that fall below 2 THz. Figure taken from Amano and Hirao (2005).

Chapter 3

Analysis of the data

This chapter is based on the published work of Sabatini et al. (2020).
The main goal of this study is testing the trend that has been already reported by Giannetti et al. (2019), where three clumps embedded in the G351 (see Fig. 3.1) complex showed that the less evolved clump (C7) presented weak emission lines of N_2D^+ and high levels of o-H_2D^+, the medium evolved clump (C5) presented similar levels of both species, and the most evolved clump (C2) showed the highest levels of N_2D^+, as can be seen in Fig. 3.2. From this anti-correlation, it has been proposed that these two molecules can work as evolutionary tracers of the high-mass star-formation process. In this study, we will test this hypothesis on a sample that is statistically representative of regions with ongoing massive star formation activity, excluding possible bias induced by specific initial chemical conditions, that in smaller samples could have affected the results. In particular, we shall focus on the o-H_2D^+ line, while N_2D^+ observations in the same sample are ongoing and a complete study of the o-H_2D^+/N_2D^+ ratio as evolutionary tracer will be pursued in a future work.

3.1 The observations

The o-H_2D^+ 1(1-0)-1(1-1) observations were made at 372.43 GHz using the on-off (ONOFF) observing mode with the FLASH345-XFFTS dual-frequency MPIfR principal investigator (PI) receiver mounted at the Atacama Pathfinder Experiment (APEX) telescope (Güsten et al., 2006); Heyminck et al. (2006). Among 99 of the 110 sources of the TOP100 sample (Giannetti et al., 2014; König et al., 2017) were observed in three projects (IDs: 0101.F-9517; M-097.F-0039-2016 and M-098.F-0013-2016) and 77 observations from July 2017 to December 2018, with a final number of 11568 spectra (about 20-30 sec. of integration time each). The subsample we analized was selected based on two spectral line surveys, both of them of the ATLASGAL sources. The first set of observations consists of 16 sources (not fully contained in the TOP100 subsample) selected by a sample of young massive clumps that showed a high degree of deuterated ammonia (Wienen et al., 2021). The second set of observations belong to a [CI] $^3P_1$ - $^3P_0$ fine structure line at 492 GHz Survey (Lee, et al., in prep.). The spectra have an extension of 2.5 GHz - from 372250 to 374750 MHz - with a native spectral resolution of 0.038 km s^{-1}.

In the next section we introduce the main properties of the clumps belonging to the aforementioned sample.

3.1.1 Main properties of the observed clumps

The dust continuum was taken from Giannetti et al. (2017); König et al. (2017) and Urquhart et al. (2018). The heliocentric distances for our sample ($d_\odot$) varies from 1 to $\sim$ 5.5 kpc were obtained from

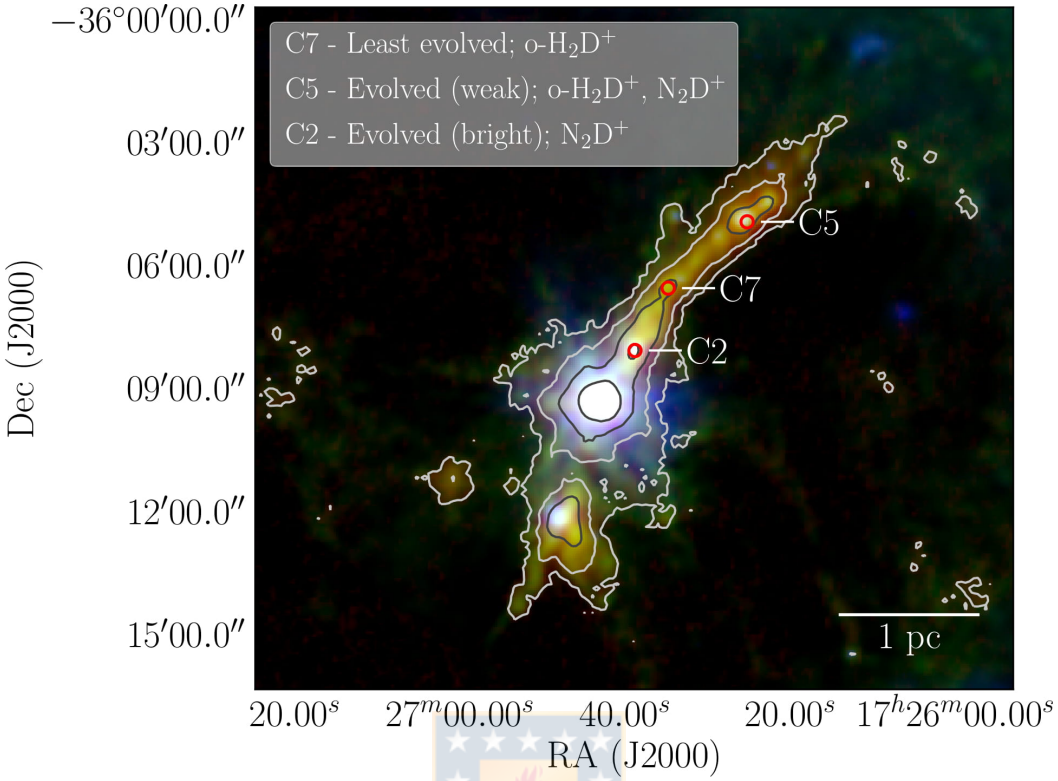

FIGURE 3.1: Three-colour image, composed by the ATLASGAL 870 μm, Hi-HAL 250 μm, and Hi-Hal 70 μm. The three clumps object of the study, showing the clump 7, 5 and 2 with red circles. Taken from Giannetti et al. (2019).

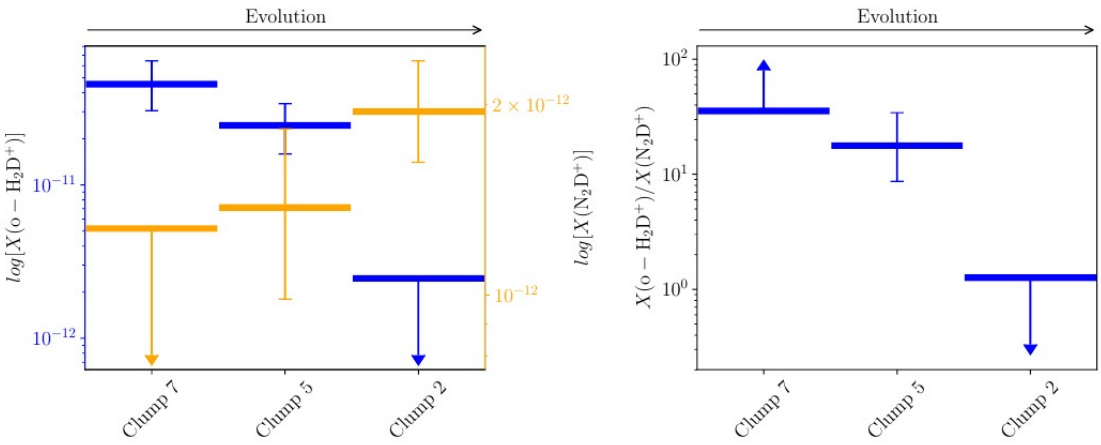

FIGURE 3.2: Left panel shows the o-H_2D^+ and N_2D^+ abundances in blue and orange, respectively. In the right panel the ratio between the two relative abundances is reported. In both figures the clumps are shown from the least evolved to the most evolved one. Taken from Giannetti et al. (2019).

König et al. (2017) and Urquhart 2018. The Galactocentric radii D_{GC} varies from 4.5 and 7.5 kpc and have a associated error $\sim$ 0.3 kpc (Urquhart et al., 2018) when the distance to the Galactic Center is assumed to be 8.35 kpc (Reid et al., 2014). The effective radii (R_{eff}) varies from 0.1 to 1.3 pc. The radial velocities (v_{lsr}) are based on $C^{13}O$ J = 3 - 2 lines taken from Giannetti et al. (2014), and also NH_3 1,1 lines from Wienen et al. (2012). The dust temperatures (T_{dust}) are between $\sim$ 10 and $\sim$ 25 K and its error associated is derived from the covarience matrix of the SED fit, obtained using a Levenberg-Marquardt least-squares minimization, all the process is explained in the Appendix D of Urquhart et al. (2018). The clump masses (M_{clump}) are between $\sim$ 100 - 5400 $M_\odot$ and the bolometric luminosities (L_{bol}). The column densities of molecular hydrogen, are measured in logarithmic scale, $\log_{10}[N(H_2)]$, between 22.7 and 23.3. More than our sample belongs to the TOP100 subsample, we use the evolutionary classification already reviewed in the subsection 1.3.4. The observed properties obtained for the sources with o-H_2D^+ detection are summarized in Table 3.1.

3.2 Data filtering and pre-processing

The main goal of this work is to check if the behaviour of o-H_2D^+ is valid also in a representative subsample of the ATLASGAL-Survey and, with this, proving that this behaviour is independent of the physical properties of the environment in which the clumps are embedded. The sample is made by a collection of multiple spectra for every source. What we want to obtain is one final spectrum per source where the o-H_2D^+ is detected. To analyse the observational data we will make use of `GILDAS`. [1] `GILDAS` (Grenoble Image and Line Data Analysis Software) is an image processing system oriented toward (sub-)-millimeter radioastronomical applications. It was built from several pre-existing programs, e.g., `CLASS` (Continuum and Line Analysis Single-dish Software) and `GREG` (Graphical Possibilities), `ASTRO` (Ephemeris and observations preparation) between others.
In this work we have been using `CLASS`, which is mainly oriented for processing spectra of single dish telescopes.

First, we enumerate the sources that are inside the sample we obtained. Then, from the `CLASS` terminal we check the spectra related to every sources, one by one looking for possible corrupted observations. With that in mind, we apply the code for every source in our sample. The code will generate an average single spectrum from all the different observations, where we will be able to check if there is an emission line or not. A second inspection of the obtained spectra generated by the code is needed. In Fig. 3.3 we present a summary of the code steps in order to obtain the final spectrum for every source.

3.2.1 Data reduction and identification of the o-H_2D^+ emission lines

The first step is to obtain the spectra useful for our purpose from the entire observations database, since it contains also other tracers and calibration data not needed in our study. We look into more than 10000 different spectra available to select only those that correspond to ATLASGAL sources and we check one by one the multiple observations associated to every source, looking for possible corrupted observations that will have to be deleted. We select the spectra according to the different sources, then we cut it around the location of the emission line, delete the damaged observations, modify the specific beam to the APEX telescope, correct the effect of the shape of the antenna visible in the spectra, average all the spectra found for all different sources individually and finally write the results as a new spectra. In this chapter we will show and explain in details the two codes used

[1] `https://www.iram.fr/IRAMFR/GILDAS/`

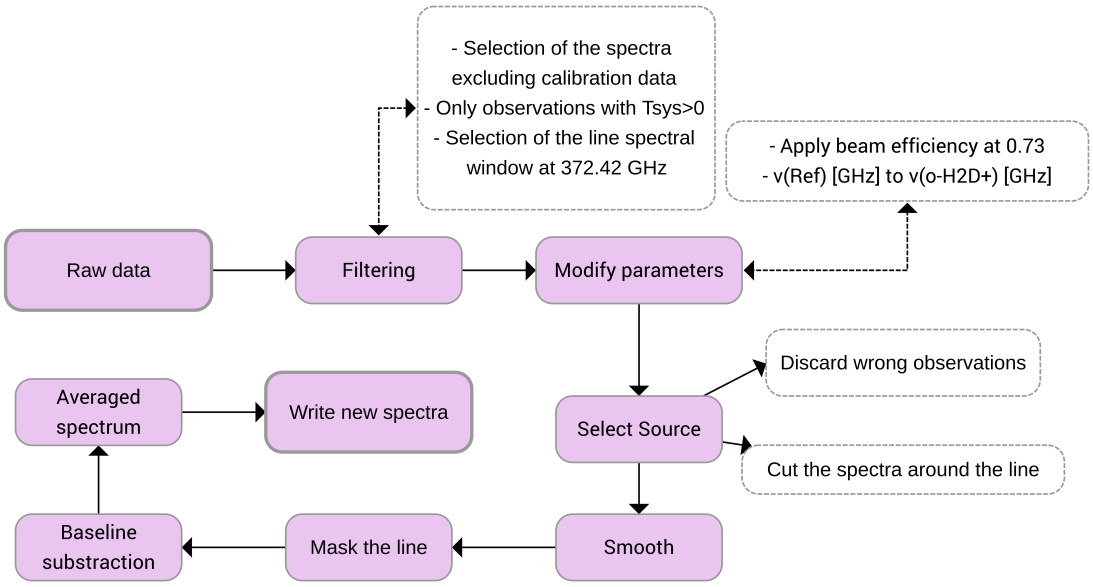

FIGURE 3.3: Flowchart that summarizes the operation of the code.

for obtaining the spectra used to study the o-H_2D^+ emission lines.
The code for the data filtering and the modification of the parameters is:

```
import pyclass as pc
import glob
import numpy as np
import pgutils


file_in = glob.glob('*.apex')


try:
  pc.comm('file out data.merged sing')
except pgutils.PygildasError:
  pc.comm('file out data.merged')

for fin in file_in:
    pc.comm('file in ' + fin)
          pc.comm('set line GHZ374')
          pc.comm('set source G*')
          pc.comm('find')
          for i in xrange(pc.gdict.found.copy()):
              pc.comm('ge n')
                  pc.comm('modi f 372421.385')
                  pc.comm('modi beam 0.73')
                  pc.comm('set wind velocity-5 velocity+5')
                  pc.comm('base 1')
```

```
21                  if      pc.gdict.tsys.copy() > 0:
22                      pc.comm('write')
```

Spectra filtering: PyClass is a module of `GILDAS` designed to create an interoperable environment in which `CLASS` can also use all Python facilities (and vice versa). This also allows `CLASS` to be used simultaneously and interactively with other python-supported software. The `glob` module finds all the path names matching a specified pattern according to the rules used by the Unix shell. The `pgutils` module is made for analyzing large datasets with PostgreSQL.
The first thing we do with the code is to open all the APEX files with the command `filein = glob.glob('*.apex')`. With `pc.comm('...')` the commands for `gildas` can be written inside the quotation marks. In line 6 inside the `try` function, with the command `file out data.merged sing` we create a new empty file called *data.merged* that will be written later. Here, the `sing` specifies that only a single spectra will be saved per source. Between our sample there is more than 10000 spectra stored in different class files, because the sources were observed at different times, and this is why in the following steps we will reorganize all the files, selecting only the spectra correspondent to sources. The next step is to make a loop for all the files contained in the `file in`. In line 11 with `'file in' + fin` the spectra contained in the `file in` are opened one by one, and in line 12 we select only the sources at the frequency of our interest, where o-H_2D^+ is detected, i.e., 372.43 GHz, which in this case are highlighted by the label `GHZ374`. This is done with the command `set line GHZ374`. With `set source G*` we take all the files of the data set with a name that starts with the letter G, which are the sources from ATLASGAL. Next, the command `find` applies all the previous filters to the search of all sources.

Baseline substraction: In line 15 a loop over all the found spectra is performed, by using `pc.gdict.found.copy()`, and it begins selecting each spectrum one by one with `get next`. Line 17 is used to modify the rest frequency, that is passed in MHz. At 372 GHz the telescope has a beam size of 14 arcsec and a main-beam efficiency of 0.73[2], this is set with `modi beam 0.73`. This beam efficiency has been applied to the antenna temperature, to obtain the spectra in T_{mb}. The efficiency factor is different for every antenna, and it is given by the antenna composition: The light from different parts of a source are received by different parts of the antenna, called receivers. Every receiver covers a different spectral window, and they do not have necessarily the same response to the same frequency. More details about the antenna functioning are in Appendix A. For each spectra the reference frequency is set at the frequency of the o-H_2D^+ 1(1-0)-1(1-1) hyperfine molecular transition at $\sim$ 372.42 GHz (Amano and Hirao, 2005) with the command `modi f 372421.385`.
For every source we apply a first degree baseline, in order to correct the curve shape given by the antenna shape itself. Some spectra will even need two baselines subtractions, and to make the code more general, this double subtraction will be applied to all the sources. To approach this we will apply once the baseline, save this spectra in the data.merged file, and then apply the second baseline in the second code.
The code is the following:

```
1  import numpy as np
2  import astropy.units as u
3  import matplotlib.pyplot as plt
4  from astropy.io import fits
```

[2] http://www.apex-telescope.org/telescope/efficiency/

```
5  import pyclass as pc
6  import glob
7  import pgutils
8  pc.comm('file in data.merged')
9  pc.comm('find')
10 pc.comm('set source G14.49-0.14')
11 pc.comm('drop 2581')
12 pc.comm('drop 2582')
13 pc.comm('drop 3235')
15 pc.comm('set unit v v')
16 pc.comm('file out G14.49-0.14.bas sing ')
17 for i in xrange(pc.gdict.found.copy()):
18            pc.comm('get next')
21            pc.comm('extract 372358 372452 frequency')
22            pc.comm('stitch /nocheck position')
23            pc.comm('smooth box 17')
24            pc.comm('set wind velocity-5 velocity+5')
25            pc.comm('base 1')
26 pc.comm('set weight sigma')
27 pc.comm('plot')
28 pc.comm("say 'sigma'")
29 pc.comm('write')
```

Cleaning procedure and corrupted spectral removal: At this point the data.merged file contains only the filtered spectra with one baseline subtraction from the previous code. The first thing we do in line 8 is open the file 'data.merged' containing all these new spectra with `file in data.merged`. Next, with the `find` command, all the spectra in the file data.merged are read. In line 10 with `set source '...'` we select every source one by one. We present the particular case of the source G14.49-0.14, because the one-by-one observation examination showed three problematic spectra, shown in Fig. 3.4. The single spectra are deleted in the following lines 11, 12 and 13, with the command `drop 'number of the observation'`.In line 15, with `set unit v v` we specify that the units to be used in the lower and upper axis, where the 'v' stands for velocity. In line 16 we create an empty file for every source, with the command `file out G14.49-0.14.bas sing`.
The next step is to create a loop with the command `pc.gdict.found.copy()`. Inside the loop, the line 18 selects the 'next' spectra with the `get next` command. In line 21 we cut the spectra from 372358 MHz to 372452, with `extract 372358 372452 frequency`.

Averaging the spectra and applying the second baseline: In order to make the average of all the single spectra for each source, we use the command `stitch`. When we add `/nocheck position` to the stitch command, the code will assume that all the spectra are aligned in the same frequency and will stack them channel by channel. This average is made in order to reduce the noise in the final spectrum and make the emission line clearer, because the noise of every observation is random and it will be considerably decreased in the stitch process.
To help improving the signal to noise ratio, we can also reduce the spectral resolution. With the command `smooth box 17` the spectrum is averaged by channels 17, which decreases the spectral resolution from 30 m s $^{-1}$ to 0.5 km s^{-1}. In line 24 we create a mask of 10 km s^{-1} defined around the line found in the spectrum with the command `set wind velocity-5 velocity+5` to apply the

baseline subtraction in line 25 with `base 1`. Now, with `set weight sigma` we set the weighting to be used in the averaging step `stitch`. The `sigma` option makes the weight equals to $1/\sigma^2$. With `plot` we apply the changes to the spectrum and shows it. In line 28 the command `say sigma` prints the sigma values associated for each source spectrum. This sigma value will be an input for MCWeeds. Finally, the new spectra with double baseline subtraction is saved with the command `write`.

Even if the code provides a fast way to obtain the final spectra, each source with a detection has to be treated individually after the code is applied, to check that the final spectra will be a clean results, without involving any problematic observations in the process.
There were some problems with a lot of spectra in the sample. Some showed a sinusoidal behaviour, some of them had a characteristic noise value considerably smaller than zero and some simply did not look as expected. For every stitch spectra that looked problematic, all the single spectra had to be checked one by one to remove the possible corrupted observations. The best example for the spectra with evident problems is the source G14.49-0.14, where three problematic spectra were found. For clarification, in the top left panel of Fig. 3.4 it is shown how a normal observation spectrum looks like, while in the other three panels of Fig. 3.4 we show the corrupted observations of G14.49-0.14. These are the observations 2581, 2582 and 3235.

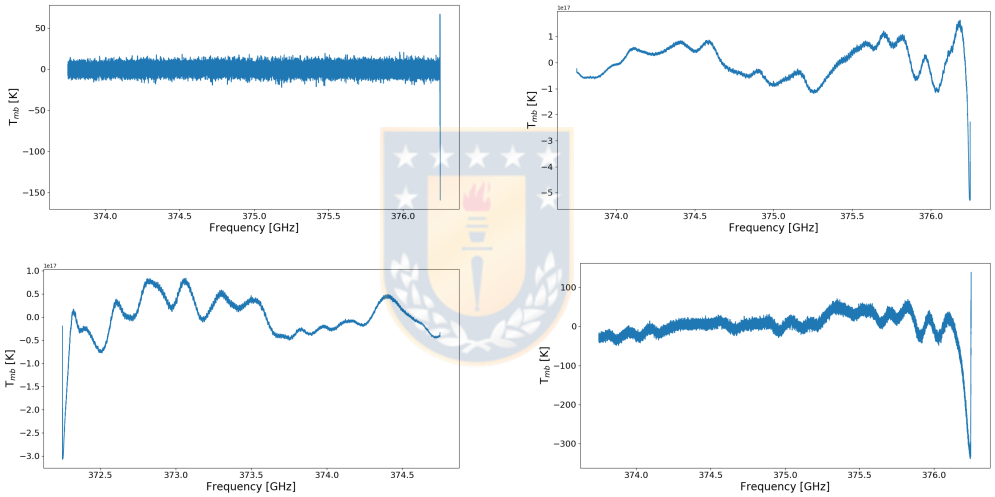

FIGURE 3.4: In the top left panel we show a normal, not problematic spectrum. In the other three panels we show the spectra with problems for the source G14.49-0.14. These spectra were not considered for creating the final spectrum of this source, as it was showed on the explanation of the codes.

After the three single spectra were deleted, as it is shown in the code above, the new spectrum shows a clear o-H_2D^+ line, see Fig. 3.5.

There are some interesting particular cases, such as G14.11-0.57, shown together with all the spectra with detections in Fig. 3.6, where the spectrum shows a double peak where the o-H_2D^+ should be found. This can be produced because of a double component of the emission line in the observed clump, with different rotations or velocities. This would require a deeper analysis and new observations, which goes beyond the scope of our work.

At the end of the entire procedure, we have the final spectra in which we are able to clearly identify if there is or not a detection of the targeted line. For the spectra showing detection, it is possible to obtain the column density of o-H_2D^+ and other relevant quantities, by using the software MCWeeds, which fit the spectra through a Bayesian analysis coupled to a Monte Carlo procedure. With the procedure described in the previous section we detected 16 sources with o-H_2D^+. We

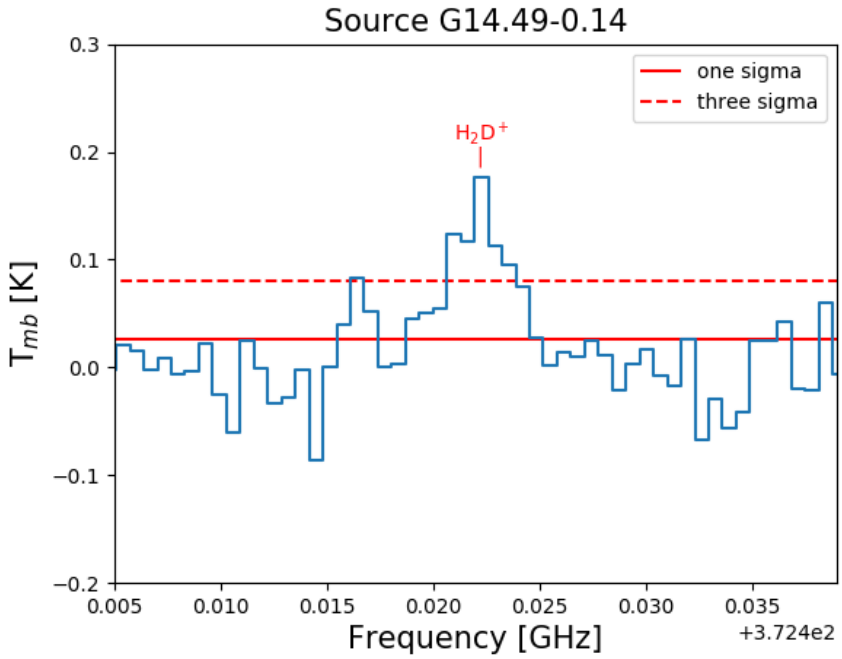

FIGURE 3.5: Detection of the source G14.49-0.14 after all the code procedure and deleting the corrupted observations. Source: This work.

applied the analysis to all our 106 sources, from which 99 belong to the ATLASGAL TOP100 sample. We found 12 detections on that belong to ATLASGAL, considering the source with the double peak as two detections, and 5 detections that are not in the TOP100 sources of ATLASGAL. The entire set of spectra is reported in Fig. 3.6 and 3.7.

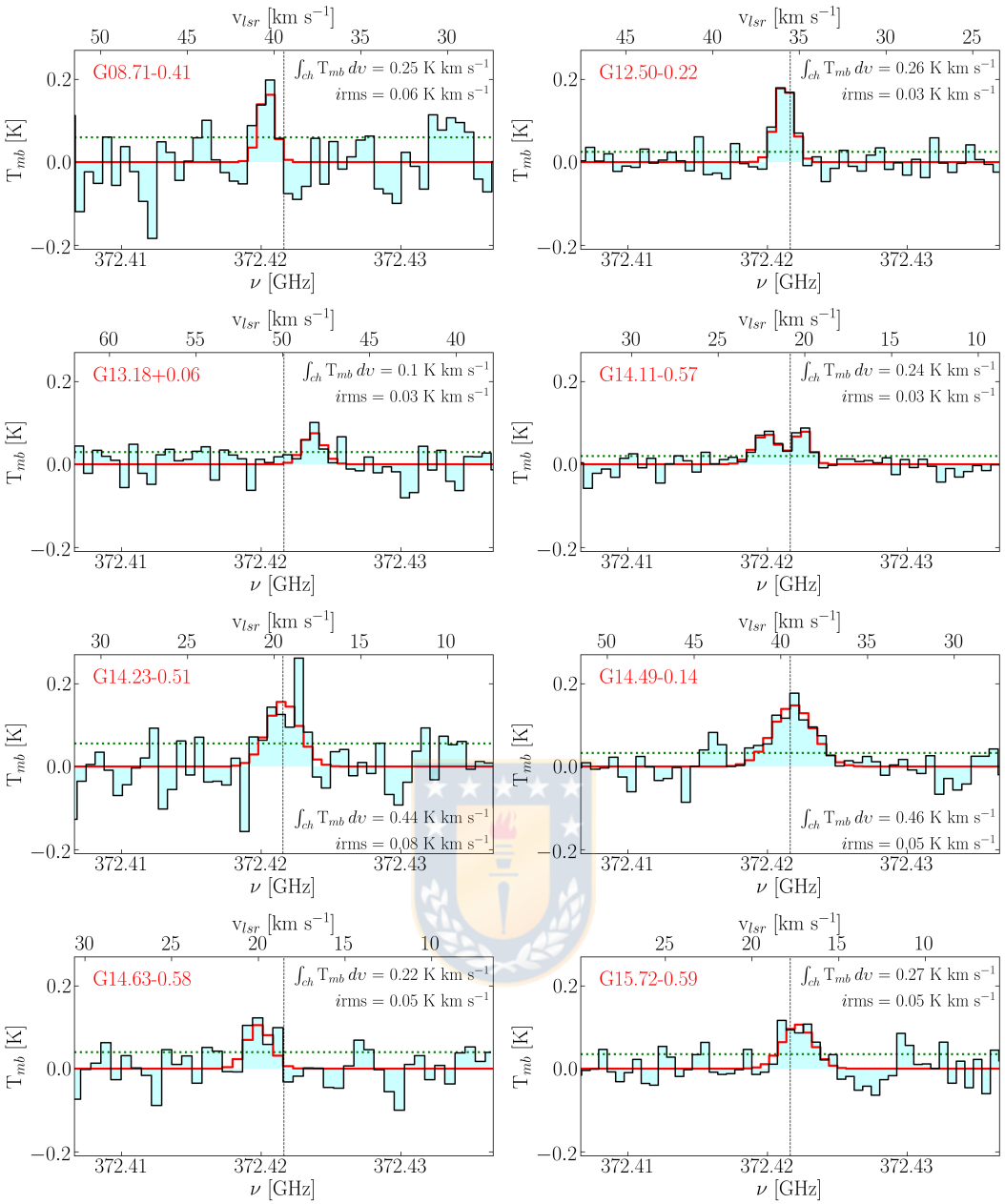

FIGURE 3.6: Half of the sources with o-H_2D^+ emission lines, represented in color cyan. The name of every source is in red on the left, the main-beam temperature ($\int_{ch} T_{\rm mb} dv$) and rms ($i$rms) are written in black to the right. 1σ of the $T_{\rm mb}$ noise levels are represented by the green dotted line. The frequency and velocity axes are cut around the line. The intensity axis, on the other hand, is fixed to the highest $T_{\rm mb}$ detected among all the detections, found in the source G14.23-0.51. The $v_{\rm lsr}$ of every source. In red it is shown the MCWeeds model performed for every source, that will be explained in section 3.3.1

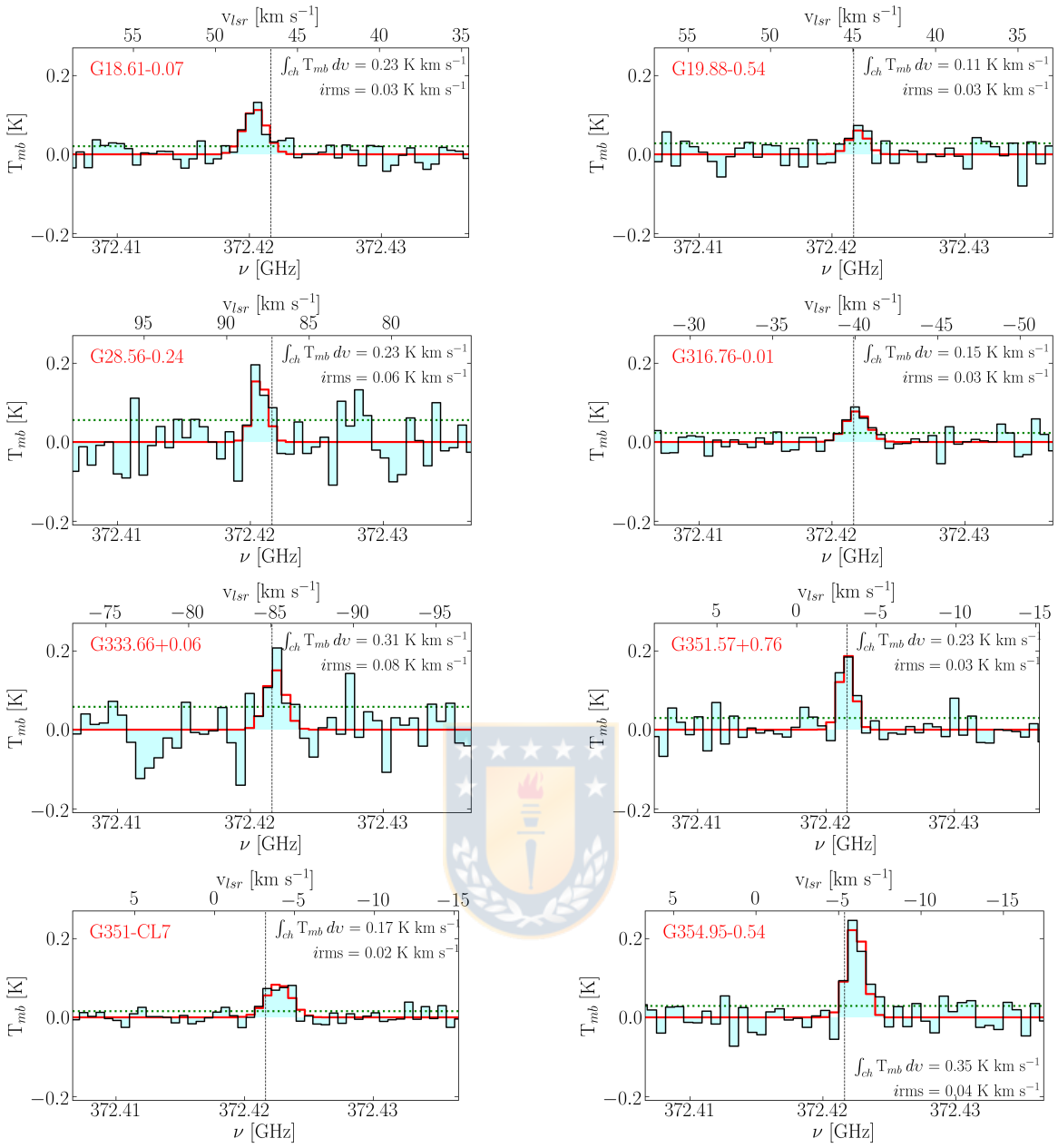

FIGURE 3.7: The second half of the sources with o-H_2D^+ detections.

3.3 Fitting procedure and column densities

In order to test the hypothesis to use o-H_2D^+ as an evolutionary tracer we have to calculate the column density. This is also very important to put constraints for theoretical models. The column density values can be obtained analytically by applying basic formulae, but in this study we decided to employ MCWeeds, an external interface between Weeds (Maret et al. 2011), simple and fast in building synthetic spectra (assuming local thermal equilibrium) and the Bayesian statistical models and adaptation algorithms of PyMC (Patil et al. 2010), which provide an accurate determination of the errors.

3.3.1 MCWeeds

MCWeeds is a code that allows to simultaneously fit the emission lines of a given spectra from an arbitrary number of species and components, assuming T_{ex} =T_{dust}, by combining multiple spectral ranges (Giannetti et al., 2017). This assumption holds for the case where the dust and the o-H_2D^+ molecules are homogeneously mixed, which might not be real for every sources since the averaged volume densities are slightly lower than the o-H_2D^+ critical density, $n_{cr} \sim 10^5$ cm^{-3} (see Caselli et al. 2008 and Vastel et al. 2012). We repeated the column density calculation considering T_{ex} = 0.6 T_{dust} according to Caselli et al. (2008), where excitation temperatures were found to be 40% lower than the dust values. For most of the cases the o-H_2D^+ column densities, are in agreement with the uncertainties reported in Table 3.2. Hence, the assumption of T_{ex} = T_{dust} does not affect the final results.
This fitting is performed using maximum a posteriori (MAP) estimates or Monte Carlo Markov Chains (MCMCs). With this, the fitting can be done in either a fast way, or a very detailed way.
The MCWeeds procedure is described in Fig. 3.8. The input includes the species to be analyzed, the spectral ranges and the data file where the spectra are stored. The fitting algorithm and the model parameters (i.e. column density, temperature, size of the emitting region, radial velocity and linewidth for every species, rms noise, and calibration uncertainty for each of the spectral ranges) should be also specified. MCWeeds generates a synthetic spectrum considering all the parameters given by the input file, a model parameter. The likelihood has to be computed assuming a Gaussian uncertainty, considering the σ as the rms noise, as the product of the probabilities for each channel. The model spectrum is compared to the observations in both of the methods: The likelihood *P(D—M)* probability of the data given the model, and the posterior method *P(M—D)* probability of the model given the data are computed. The new generated parameters are accepted or rejected as the process is repeated according to the iterations previously defined. MCWeeds provides the best fit parameters with their uncertainties, the best fit spectrum superimposed on the observations, the probability distribution function the trace plot and the autocorrelation of each the fit parameters.

3.3.2 The results

The column densities were computed assuming extended emission relative to the FWHM APEX-beam, supported by the results of Pillai et al. (2012), where the o-H_2D^+ line scale was found close to the typical clump size. For all the clumps with a clear o-H_2D^+ emission line, the angular size that corresponds to the effective radii, R_{eff} (König et al. 2017 and Urquhart et al. 2018) is larger than the APEX beam sizes, with the exception of the source G316.76-0.01 (Pillai et al. 2012). We performed fits with the assumption of the o-H_2D^+ molecular line parameters provided by the CDMS database. These parameters are: an Einstein-A coefficient $\log_{10}(A_{ul}\ [s^{-1}]) = -3.96$, a statistical weight of $g_{ul} = 9$, an energy gap between the two quantum levels $E_{ul} = 17.9$ K, and the partition function, $Q(T)$ for o-H_2D^+ in the relevant temperature range [9.375 K: 10.3375, 18.750 K: 12.5068, 37.500 K: 15.5054] (for more information, see Giannetti et al. 2019).

The results of the fitting performed by MCWeeds are reported in figures 3.6 and 3.7. where the noise level 1σ is the green dotted lines, the integrated mean-beam temperature and the rms for each line are indicated in every source. For every spectrum there is a clear emission line, with the exception of G14.11-0.57 (in Fig. 3.6) where two emission lines can be observed, with a distance of $\sim$ 2.5 km s^{-1} from each other. Since both of the emission lines have iSNR $\gg$ 3, they are treated separately on the following analysis, as C1 and C2. For all the sample the averaged noise remains more or less constant, from around 0.02 to 0.03 K (T_{mb} scale), with only a few cases with the noise around 0.05 K higher. All the FWHM line widths of the sample are between 1 and 3 km s $^{-1}$, which

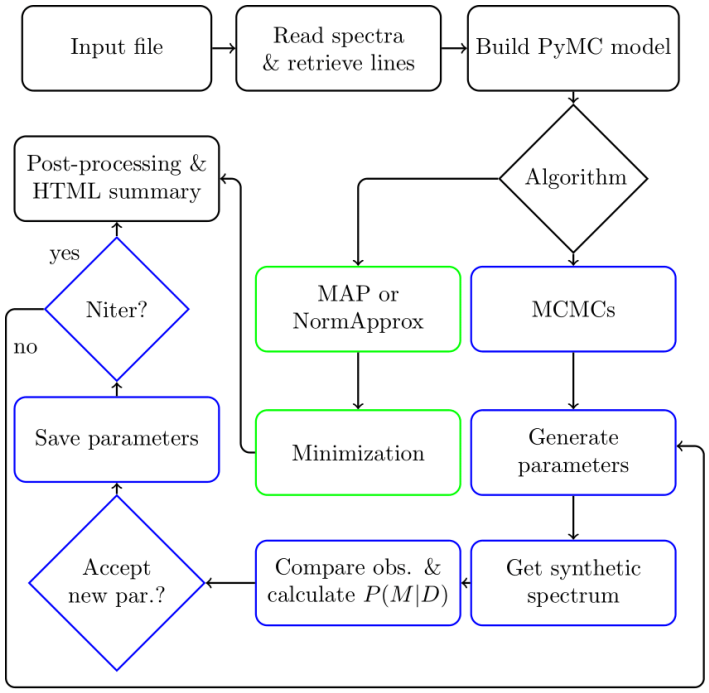

FIGURE 3.8: Schematic representation of the MCWeeds procedure. The MCMC steps are circled in blue and the MAP or normal approximation are in green. Figure taken from König et al. (2017).

is in concordance with the values for the $C^{13}O$ J = 3 - 2 and NH_3 (1,1) lines reported by Giannetti et al. (2014) and Wienen et al. (2012). In table 3.2 we show some of the sources with detections and their column density values obtained with MCWeeds.

The o-H_2D^+ column densities calculated for our sample varies from 1.3×10^{12} to 10^{13} cm^{-2},i.e., less than an order of magnitude. The abundances $X(o\text{-}H_2D^+)$ = N[o-H_2D^+]/N[H_2] were obtained with respect to the H_2 column densities reported by Giannetti et al. (2017) and Urquhart et al. (2018), and they vary from $\sim$ (1–12.6) $\times 10^{-11}$. The $X(o\text{-}H_2D^+)$ values found are slightly lower than the ones found previously on studies of low-mass pre- and protostellar regions (e.g. Vastel et al. 2006; Harju et al. 2008; Caselli et al. 2008; Friesen et al. 2010, 2014; and Miettinen 2020), but, on the other hand, were in agreement with previous studies focused on high-mass star-forming regions, i.e.,Harju et al. (2006), Pillai et al. (2012), and Giannetti et al. (2019); with the exceptional case of the study of Swift (2009) where significant lower values were found, with abundances around $\sim (3-5) \times 10^{-13}$ in the IRDCs G030.88+00.13 and G028.53-00.25.

The o-H_2D^+ detection limits were calculated using sources with no detection. These sources were classified according to their evolutionary stages, and between them were selected some observations with rms within a factor of 2 with respect to the average noise of the spectra with detection. From this filtering, we obtain 8 sources belonging to the 70w class, 16 to the IRw, 9 to the IRb and 13 sources HII regions. Since for each kind of source there are similar numbers, it is possible to compare the detections to a progressive evolution of clumps, with a o-H_2D^+ detection rate of 47% in 70w, 27% in IRw, 18% in IRb, and 7% in HII.

For each one of the sources (with and without the presence of the emission line) the detection limits were calculated, for obtaining the column density value corresponding to a 3σ detection. The FWHM line width used corresponds to 1.5 km s^{-1}, the average value of the detection sources. This only provides qualitative information when it comes to the upper limits of the o-H_2D^+ emission lines, and it is heavily influenced by the quality of the observations, more than by the physics of

the source. This problem is solved with only selecting the spectra with noise levels comparable to the ones with detected sources. In this way, the upper limit or relative abundance are inside the expected and typically observed range.

3.3.3 Dynamical quantities

From the spectra analysis MCWeeds also obtains an estimate of the FWHMs (Δv_{obs} in Table 3.2) for each emission line, and from this the virial parameter (α) and the Mach number ($\mathcal{M}$), both dynamical quantities, can be calculated. The virial parameter can indicate the dynamical state of every source, how far they are from the virial equilibrium (Chandrasekhar and Fermi 1953). The Mach number indicates the turbulence of the gas, if a clump is supported by dynamical or by thermal motions. The α values were estimated using the definition of Bertoldi and McKee (1992):

$$\alpha = \frac{M_{vir}}{M_{clump}} \quad \text{with} \quad M_{vir} = k_2 \left(\frac{\mathrm{R}_{clump}}{\mathrm{pc}}\right) \left(\frac{\Delta v_{dyn}}{\mathrm{km\,s^{-1}}}\right)^2. \tag{3.1}$$

Where M_{vir} is the source virial mass; M_{clump} is the total mass of the clump; R_{clump} is the clump size measured in parsec, and assumed to be equal to R_{eff}; and $k_2 = 210$ because the core is assumed to be homogeneous (MacLaren et al. 1988); Δv_{dyn} is the combination of the thermal motions of the particle of mean mass, $\Delta v_{th,\langle m\rangle}$ is the thermal gas motions, and Δv_{th} is the observed FWHM. The Δv_{dyn} values were derived using the values of Δv_{obs} in Table 3.2, following Kauffmann et al. (2013):

$$\Delta v_{dyn} = \sqrt{\Delta v_{obs}^2 + \Delta v_{th,\langle m\rangle}^2 - \Delta v_{th}^2}. \tag{3.2}$$

The Mach number was computed from the o-H_2D^+ line width provided by MCWeeds with:

$$\mathcal{M} = \frac{\sigma_{turb}}{\sigma_{\mathrm{H_2}}}, \tag{3.3}$$

where $\sigma_{turb} = (\sigma_{obs}^2 - \sigma_{th}^2)^{1/2}$ and $\sigma_{\mathrm{H_2}} = \Delta v_{th,\langle m\rangle}/(2\sqrt{2ln2})$.

$$\sigma_{turb} = \sqrt{\sigma_{obs}^2 - \sigma_{th}^2} \quad \text{and} \quad \sigma_{\mathrm{H_2}} = \frac{\Delta v_{th,\langle m\rangle}}{2\sqrt{2ln2}}. \tag{3.4}$$

$\sigma_{\mathrm{obs}} = \Delta v_{\mathrm{obs}}/(2\sqrt{2ln2})$ is the observed velocity dispersion of o-H_2D^+ lines, and $\sigma_{th} = [k_B\, T_{\mathrm{dust}}/m(\mathrm{H_2D^+})]^{1/2}$ is the gas thermal component at T_{dust}, where k_B is the *Boltzmann constant* and $m(\mathrm{H_2D^+})$ is the mass of the H_2D^+ molecule. The Mach number varies between $1.3 - 5.4$, which implies the presence of supersonic turbulent motions for the gas involved in all the sources. From the derived properties summarized in Table 3.2 it is suggested that the sample of this study is generally heterogeneous, and also composed of highly unstable sources (with $\alpha < 1$). There is also agreement between our results and the results presented on Kauffmann et al. (2013), where it was found that $M_{\mathrm{clump}} \gtrsim 10\, M_\odot$, $\alpha \ll 2$, and that these low α values are possibly consequences of a rapid collapse.

TABLE 3.1: Summary of the physical properties of the ATLASGAL sources in our sample.

ATLASGAL-ID	$d_\odot$ (kpc)	D_{GC} (kpc)	R^b_{eff} (pc)	T_{dust} (K)	ΔT^c_{dust} (K)	v_{lsr} (km s^{-1})	M_{clump} ($10^2 M_\odot$)	L_{bol} ($10^2 L_\odot$)	N(H_2) $\log_{10}$(cm^{-2})	n(H_2) $\log_{10}$(cm^{-3})	Class^{b,c}
G08.71–0.41^a	4.8	4.0	1.0	11.8	0.3	39.4	16.6	5.0	22.8	4.0	IRw
G13.18+0.06^a	2.4	5.9	0.5	24.2	0.8	49.9	3.7	83.2	22.9	4.3	70w
G14.11–0.57^a	2.6	6.9	0.5	22.4	0.8	20.8	3.5	31.8	22.9	4.4	IRw
G14.49–0.14^a	3.9	5.4	0.8	12.4	0.4	39.5	19.2	7.5	23.1	4.4	70w
G14.63–0.58^a	1.8	6.9	0.4	22.5	0.4	18.5	2.5	27.8	23.0	4.6	IRw
G18.61–0.07^a	4.3	5.3	0.7	13.8	0.3	46.6	8.8	5.9	22.8	4.2	IRw
G19.88–0.54^a	3.7	5.3	0.6	24.2	1.4	44.9	8.0	124.0	23.1	4.5	IRb
G28.56–0.24^a	5.5	4.8	1.3	11.7	0.1	87.3	54.1	17.7	23.1	4.2	IRw
G333.66+0.06^a	5.3	4.5	1.1	17.8	0.3	-85.1	14.2	42.7	22.7	3.9	70w
G351.57+0.76^a	1.3	7.0	0.3	17.0	0.1	-3.2	1.6	4.3	22.7	4.4	70w
G354.95–0.54^a	1.9	7.4	0.4	19.1	1.3	-5.4	1.5	4.8	22.6	4.2	70w
G12.50-0.22^b	2.6	5.9	0.4	13.0	0.2	35.5	1.2	1.6	22.8	4.4	70w
G14.23–0.51^b	1.5	6.9	0.5	17.2	2.7	19.5	7.2	14.8	23.3	4.8	HII
G15.72–0.59^b	1.8	6.6	0.2	12.1	0.5	17.8	1.7	0.3	22.8	4.7	IRw
G316.76-0.01^b	2.5	6.7	0.2	18.9	0.4	-39.9	4.7	24.2	23.1	5.0	IRb
G351.77-CL7	1.0	7.4	0.1^d	13.0	0.3^e	-3.2^d	1.2	0.2	22.9	5.2	70w

Physical properties (observed and derived) of our sample. In the top row we present the TOP100 sources belonging to our sample, meanwhile in the bottom row we present the ATLASGAL samples that do not belong to the TOP100 subsample. The T_{dust} and L_{bol} are obtained from a SED fit described in Urquhart et al. 2018, and their uncertainties as described in the introduction; R_{eff} values are evaluated from the dust continuum at 870 μm, the N(H_2) values are taken from the 870 μm flux peaks and the M_{clump} come from the integrated flux density over R_{eff}, respectively; n(H_2) values are lower limits derived here as n(H_2) = N(H_2)/2R_{eff}; $^{(a)}d_\odot$, M_{clump}, L_{bol}, and N(H_2) from Giannetti et al. (2017); R_{eff} and T_{dust} from König et al. (2017); D_{GC} from Urquhart et al. (2018); v_{lsr} from Giannetti et al. (2014) and derived from the $C^{17}O$ $J = 3-2$ transition; $^{(b)}$distances, R_{eff}, v_{lsr}, M_{clump}, L_{bol}, N(H_2), and classification from Urquhart et al. (2018). The uncertainty associated with N(H_2) is 20% for each source and the v_{lsr} values are derived from the NH_3 (1, 1) data taken from Wienen et al. (2012); $^{(c)}$Classification scheme from König et al. (2017); $^{(d)}R_{eff}$ and v_{lsr} are from Leurini et al. (2011); v_{lsr} is derived from the $C^{18}O$ $J = 2-1$; $^{(e)}T_{dust}$ standard deviation computed on the clump region defined in Leurini et al. (2019).

TABLE 3.2: Summary of derived properties of the ATLASGAL-sources in our sample.

ATLASGAL-ID	$L/M^{(a)}$ $L_\odot M_\odot^{-1}$	$[f_D]^{(b)}$ $(\chi^{obs}_{C^{17}O}/\chi^{exp}_{C^{17}O})$	N(o-H_2D^+)$^{(c)}$ $\log_{10}(cm^{-2})$	X(o-H_2D^+) $\log_{10}(N[o\text{-}H_2D^+]/N[H_2])$	$\Delta v^{(c)}_{obs}$ (km s^{-1})	$\alpha^{(d)}$	$\mathcal{M}^{(e)}$
G08.71–0.41	0.3	8.2	$12.7^{+0.2}_{-0.2}$	-10.1 ± 0.2	$1.0^{+0.7}_{-0.7}$	0.1	1.9
G13.18+0.06	22.5	6.6	$12.2^{+0.3}_{-0.3}$	-10.6 ± 0.3	$1.4^{+1.2}_{-1.0}$	0.5	1.9
G14.11–0.57 (C1)	9.1	2.1	$12.1^{+0.3}_{-0.3}$	-10.8 ± 0.3	$1.0^{+0.7}_{-0.7}$	0.3	1.3
G14.11–0.57 (C2)	9.1	2.1	$12.3^{+0.3}_{-0.3}$	-10.6 ± 0.3	$1.6^{+1.1}_{-0.9}$	0.9	2.3
G14.49–0.14	0.4	9.0	$13.0^{+0.1}_{-0.1}$	-10.1 ± 0.1	$2.7^{+0.9}_{-0.8}$	0.6	5.4
G14.63–0.58	11.1	2.0	$12.4^{+0.2}_{-0.2}$	-10.6 ± 0.2	$1.7^{+1.0}_{-0.7}$	1.0	2.4
G18.61–0.07	0.7	4.2	$12.6^{+0.1}_{-0.1}$	-10.2 ± 0.1	$1.6^{+0.7}_{-0.6}$	0.5	3.0
G19.88–0.54	15.5	2.2	$12.1^{+0.3}_{-0.3}$	-11.0 ± 0.3	$1.3^{+1.0}_{-0.9}$	0.3	1.7
G28.56–0.24	0.3	6.3	$12.7^{+0.3}_{-0.4}$	-10.4 ± 0.4	$1.0^{+0.6}_{-0.7}$	0.1	1.9
G333.66+0.06	3.0	4.0	$12.6^{+0.3}_{-0.4}$	-10.1 ± 0.4	$1.3^{+1.0}_{-1.0}$	0.2	2.1
G351.57+0.76	2.7	2.4	$12.6^{+0.2}_{-0.2}$	-10.1 ± 0.2	$1.0^{+0.5}_{-0.4}$	0.4	1.5
G354.95–0.54	3.2	3.8	$12.7^{+0.1}_{-0.1}$	-9.9 ± 0.1	$1.2^{+0.4}_{-0.4}$	0.8	1.8
G12.50-0.22	1.3	–	$12.8^{+0.1}_{-0.1}$	-10.0 ± 0.1	$1.2^{+0.4}_{-0.3}$	1.0	2.2
G14.23–0.51	2.1	–	$12.8^{+0.2}_{-0.2}$	-10.5 ± 0.2	$2.1^{+0.8}_{-0.8}$	0.6	3.5
G15.72–0.59	0.2	–	$12.8^{+0.2}_{-0.2}$	-10.0 ± 0.2	$1.9^{+1.0}_{-1.0}$	1.0	3.8
G316.76-0.01	5.2	–	$12.3^{+0.2}_{-0.2}$	-10.8 ± 0.2	$1.6^{+0.6}_{-0.5}$	0.2	2.5
G351.77-CL7	0.7^f	3.4 ^{g}	$12.6^{+0.1}_{-0.1}$	-10.3 ± 0.1	$1.8^{+0.6}_{-0.6}$	0.5	3.5

Derived properties of our sample. Top panel: TOP100 sources; Bottom panel: ATLASGAL sources not in the TOP100. The symbol "–" means that this value is not available in the literature (in the case of f_D) or that it could not be calculated (in the case of the source G14.11-0.57).

$^{(a)}$data from columns (5) and (6) in Table 3.1; $^{(b)}$we consider a 15% of uncertainty for the CO depletion factor, f_D, considering recent results discussed in Sabatini et al. (2019), where there is a detailed CO-depletion study in a local filament of massive star-forming regions; $^{(c)}$ computed using MCweeds ((Giannetti et al., 2017)) assuming $T_{ex} = T_{dust}$; $^{(d)}$ virial parameters derived following MacLaren et al. (1988), as explained in 3.3.3 ; $^{(e)}$Mach numbers derived according to the equation 3.3 in the subsection 3.3.3, where the mass of the H_2D^+ molecule is $m(H_2D^+) = 6,692 \times 10^{-24}$ g; $^{(f)}$this value was calculated from Giannetti et al. (2019) as the integrated values on the APEX-beam at 230 GHz (i.e. 28 arcsec); $^{(g)}$data from Sabatini et al. (2019); explained in 3.3.3.

Chapter 4

Correlations between o-H_2D^+ and physical quantities

This chapter is based on the published work of Sabatini et al. (2020). In order to study the possible presence of trends or correlations in the results between o-H_2D^+ column densities and the observed and derived properties, we apply a Markov Chain Monte-Carlo as a tool of probabilistic inference.

4.1 Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) is an statistical tool composed by two different techniques: *Monte Carlo*, that randomly samples a probability distribution and approximating a desired quantity, and *Markov Chain*, which is a method for generating a sequence of random variables where the current value depends probabilistically on the value of prior variable (Dani Gamerman, 2006). With these two methods together, it is possible to randomly sample probability distribution of high dimensions, with a probabilistic dependence between the samples.
Using MCMC it is possible to make use of Bayesian inference. The Bayes theorem is defined as:

$$p(m,b,f|x,y,\sigma) \propto p(m,b,f) \cdot p(y|x,\sigma,m,b,f) \tag{4.1}$$

being $p(y|x,\sigma,m,b,f)$ the maximum likelihood function, i.e. the probability of a data set given the model parameters and $p(m,b,f)$ the prior function.
The parameters m,b and f are the set of parameters of interest and the parameters x,y and σ are given from the data. Then, $p(m,b,f|x,y,\sigma)$ indicates the posterior or the probability of m,b,f given the data, $p(y|x,\sigma,m,b,f)$ is the likelihood or the probability of the data given m,b,f, and $p(m,b,f)$ indicates the a-priori probability of m,b,f (van Ravenzwaaij et al., 2018).
The maximum likelihood estimation is given by

$$\ln p(y|x,\sigma,m,b,f) = -\frac{1}{2}\sum_n \left[\frac{(y_n - mx_n - b)^2}{s_n^2} + \ln(2\pi s_n^2)\right] \tag{4.2}$$

where

$$s_n^2 = \sigma_n^2 + f^2(mx_n + b)^2 \tag{4.3}$$

This likelihood estimation assumes the error bars as correct, Gaussian and independent, and it provides the estimation of the slope of the fitting m, b that points the origin of the fitting line and f that corresponds to the factor of variance underestimation.

4.1.1 Running the iterations

For our purpose we start by setting the likelihood function as a weighted fit to our sample. MCMC initiates walkers starting inside small Gaussian distributions in the given linear fit and then starts exploring the prior function previously estimated. We run our fitting with 5000 iterations, discarding the first 100 steps. This is because there is a small number of iterations for every MCMC fitting that do not vary much from the given values, this initial iterations are called *flat samples* and can be observed in Fig 4.1. In Fig. 4.2 we show a corner plot, reporting the dimensional projections of the posterior probability distribution of the parameters. In Fig. 4.3 and 4.4 are shown the different fitting performed with MCMC, together with the fitting parameters in each particular panel. Both the trace plot and the corner plot correspond to the run of the N(o-H_2D^+) vs. the virial parameter α.

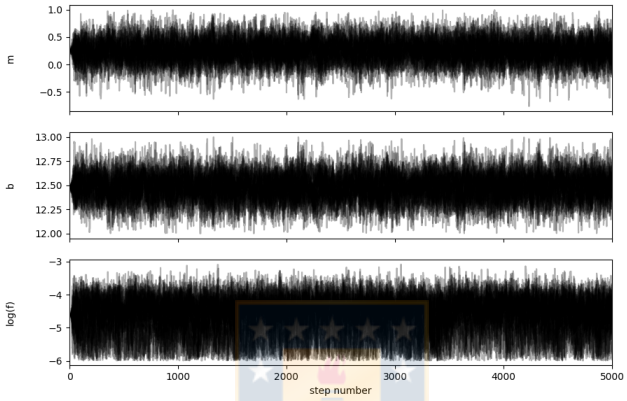

FIGURE 4.1: Trace plots of the MCMC chains. Here are shown the iterations of the three different parameters that build the line approach. Source: This work.

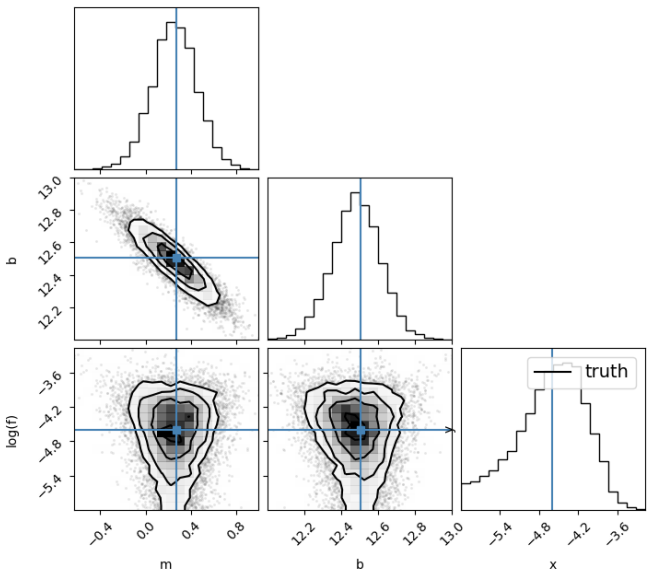

FIGURE 4.2: Corner plot of the MCMC fitting. Here it is shown the one and two dimensional projections of the posterior probability distribution of the given parameters. The blue lines represent where the iterations converge to. Source: This work.

The results of the MCMC fitting applied to our study will be shown in the following section.

4.2 Correlations

The following figures are the correlation plots between the N(o-H_2D^+) and some of the parameters presented in the tables 3.1 and 3.2.

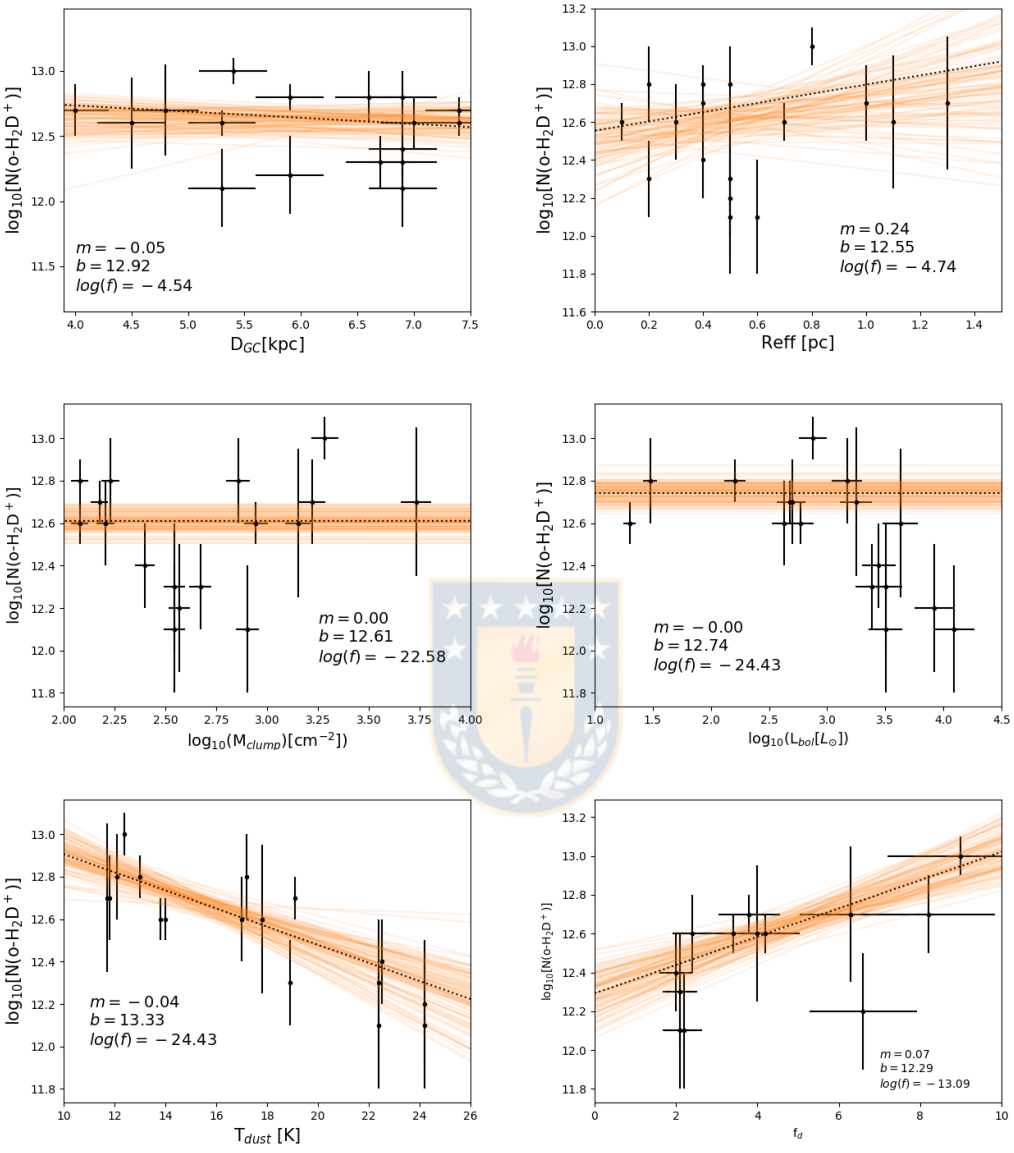

FIGURE 4.3: Different correlations obtained between $N(o - H_2D^+)$ and the observed and derived properties presented in tables 3.1 and 3.2, respectively. In each panel the black dots represent the sources and the error bars represent the uncertainties. The orange regions represent the results of the MCMC linear-fit. The correlation plots are in log-linear scale, except for the mass plot, the luminosity plot and the L/M plot, which are in log-scale. Source: This work.

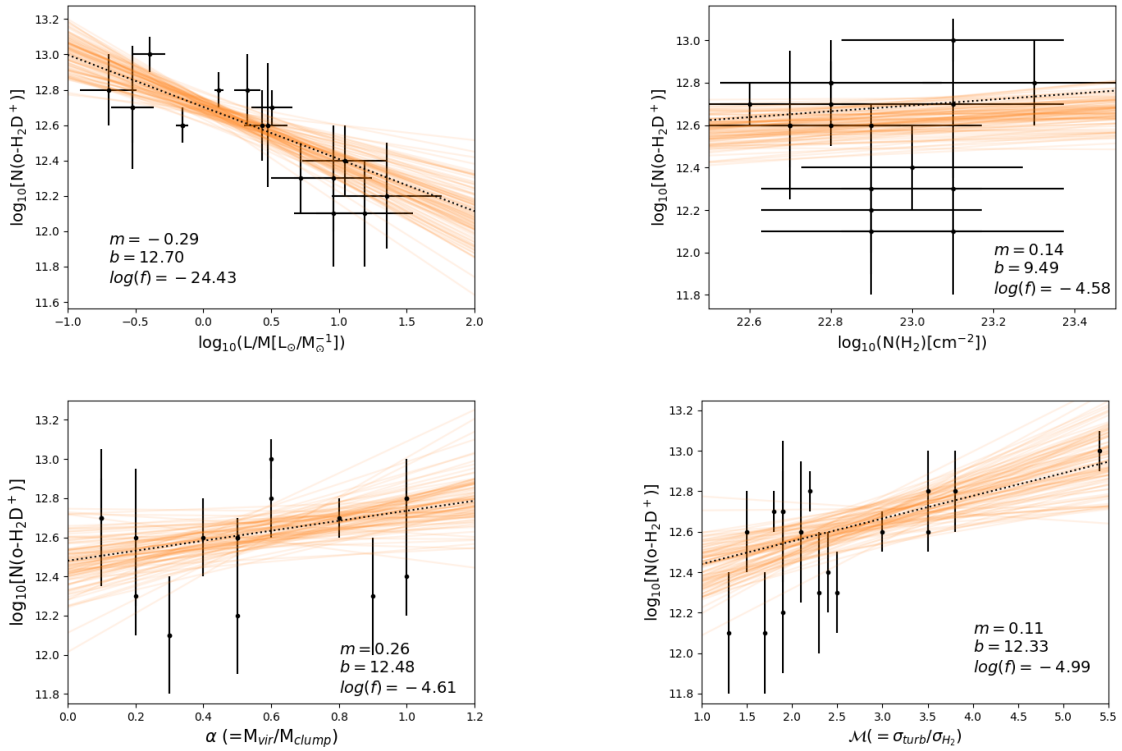

FIGURE 4.4: Second part of the correlations with a MCMC fitting. Source: This work.

The figures 4.3 and 4.4 show the different correlations between N(o-H_2D^+) and the physical parameters that can be found in tables 3.1 and 3.2. In this section we will analyze the different correlations and how they are related with previous studies.

The parameters D_{GC}, R_{eff} and M_{clump} are identified as *source parameters*, examined looking for a relation between the amount of o-H_2D^+ and the characteristics of the structure of the clumps and their distribution on the galactic plane. These parameters are presented in the three first panels of 4.3 where it is observed an uniform distribution of the o-H_2D^+ over time, not showing any notorious change with the evolution of the clumps. This is in agreement with some previous information about the deuterium behaviour: in the galactic disk the variations in the deuterium abundance with respect to the hydrogen, [D/H] are expected, due to the fact that deuteration occurs on the very early stages of the universe and then it has been consuming via thermonuclear reactions that take place inside stellar cores (Ceccarelli et al., 2014; Lubowich et al., 2000). While on the other hand, it is also known that deuterium fractionation is boosted by the star formation process, creating a very unfavouring environment for the deuteration to occur (Bacmann et al., 2003; Caselli, 2002; Lis et al., 2006; Pillai et al., 2012; Ceccarelli et al., 2014; Giannetti et al., 2019). The latter affirmation is consistent with the L_{bol} vs N(o-H_2D^+) plot, where the amount of molecule observed decreases as the luminosity increase.

The parameters identified as *evolutionary parameters* are CO-depletion factor [f_D] (Caselli et al., 1999; Fontani et al., 2012), the dust temperature T_{dust} and the L/M ratio, parameters that highlight trends during the star formation process. All of them are known to be evolutionary tracers, and the three of them show strong trends. The luminosity-to-mass ratio outstands between the evolutionary parameters, since it is the main indicator for the evolutionary stage of the high-mass star formation processes, according to Körtgen et al. (2017); Saraceno et al. (1996); Molinari et al. (2008); Urquhart et al. (2018) and Giannetti et al. (2019).

The *dynamical parameters* are the H_2 column density N(H_2), the mach number $\mathcal{M}$ and the virial parameter α. These parameters are related to the dynamical stage of the clumps. For the H_2 column density there is no trend found, but for the virial parameter α and Mach number $\mathcal{M}$ there are weak trends, not very crucial as it was expected, due their correlation with other parameters that do depend on the amount of N(o-H_2D^+), i.e., FWHM or the T_{dust}.
From the strong correlations found suggest that the deuterium fractionation is driven by a temporal evolution more than a dynamical evolution of the star-forming regions. In the following chapter we will discuss the results presented in this section, and explain the conclusions we reached.

Chapter 5

Discussion and conclusions

In this work we collected more than 10000 spectra in 106 sources of the ATLASGAL survey, with the aim of confirming that o-H_2D^+ can be used as a chemical clock to follow the evolution of the regions where high-mass star forming process takes place, massive clumps. We started working with observations from the APEX Telescope, processing the spectra with a code wrote specifically with the purpouse of filtering, cleaning and cutting the spectra, removing the corrupted observations among them, applying a baseline correction and averaging the multiple observations correspondent to each spectra into a single spectrum with a better signal-to-noise ratio. At the end of this procedure it is possible to identify whether or not a spectrum presents an o-H_2D^+ emission line. We found 16 sources with emission lines greater than 3σ, i.e. three times the average noise of the T_{mb}, becoming the largest sample of o-H_2D+ detected in high-mass star forming regions ever found to this date. From the obtained spectra we calculated the column densities by fitting line profiles using MCWeeds, assuming the local thermodynamical equilibrium. From those results we obtained the most important parameters of each clumps that are summarized in Tables 3.1 and 3.2.

5.1 Discussion

Since our study was motivated by the results presented in Giannetti et al. (2019), it is necessary to make a comparison between them and the results obtained in this work. In Fig. 5.1, presented in Sabatini et al. (2020) it is shown the abundance of o-H_2D^+ as a function of the evolutionary class of the TOP100 clumps (represented with red markers), while the work from Sabatini et al. (2020) (which is derived from the analysis presented in this thesis) is reported in blue. The same trend of o-H_2D^+ can be observed with the evolution of the clumps. The X(o-H_2D^+) abundance decreases by more than an order of magnitude when going from the least to a the most evolved sources. From this it is suggested that the deuteration of H_3^+ is enhanced under the specific conditions of the initial phases of the star formation process, showing that as the clump evolves and the environment becomes warmer the levels of o-H_2D^+ decrease.

The trend shown in Fig. 5.1 is similar to the result shown in Fig. 4.3, in the N(o-H_2D^+) vs. L/M panel, where the luminosity-to-mass ratio increases with the evolution of the clump (Saraceno et al., 1996) while the N(o-H_2D^+) decreases. This can be also observed in the N(o-H_2D^+) vs. L_{bol} panel, with the bolometric luminosity working as a tracer for the most evolved clumps. This is another indication that when a protostellar object is formed and the temperature increases, the deuterated species start to be destroyed triggered by the evaporation of CO back to gas-phase.

In addition, in a recent study where three prestellar and three protostellar low-mass cores where targeted in the Orion-B9 Miettinen (2020) similar results were presented, particularly for the less evolved sources, 70w and IRw. On the other hand, the IRb sources mapped in that study show X(o-H_2D^+) $\sim$ 5 times greater that the limits provided by Giannetti et al. (2019). The discrepancy between

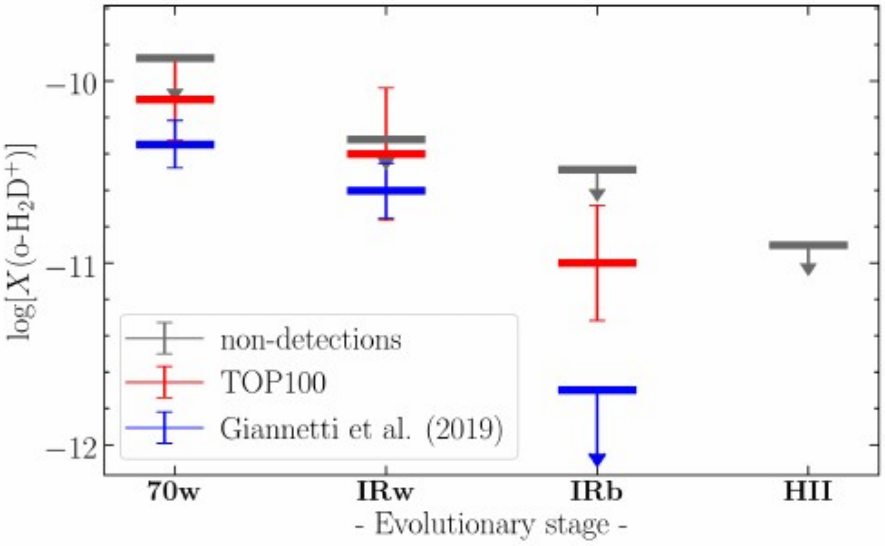

FIGURE 5.1: In this figure it is shown the average $X(o - H_2D^+)$, as a function of the evolutionary classes in the sample. Blue markers represent the recent estimates of Giannetti et al. (2019), the red ones represent the TOP100 sources and grey ones are the non-detection sources from the sample selected. Figure taken from Sabatini et al. (2020).

these two results can be attributed to the physical and chemical initial conditions of the environment of the clumps studied.

5.1.1 o-H_2D^+ as a chemical clock

The chemistry of o-H_2D+ is highly sensitive to the changes of density and temperature during the cloud contraction, in the massive star formation scenario. This is why it is thought that this deuterated molecule is considered as a potential chemical clock of these regions. The formation of H_3^+ in star-forming regions starts with interactions driven by cosmic rays (CR), starting from molecular hydrogen H_2 being ionized by a CR particle producing a reaction exchange, as follows:

$$H_2 + \text{c.r.} \leftarrow H_2^+ + e^+, \tag{5.1}$$

$$H_2^+ + H_2 \leftarrow H_3^+ + H \tag{5.2}$$

Then, the deuteration of H_3^+ produces H_2D^+ following this proton transfer reaction:

$$H_3^+ + HD \rightleftharpoons H_2D^+ + H_2 + \Delta E \tag{5.3}$$

In the case of the massive star formation environment, with temperatures lower than 20 K there is a trend to form more likely the *ortho* state of the H_2D^+ (Pagani et al., 1992). As was already presented from our results, the temperature of the clumps is strongly related to the evolution of them and the behaviour of o-H_2D^+ in the different evolutionary stages. In Urquhart et al. (2018) it was proposed that there is a strong correlation that link the gas temperature of the clumps to their embedded massive protostars luminosity of to the L/M ratio or each clump, which was also proposed in König et al. (2017), where it was found that there notorious difference between the temperatures associated to the different evolutive classes, with temperature changes from 10 K to 40 K. In the results obtained

from this study we confirmed the relation between the abundance of this molecule and T_{dust}, where the highest abundance of o-H_2D^+ are found in the coldest environments, and tending to descend with the temperature raising.
The CO depletion factor is also strongly dependent on the temperature, since when there is more depletion factor, i.e., less levels of CO observed as temperature decreases. This is supported by the cold observed clouds that show low concentrations of CO in their gas phase found in Roberts and Millar (2000), Bacmann et al. (2003) and Ceccarelli et al. (2014) and also in low mass star formation clouds by Caselli et al. (1998) and Bacmann et al. (2002). High-mass star forming regions showed decreasing f_d when L/M was increasing, and with this also confirming that the temperature of the clumps also increases over time, as proposed in Giannetti et al. (2017) and Giannetti et al. (2014).
The formation of H_2D^+ in CO-depleted environments can be more efficient by two events. First, the H_3^+ reservoir is held for reaction (5.3) making the following proton-transfer reaction slows down:

$$H_3^+ + CO \rightarrow HCO^+ + H_2. \tag{5.4}$$

The second one is provoked by the same effect, preventing H_2D^+ from reacting with CO and on a smaller rate N_2, that usually depletes slowly, to form DCO^+ and N_2D^+, remaining observable in the gas phase. The N(o-H_2D^+) vs. f_D correlation shown in Fig. 4.3 is in concordance with this, since we can observe at the highest N(o-H_2D^+) levels we also have the highest f_D values. In the case of the T_{dust} of the clumps it is observed that for larger N(o-H_2D^+) the temperature is lower, which is expected in this dense and cold environments where CO-depletion is boosted (e.g. Kramer et al. 1999; Caselli et al. 1999; Crapsi et al. 2005; Fontani et al. 2012; Wiles et al. 2016; Sabatini et al. 2019).
In our results we found no evident trend α and $\mathcal{M}$ over the evolution in time of the clumps of TOP100 sources studied. For all the clumps of our sample the virial parameter $\alpha \lesssim 1$, indicating that they are dynamical unstable and will collapse fast. This is supported by the results from (Körtgen et al., 2017, 2018) where the found that dynamical parameters (Mach number, magnetic field magnitude distribution and cloud surface density) affect the results (o-H_2D^+ levels) but not as much as the time evolution of the clumps. This results leads us to interpretate that the plots of mach number and virial parameters, where there are no correlation found, are not necessarily a signal that these dynamical parameters are not empirically correlated.

From the results we have presented in this work it is clear that there are still many questions surrounding the observation of o-H_2D^+ is related to the dynamical characteristics of the clumps, and their evolution over time. Also, it is possible that the H_2D^+ velocity dispersion can be not enough representative of the clump dynamics, since in the center of the clump it is replaced with by N_2D^+ or D_2H^+ or destroyed by recently desorbed CO.
From our results we conclude that the o-H_2D^+ behaviour is strongly linked with the physical features of every different phase of the clump evolution. They are related as follows: We found high o-H_2D^+ abundances in the cold starless or prestellar stages, 70w and IRw. Then, a decrease as the protostellar objects evolve reaching to the HII regions. The latter is particularly confirmed by our results where it is possible to see that deuteration proceeds faster in the first stages, due the higher CO freeze-out degree, and with the increase of temperature CO returns to the gas phase due the formed YSO.

5.2 Conclusions

From the correlation plots presented in Fig. 4.3 and Fig. 4.4, we reached the following conclusions:

- It is confirmed that in the high-mass regime of star formation the correlation between N(o-H_2D^+) and $T_{\rm dust}$ holds.
- New correlations were found between N(o-H_2D^+) and the evolutionary parameters $L_{\rm bol}$ and L/M.
- Correlations were found with the parameters related to the evolutionary stages (i.e. f_d, T_{dust} and L/M).
- It is confirmed that N(o-H_2D^+) shows a trend diminishing the amount of o-H_2D^+ with the evolution of the clump as it was proposed by Giannetti et al. (2019), also extending it to the most advanced phase, i.e., HII regions.
- The correlation between N(o-H_2D^+) and the dynamical parameters is not as evident compared to the other parameters.
- H_3^+ deuteration appears to be driven more by thermal than by the dynamical evolution of high-mass star forming regions, as proposed by Bovino et al. (2019) and Körtgen et al. (2017), but a certain degree of degeneracy is expected in cases of slow collapse where there is enough time to reach high deuteration levels.

With the fact that we can observe the same behaviour in Giannetti et al. (2019), Miettinen (2020) and in our work, it is possible to confirm that the evolutionary sequence of massive star forming regions goes along with a downward trend for o-H_2D^+, providing the scientific community with a new observational tool that will help to narrow down the questions that are still surrounding the massive star formation scenario.

5.3 Outlook

This work can be continued applying the same study to observations of N_2D^+, in order to confirm that this molecule is inversely proportional to o-H_2D^+ and shows an increasing trend that goes along with the evolution of the clumps, as it was also proposed by Giannetti et al. (2019) in the study of the three massive clumps embedded in G351.77-0.51.
Another possible complement to this study is performing synthetic observations from simulations to be compared with the clumps belonging to our sample, providing a more complete scenario on the study of deuterated molecules.

Appendix A

Basic concepts of Radioastronomy

For the proper understanding of the APEX telescope, it is necessary to review some of the most important radioastronomy concepts. Most of the following explanations are based on "Tools of Radio Astronomy" (Wilson et al., 2009).
First, the infinitesimal power dP that is received by an infinitesimal surface dσ is defined as

$$\mathrm{d}P = I_\nu cos\theta d\Omega d\sigma d\nu \tag{A.1}$$

where dP is the infinitesimal power [10^7 erg s^{-1}], dσ is the infinitesimal area of surface [cm^2], dν is the infinitesimal bandwidth [Hz], θ is the angle between a normal vector to dσ and the direction to dΩ and I_ν is the brightness or intensity [10^{11} erg s^{-1} cm^{-2} Hz^{-1} sr^{-1}].
The brightness that the antenna will receive, depends on the direction at where the source is located, making B = B(θ, ϕ).
The main beam solid angle is given by

$$\Omega_\mathrm{A} = \int\int_{4\pi} P_\mathrm{n}(\theta,\phi) d\Omega = \int_0^{2\pi}\int_0^{\pi} P_\mathrm{n}(\theta,\phi) sin\theta d\theta d\phi \tag{A.2}$$

which is measured in steradians (sr). Here, $P_\mathrm{n}(\theta, \phi)$ is the normalized power pattern,

$$P_\mathrm{n}(\theta,\phi) = \frac{1}{P_\mathrm{max}} P(\theta,\phi) \tag{A.3}$$

In the ideal antenna we would have $P_\mathrm{n} = 1$ for all Ω_A and $P_\mathrm{n} = 0$ everywhere else. In the real antennas, most of the power pattern can be found in a region known as main lobe, while the residuals of the signal are received by the side or secondary lobes (see Fig. A.1). Then, the mean beam solid angle is

$$\Omega_\mathrm{MB} = \iint_{main\ lobe} P_n(\theta,\phi) d\Omega \tag{A.4}$$

The amount of signal coming from the sources that is actually received by the main lobe of the antenna is known as *beam efficiency*. In the ideal antenna, all the signal from the source is received by the main lobe, and the beam efficiency is equal to one.
This beam efficiency η_B is defined as

$$\eta_\mathrm{B} = \frac{\Omega_\mathrm{MB}}{\Omega_\mathrm{A}} \tag{A.5}$$

The angular extent of the main beam, i.e. the angle between points of the main beam where the normalized power pattern falls to 1/2 of its maximum. This is called *half power beam width* (HPBW) or also *full width to half power* (FWHP) and can be seen in Fig. A.2.

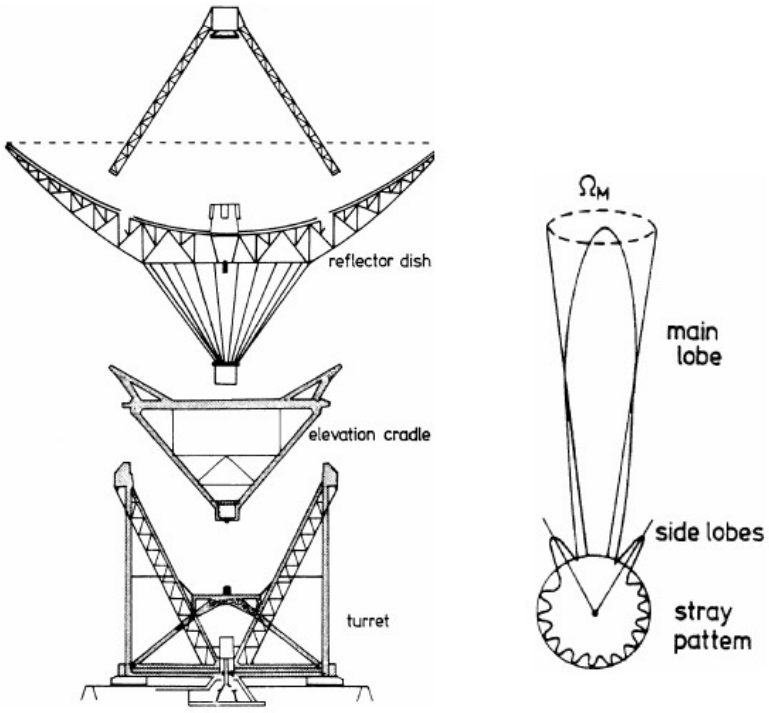

FIGURE A.1: Sketch of the antenna structure on the left, showing the main and more important parts of the structure of a single dish antenna. On the right the representation of the normalized power pattern. In this sketch is also the main lobe, the receiver of most of the signal and the side lobes.

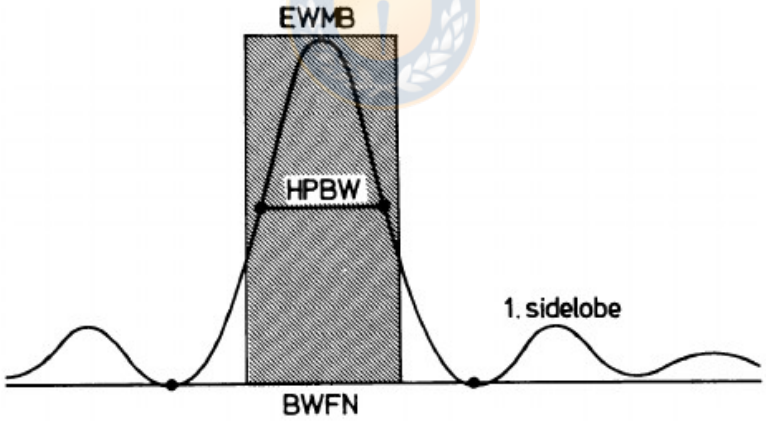

FIGURE A.2: Graphic representation of a single-dish radiotelescope beamwidth. Other measurements to the main lobe are the beam width between first nulls (BWFN) and the equivalent width of the main beam (EWMB). The HPBW, EWMB and BWFN are drawn in the tallest wave of the drawing that represents the main lobe in the signal received by the antenna.

The total power received by an antenna pointing to a source with a brightness distribution of B_ν, is defined as

$$P_\nu = \frac{1}{2} A_\mathrm{e} \iint B_\nu(\theta,\phi) P_\mathrm{n}(\theta,\phi) d\Omega \tag{A.6}$$

Where A_e is the effective aperture, the flat surface sustained by the antenna semi circumference collecting area of the antenna.

The brightness distribution of the antenna can be transformed to a temperature distribution according to the Rayleigh-Jeans law, since we are at low frequencies. Here, using the antenna temperature

T_A, we have

$$P_\nu = kT_A \tag{A.7}$$

and the equation A.6 can be rewritten as

$$T_A(\theta_0, \phi_0) = \frac{\int T_b(\theta, \phi) P_n(\theta - \theta_0, \phi - \phi_0) sin\theta d\theta d\phi}{\int P_n(\theta, \phi) d\Omega} \tag{A.8}$$

which corresponds to the convolution of the brightness temperature with the beam pattern of the telescope. Here $T_b(\theta, \phi)$ is the brightness temperature associated with the thermodynamic energy of the radiating material only for thermal radiation in the Rayleigh-Jeans limit from an optically thick source. For other cases, T_b will depend on the frequency.
Now, the signal received by the antenna consists on a sum of different contributions.

$$T = T_A + T_{cont} + T_N, \tag{A.9}$$

which are the antenna temperature, the continuum contribution and the contribution of the noise, both atmospheric and electronics, respectively. Is important to notice that T_A at this point correspond to an ideal antenna, with no losses and no atmospheric related absorption. We will call T_A^* the real antenna temperature, corrected for atmospheric losses.
The main beam temperature is obtained with the following conversion

$$T_{MB} = T_A^* / \eta_B \tag{A.10}$$

For the completeness of Radioastronomy theory it is recommended to review specific bibliography, i.e., Wilson et al. (2009).

Bibliography

Amano, T., and Hirao, T. (2005). "Accurate rest frequencies of submillimeter-wave lines of H_2D^+ and D_2H^+." *Journal of Molecular Spectroscopy*, *233*(1), 7–14.

Bacmann, A., Lefloch, B., Ceccarelli, C., Castets, A., Steinacker, J., and Loinard, L. (2002). "The degree of CO depletion in pre-stellar cores." , *389*, L6–L10.

Bacmann, A., Lefloch, B., Ceccarelli, C., Steinacker, J., Castets, A., and Loinard, L. (2003). "CO Depletion and Deuterium Fractionation in Prestellar Cores." , *585*(1), L55–L58.

Batten, A. H. (1973). *Binary and multiple systems of stars*.

Bergin, E. A., and Tafalla, M. (2007). "Cold Dark Clouds: The Initial Conditions for Star Formation." , *45*(1), 339–396.

Bertoldi, F., and McKee, C. F. (1992). "Pressure-confined Clumps in Magnetized Molecular Clouds." , *395*, 140.

Bonnell, I. A., Vine, S. G., and Bate, M. R. (2004). "Massive star formation: nurture, not nature." , *349*(2), 735–741.

Bovino, S., Ferrada-Chamorro, S., Lupi, A., Sabatini, G., Giannetti, A., and Schleicher, D. R. G. (2019). "The 3D Structure of CO Depletion in High-mass Prestellar Regions." , *887*(2), 224.

Bovino, S., and Galli, D. (2019). *Thermodynamics and chemistry of the early universe*, 45–66.

Brand, J., and Blitz, L. (1993). "The velocity field of the outer galaxy." , *275*, 67–90.

Burkert, A., and Hartmann, L. (2004). "Collapse and Fragmentation in Finite Sheets." , *616*(1), 288–300.

Carey, S. J., Clark, F. O., Egan, M. P., Price, S. D., Shipman, R. F., and Kuchar, T. A. (1998). "The Physical Properties of the Midcourse Space Experiment Galactic Infrared-dark Clouds." , *508*(2), 721–728.

Caselli, P. (2002). "Deuterated molecules as a probe of ionization fraction in dense interstellar clouds." , *50*(12-13), 1133–1144.

Caselli, P., and Ceccarelli, C. (2012). "Our astrochemical heritage." , *20*, 56.

Caselli, P., van der Tak, F. F. S., Ceccarelli, C., and Bacmann, A. (2003). "Abundant H_2D^+ in the pre-stellar core L1544." , *403*, L37–L41.

Caselli, P., Vastel, C., Ceccarelli, C., van der Tak, F. F. S., Crapsi, A., and Bacmann, A. (2008). "Survey of ortho-H_2D^+ ($1_{1,0}$-$1_{1,1}$) in dense cloud cores." , *492*(3), 703–718.

Caselli, P., Walmsley, C. M., Tafalla, M., Dore, L., and Myers, P. C. (1999). "CO Depletion in the Starless Cloud Core L1544." , *523*(2), L165–L169.

Caselli, P., Walmsley, C. M., Terzieva, R., and Herbst, E. (1998). "The Ionization Fraction in Dense Cloud Cores." , *499*(1), 234–249.

Ceccarelli, C., Caselli, P., Bockelée-Morvan, D., Mousis, O., Pizzarello, S., Robert, F., and Semenov, D. (2014). "Deuterium Fractionation: The Ariadne's Thread from the Precollapse Phase to Meteorites and Comets Today." In H. Beuther, R. S. Klessen, C. P. Dullemond, and T. Henning (Eds.), *Protostars and Planets VI*, 859.

Chandrasekhar, S., and Fermi, E. (1953). "Problems of Gravitational Stability in the Presence of a Magnetic Field." , *118*, 116.

Colombo, D., Meidt, S. E., Schinnerer, E., García-Burillo, S., Hughes, A., Pety, J., Leroy, A. K., Dobbs, C. L., Dumas, G., Thompson, T. A., Schuster, K. F., and Kramer, C. (2014). "The PdBI Arcsecond Whirlpool Survey (PAWS): Multi-phase Cold Gas Kinematic of M51." , *784*(1), 4.

Crapsi, A., Caselli, P., Walmsley, C. M., Myers, P. C., Tafalla, M., Lee, C. W., and Bourke, T. L. (2005). "Probing the Evolutionary Status of Starless Cores through N_2H^+ and N_2D^+ Observations." , *619*, 379–406.

Crutcher, R. M. (2012). "Magnetic Fields in Molecular Clouds." , *50*, 29–63.

Dame, T. M., Hartmann, D., and Thaddeus, P. (2001). "The Milky Way in Molecular Clouds: A New Complete CO Survey." , *547*(2), 792–813.

Dani Gamerman, H. F. L. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Texts in Statistical Science, Chapman and Hall/CRC, 2 edn.

de Wit, W. J., Testi, L., Palla, F., Vanzi, L., and Zinnecker, H. (2004). "The origin of massive O-type field stars. I. A search for clusters." , *425*, 937–948.

Evans, I., Neal J. (1991). "Star Formation - Observations." In D. L. Lambert (Ed.), *Frontiers of Stellar Evolution*, *Astronomical Society of the Pacific Conference Series*, vol. 20, 45.

Fontani, F., Giannetti, A., Beltrán, M. T., Dodson, R., Rioja, M., Brand, J., Caselli, P., and Cesaroni, R. (2012). "High CO depletion in southern infrared dark clouds." , *423*(3), 2342–2358.

Friesen, R. K., Di Francesco, J., Bourke, T. L., Caselli, P., Jørgensen, J. K., Pineda, J. E., and Wong, M. (2014). "Revealing H_2D^+ Depletion and Compact Structure in Starless and Protostellar Cores with ALMA." , *797*(1), 27.

Friesen, R. K., Di Francesco, J., Myers, P. C., Belloche, A., Shirley, Y. L., Bourke, T. L., and André, P. (2010). "The Initial Conditions of Clustered Star Formation. III. The Deuterium Fractionation of the Ophiuchus B2 Core." , *718*(2), 666–682.

Giannetti, A., Bovino, S., Caselli, P., Leurini, S., Schleicher, D. R. G., Körtgen, B., Menten, K. M., Pillai, T., and Wyrowski, F. (2019). "A timeline for massive star-forming regions via combined observation of o-H_2D^+ and N_2D^+." , *621*, L7.

Giannetti, A., Leurini, S., Wyrowski, F., Urquhart, J., Csengeri, T., Menten, K. M., König, C., and Güsten, R. (2017). "ATLASGAL-selected massive clumps in the inner Galaxy. V. Temperature structure and evolution." , *603*, A33.

Giannetti, A., Wyrowski, F., Brand, J., Csengeri, T., Fontani, F., Walmsley, C. M., Nguyen Luong, Q., Beuther, H., Schuller, F., Guesten, R., and Menten, K. M. (2014). "VizieR Online Data Catalog: ATLASGAL massive clumps CO depletion (Giannetti+, 2014)." *VizieR Online Data Catalog*, J/A+A/570/A65.

Girichidis, P., Federrath, C., Banerjee, R., and Klessen, R. S. (2011). "Importance of the initial conditions for star formation – I. Cloud evolution and morphology." *Monthly Notices of the Royal Astronomical Society*, *413*(4), 2741–2759.

Girichidis, P., Federrath, C., Banerjee, R., and Klessen, R. S. (2012). "Importance of the initial conditions for star formation - II. Fragmentation-induced starvation and accretion shielding." , *420*(1), 613–626.

Güsten, R., Nyman, L. Å., Schilke, P., Menten, K., Cesarsky, C., and Booth, R. (2006). "The Atacama Pathfinder EXperiment (APEX) - a new submillimeter facility for southern skies -." , *454*(2), L13–L16.

Harju, J., Haikala, L. K., Lehtinen, K., Juvela, M., Mattila, K., Miettinen, O., Dumke, M., Güsten, R., and Nyman, L. Å. (2006). "Detection of H_2D^+ in a massive prestellar core in Orion B." , *454*(2), L55–L58.

Harju, J., Juvela, M., Schlemmer, S., Haikala, L. K., Lehtinen, K., and Mattila, K. (2008). "Detection of 6 K gas in Ophiuchus D." , *482*(2), 535–539.

Hartmann, L. (2002). "Flows, Fragmentation, and Star Formation. I. Low-Mass Stars in Taurus." , *578*(2), 914–924.

Heyminck, S., Kasemann, C., Güsten, R., de Lange, G., and Graf, U. U. (2006). "The first-light APEX submillimeter heterodyne instrument FLASH." , *454*(2), L21–L24.

Ho, P. T. P., and Haschick, A. D. (1981). "Formation of OB clusters: VLA observations." , *248*, 622–637.

Hoare, M. G., Kurtz, S. E., Lizano, S., Keto, E., and Hofner, P. (2007). "Ultracompact Hii Regions and the Early Lives of Massive Stars." In B. Reipurth, D. Jewitt, and K. Keil (Eds.), *Protostars and Planets V*, 181.

Hoare, M. G., Purcell, C. R., Churchwell, E. B., Diamond, P., Cotton, W. D., Chandler, C. J., Smethurst, S., Kurtz, S. E., Mundy, L. G., Dougherty, S. M., Fender, R. P., Fuller, G. A., Jackson, J. M., Garrington, S. T., Gledhill, T. R., Goldsmith, P. F., Lumsden, S. L., Martí, J., Moore, T. J. T., Muxlow, T. W. B., Oudmaijer, R. D., Pandian, J. D., Paredes, J. M., Shepherd, D. S., Spencer, R. E., Thompson, M. A., Umana, G., Urquhart, J. S., and Zijlstra, A. A. (2012). "The Coordinated Radio and Infrared Survey for High-Mass Star Formation (The CORNISH Survey). I. Survey Design." , *124*(919), 939.

Kauffmann, J., Pillai, T., and Goldsmith, P. F. (2013). "Low Virial Parameters in Molecular Clouds: Implications for High-mass Star Formation and Magnetic Fields." , *779*(2), 185.

Khersonkii, V., Varshalovich, D., and Opendak, M. (1987). "Rotational transition of the interstellar h2d+ molecule." *Soviet Astronomy*, *31*, 274.

Klessen, R. S., and Glover, S. C. O. (2016). "Physical Processes in the Interstellar Medium." *Saas-Fee Advanced Course*, *43*, 85.

Kolmogorov, A. (1941). "The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds' Numbers." *Akademiia Nauk SSSR Doklady*, *30*, 301–305.

König, C., Urquhart, J. S., Csengeri, T., Leurini, S., Wyrowski, F., Giannetti, A., Wienen, M., Pillai, T., Kauffmann, J., Menten, K. M., and Schuller, F. (2017). "ATLASGAL-selected massive clumps in the inner Galaxy. III. Dust continuum characterization of an evolutionary sample." , *599*, A139.

Körtgen, B., Bovino, S., Schleicher, D. R. G., Giannetti, A., and Banerjee, R. (2017). "Deuterium fractionation and H_2D^+ evolution in turbulent and magnetized cloud cores." , *469*(3), 2602–2625.

Körtgen, B., Bovino, S., Schleicher, D. R. G., Stutz, A., Banerjee, R., Giannetti, A., and Leurini, S. (2018). "Fast deuterium fractionation in magnetized and turbulent filaments." , *478*, 95–109.

Kramer, C., Alves, J., Lada, C. J., Lada, E. A., Sievers, A., Ungerechts, H., and Walmsley, C. M. (1999). "Depletion of CO in a cold dense cloud core of IC 5146." , *342*, 257–270.

Kroupa, P. (2001). "On the variation of the initial mass function." , *322*(2), 231–246.

Krumholz, M. R., and Bonnell, I. A. (2009). *Models for the formation of massive stars*, 288–320.

Lada, C. J. (2006). "Stellar Multiplicity and the Initial Mass Function: Most Stars Are Single." , *640*(1), L63–L66.

Lada, C. J., and Lada, E. A. (2003). "Embedded Clusters in Molecular Clouds." , *41*, 57–115.

Larson, R. B. (1998). "Early star formation and the evolution of the stellar initial mass function in galaxies." , *301*(2), 569–581.

Lee, ., M.-Y. (et al., in prep.).

Leurini, S., Pillai, T., Stanke, T., Wyrowski, F., Testi, L., Schuller, F., Menten, K. M., and Thorwirth, S. (2011). "The molecular distribution of the IRDC G351.77-0.51." , *533*, A85.

Leurini, S., Schisano, E., Pillai, T., Giannetti, A., Urquhart, J., Csengeri, T., Casu, S., Cunningham, M., Elia, D., Jones, P. A., König, C., Molinari, S., Stanke, T., Testi, L., Wyrowski, F., and Menten, K. M. (2019). "Characterising the high-mass star forming filament G351.776-0.527 with Herschel and APEX dust continuum and gas observations." , *621*, A130.

Lis, D. C., Gerin, M., Roueff, E., Vastel, C., and Phillips, T. G. (2006). "Ground State Rotational Lines of Doubly Deuterated Ammonia as Tracers of the Physical Conditions and Chemistry of Cold Interstellar Medium." , *636*(2), 916–922.

Lubowich, D. A., Pasachoff, J. M., Balonek, T. J., Millar, T. J., Tremonti, C., Roberts, H., and Galloway, R. P. (2000). "Deuterium in the Galactic Centre as a result of recent infall of low-metallicity gas." , *405*(6790), 1025–1027.

MacLaren, I., Richardson, K. M., and Wolfendale, A. W. (1988). "Corrections to Virial Estimates of Molecular Cloud Masses." , *333*, 821.

Maret, S., Hily-Blant, P., Pety, J., Bardeau, S., and Reynier, E. (2011). "Weeds: a CLASS extension for the analysis of millimeter and sub-millimeter spectral surveys." , *526*, A47.

Miettinen, O. (2020). "APEX observations of ortho-H_2D^+ towards dense cores in the Orion B9 filament." , *634*, A115.

Molinari, S., Pezzuto, S., Cesaroni, R., Brand, J., Faustini, F., and Testi, L. (2008). "The evolution of the spectral energy distribution in massive young stellar objects." , *481*(2), 345–365.

Moore, T. J. T., Urquhart, J. S., Morgan, L. K., and Thompson, M. A. (2012). "The effect of spiral arms on star formation in the Galaxy." , *426*(1), 701–707.

Motte, F., Bontemps, S., and Louvet, F. (2018). "High-Mass Star and Massive Cluster Formation in the Milky Way." , *56*, 41–82.

Mouschovias, T. C. (1979). "Ambipolar diffusion in interstellar clouds: a new solution." , *228*, 475–481.

Murray, N. (2011). "Star Formation Efficiencies and Lifetimes of Giant Molecular Clouds in the Milky Way." , *729*(2), 133.

Nakano, T., and Nakamura, T. (1978). "Gravitational Instability of Magnetized Gaseous Disks 6." , *30*, 671–680.

Pagani, L., Salez, M., and Wannier, P. G. (1992). "The chemistry of H2D+ in cold clouds." , *258*, 479–488.

Patil, A., Huard, D., and Fonnesbeck, C. (2010). "Pymc: Bayesian stochastic modelling in python." *Journal of Statistical Software, Articles*, *35*(4), 1–81.

Peretto, N., André, P., and Belloche, A. (2006). "Probing the formation of intermediate- to high-mass stars in protoclusters. A detailed millimeter study of the NGC 2264 clumps." , *445*(3), 979–998.

Pillai, T., Caselli, P., Kauffmann, J., Zhang, Q., Thompson, M. A., and Lis, D. C. (2012). "H_2D^+ in the High-mass Star-forming Region Cygnus X." , *751*(2), 135.

Price, N. M., and Podsiadlowski, P. (1995). "Dynamical interactions between young stellar objects and a collisional model for the origin of the stellar mass spectrum." , *273*(4), 1041–1068.

Punanova, A., Caselli, P., Pineda, J. E., Pon, A., Tafalla, M., Hacar, A., and Bizzocchi, L. (2018). "Kinematics of dense gas in the l1495 filament." *Astronomy Astrophysics*, *617*, A27.

Purcell, C. R., Hoare, M. G., Cotton, W. D., Lumsden, S. L., Urquhart, J. S., Chandler, C., Churchwell, E. B., Diamond, P., Dougherty, S. M., Fender, R. P., Fuller, G., Garrington, S. T., Gledhill, T. M., Goldsmith, P. F., Hindson, L., Jackson, J. M., Kurtz, S. E., Marti, J., Moore, T. J. T., Mundy, L. G., Muxlow, T. W. B., Oudmaijer, R. D., Pandian, J. D., Paredes, J. M., Shepherd, D. S., Smethurst, S., Spencer, R. E., Thompson, M. A., Umana, G., and Zijlstra, A. A. (2013). "VizieR Online Data Catalog: CORNISH project. II. Source catalog (Purcell+, 2013)." *VizieR Online Data Catalog*, J/ApJS/205/1.

Ragan, S. E., Bergin, E. A., Plume, R., Gibson, D. L., Wilner, D. J., O'Brien, S., and Hails, E. (2006). "Molecular Line Observations of Infrared Dark Clouds: Seeking the Precursors to Intermediate and Massive Star Formation." , *166*(2), 567–584.

Rathborne, J. M., Jackson, J. M., and Simon, R. (2006). "Infrared Dark Clouds: Precursors to Star Clusters." , *641*(1), 389–405.

Redaelli, E., Bizzocchi, L., Caselli, P., Sipilä, O., Lattanzi, V., Giuliano, B. M., and Spezzano, S. (2019). "High-sensitivity maps of molecular ions in l1544." *Astronomy Astrophysics*, *629*, A15.

Reid, M. J., Menten, K. M., Brunthaler, A., Zheng, X. W., Dame, T. M., Xu, Y., Wu, Y., Zhang, B., Sanna, A., Sato, M., Hachisuka, K., Choi, Y. K., Immer, K., Moscadelli, L., Rygl, K. L. J., and Bartkiewicz, A. (2014). "Trigonometric Parallaxes of High Mass Star Forming Regions: The Structure and Kinematics of the Milky Way." , *783*(2), 130.

Roberts, H., and Millar, T. J. (2000). "Modelling of deuterium chemistry and its application to molecular clouds." , *361*, 388–398.

Roman-Duval, J., Jackson, J. M., Heyer, M., Rathborne, J., and Simon, R. (2010). "Physical Properties and Galactic Distribution of Molecular Clouds Identified in the Galactic Ring Survey." , *723*(1), 492–507.

Sabatini, G., Bovino, S., Giannetti, A., Wyrowski, F., Órdenes, M. A., Pascale, R., Pillai, T., Wienen, M., Csengeri, T., and Menten, K. M. (2020). "Survey of ortho-H_2D^+ in high-mass star-forming regions." , *644*, A34.

Sabatini, G., Giannetti, A., Bovino, S., Brand, J., Leurini, S., Schisano, E., Pillai, T., and Menten, K. M. (2019). "On the size of the CO-depletion radius in the IRDC G351.77-0.51." , *490*(4), 4489–4501.

Saraceno, P., Andre, P., Ceccarelli, C., Griffin, M., and Molinari, S. (1996). "An evolutionary diagram for young stellar objects." , *309*, 827–839.

Schuller, F., Menten, K. M., Contreras, Y., Wyrowski, F., Schilke, P., Bronfman, L., Henning, T., Walmsley, C. M., Beuther, H., Bontemps, S., Cesaroni, R., Deharveng, L., Garay, G., Herpin, F., Lefloch, B., Linz, H., Mardones, D., Minier, V., Molinari, S., Motte, F., Nyman, L. Å., Reveret, V., Risacher, C., Russeil, D., Schneider, N., Testi, L., Troost, T., Vasyunina, T., Wienen, M., Zavagno, A., Kovacs, A., Kreysa, E., Siringo, G., and Weiß, A. (2009). "ATLASGAL - The APEX telescope large area survey of the galaxy at 870 μm." , *504*(2), 415–427.

Shirley, Y. L. (2015). "The Critical Density and the Effective Excitation Density of Commonly Observed Molecular Dense Gas Tracers." , *127*(949), 299.

Shu, F. H., Adams, F. C., and Lizano, S. (1987). "Star formation in molecular clouds: observation and theory." , *25*, 23–81.

Sipilä, O., Hugo, E., Harju, J., Asvany, O., Juvela, M., and Schlemmer, S. (2010). "Modelling line emission of deuterated H_3^+ from prestellar cores." , *509*, A98.

Smith, I. W. M., Cockell, C. S., and Sydney, L. (2013). *Astrochemistry and Astrobiology*. Springer-Verlag Berlin Heidelberg.

Sochi, T., and Tennyson, J. (2010). "A computed line list for the H_2D^+ molecular ion." , *405*(4), 2345–2350.

Sridharan, T. K., Beuther, H., Saito, M., Wyrowski, F., and Schilke, P. (2005). "High-Mass Starless Cores." , *634*(1), L57–L60.

Swift, J. J. (2009). "Two Massive, Low-Luminosity Cores Toward Infrared Dark Clouds." , *705*(2), 1456–1461.

Tachihara, K., Onishi, T., Mizuno, A., and Fukui, Y. (2002). "Statistical study of $C^{18}O$ dense cloud cores and star formation." , *385*, 909–920.

Teyssier, D., Hennebelle, P., and Pérault, M. (2002). "Radio-millimetre investigation of galactic infrared dark clouds." , *382*, 624–638.

Tielens, A. G. G. M. (2005). *The Physics and Chemistry of the Interstellar Medium*. Cambridge University Press.

Urquhart, J. S., Busfield, A. L., Hoare, M. G., Lumsden, S. L., Oudmaijer, R. D., Moore, T. J. T., Gibb, A. G., Purcell, C. R., Burton, M. G., and Marechal, L. J. L. (2007). "VizieR Online Data Catalog: RMS 13CO obs. of YSOs in South Galactic plane (Urquhart+, 2007)." *VizieR Online Data Catalog*, J/A+A/474/891.

Urquhart, J. S., Hoare, M. G., Lumsden, S. L., Oudmaijer, R. D., Moore, T. J. T., Brook, P. R., Mottram, J. C., Davies, B., and Stead, J. J. (2009). "The RMS survey. H_2O masers towards a sample of southern hemisphere massive YSO candidates and ultra compact HII regions." , *507*(2), 795–802.

Urquhart, J. S., König, C., Giannetti, A., Leurini, S., Moore, T. J. T., Eden, D. J., Pillai, T., Thompson, M. A., Braiding, C., Burton, M. G., Csengeri, T., Dempsey, J. T., Figura, C., Froebrich, D., Menten, K. M., Schuller, F., Smith, M. D., and Wyrowski, F. (2018). "ATLASGAL - properties of a complete sample of Galactic clumps." , *473*(1), 1059–1102.

Urquhart, J. S., Moore, T. J. T., Csengeri, T., Wyrowski, F., Schuller, F., Hoare, M. G., Lumsden, S. L., Mottram, J. C., Thompson, M. A., Menten, K. M., Walmsley, C. M., Bronfman, L., Pfalzner, S., König, C., and Wienen, M. (2014). "ATLASGAL - towards a complete sample of massive star forming clumps." , *443*(2), 1555–1586.

Urquhart, J. S., Moore, T. J. T., Menten, K. M., König, C., Wyrowski, F., Thompson, M. A., Csengeri, T., Leurini, S., and Eden, D. J. (2015). "The almost ubiquitous association of 6.7-GHz methanol masers with dust." , *446*(4), 3461–3477.

van Ravenzwaaij, D., Cassey, P., and Brown, S. (2018). "A simple introduction to Markov Chain Monte–Carlo sampling." *Psychonomic Bulletin Review*, *25*, 143–154.

Vastel, C., Caselli, P., Ceccarelli, C., Bacmann, A., Lis, D. C., Caux, E., Codella, C., Beckwith, J. A., and Ridley, T. (2012). "Upper limit for the D_2H^+ ortho-to-para ratio in the prestellar core 16293E (CHESS)." , *547*, A33.

Vastel, C., Caselli, P., Ceccarelli, C., Phillips, T., Wiedner, M. C., Peng, R., Houde, M., and Dominik, C. (2006). "The Distribution of Ortho-H_2D^+($1_{1,0}$-$1_{1,1}$) in L1544: Tracing the Deuteration Factory in Prestellar Cores." , *645*(2), 1198–1211.

Vázquez-Semadeni, E., Gómez, G. C., Jappsen, A. K., Ballesteros-Paredes, J., and Klessen, R. S. (2009). "High- and Low-Mass Star-Forming Regions from Hierarchical Gravitational Fragmentation. High Local Star Formation Rates with Low Global Efficiencies." , *707*(2), 1023–1033.

Wienen, M., Wyrowski, F., Schuller, F., Menten, K. M., Walmsley, C. M., Bronfman, L., and Motte, F. (2012). "Ammonia from cold high-mass clumps discovered in the inner Galactic disk by the ATLASGAL survey." , *544*, A146.

Wienen, M., Wyrowski, F., Walmsley, C. M., Csengeri, T., Pillai, T., Giannetti, A., and Menten, K. M. (2021). "ATLASGAL-selected massive clumps in the inner Galaxy. IX. Deuteration of ammonia." , *649*, A21.

Wiles, B., Lo, N., Redman, M. P., Cunningham, M. R., Jones, P. A., Burton, M. G., and Bronfman, L. (2016). "Scaled up low-mass star formation in massive star-forming cores in the g333 giant molecular cloud." *MNRAS*, *458*(4), 3429–3442.

Williams, J. P., Blitz, L., and McKee, C. F. (2000). "The Structure and Evolution of Molecular Clouds: from Clumps to Cores to the IMF." In V. Mannings, A. P. Boss, and S. S. Russell (Eds.), *Protostars and Planets IV*, 97.

Wilson, T. L., Rohlfs, K., and Hüttemeister, S. (2009). *Tools of Radio Astronomy*.

Wyrowski, F., Güsten, R., Menten, K. M., Wiesemeyer, H., Csengeri, T., Heyminck, S., Klein, B., König, C., and Urquhart, J. S. (2016). "Infall through the evolution of high-mass star-forming clumps." , *585*, A149.