Alma Mater Studiorum · Università di Bologna

School of Science
Department of Physics and Astronomy
Master's Degree in Astrophysics and Cosmology

THE SEARCH FOR CHEMICAL COMPLEXITY IN PROTOPLANETARY DISKS:

ALMA Wideband Preview of HL Tauri

Candidate:
Davide Napolitano

Supervisors:
Prof. Leonardo Testi
Dr. Anna Miotello

Academic Year 2024/2025

"Astonishment and wonder are what we feel on encountering something that differs from what is normal, or at least from what is for some reason or other expected. But this whole world is something we encounter only once. We have nothing with which to compare it, and it is impossible to see how we can approach it with any particular expectation. And yet we are astonished; we are puzzled by what we find, yet are unable to say what we should have to have found in order not to be surprised, or how the world would have to have been constructed in order not to constitute a riddle."

Erwin Schrödinger, *My view of the world*

ABSTRACT

Planets form in the so-called protoplanetary disks in the first few Myr since the formation of stars. The study of the chemical content of protoplanetary disks is thus essential to understand the pristine composition of planetary atmospheres and surfaces, and the chemical budget available for the possible emergence of life. In the context of the preparations for the Wideband Sensitivity Upgrade (WSU), the Acatama Large Millimiter and sub-millimiter Array has acquired a demonstration dataset, the largest ever produced of any protoplanetary disk, consisting of 32 separate observations of the HL Tauri disk with high angular and spectral resolution. In this thesis I analyze the dataset with the aims to investigate the chemical composition of the HL Tau disk, especially in relation to the late accretion events described in the literature, which are thought to favor the enrichment of the gas phase with S-bearing molecules. I find sixteen line detections at high confidence belonging to nine molecules, namely four isotopologues of CO, HCO^+,CN,HCN,HNC, and SO, originating from three spatially distinct regions of the HL Tau system. I derive spatial and line properties and compare them with those found in the literature. I derive column densities and temperature of SO using additional detections from the literature, better constraining them and confirming the origin of the molecule in relation to late accretion events. I also perform an analysis of the continuum emission of HL Tau, deriving the SED and the spectral index map.

Contents

Introduction and objectives

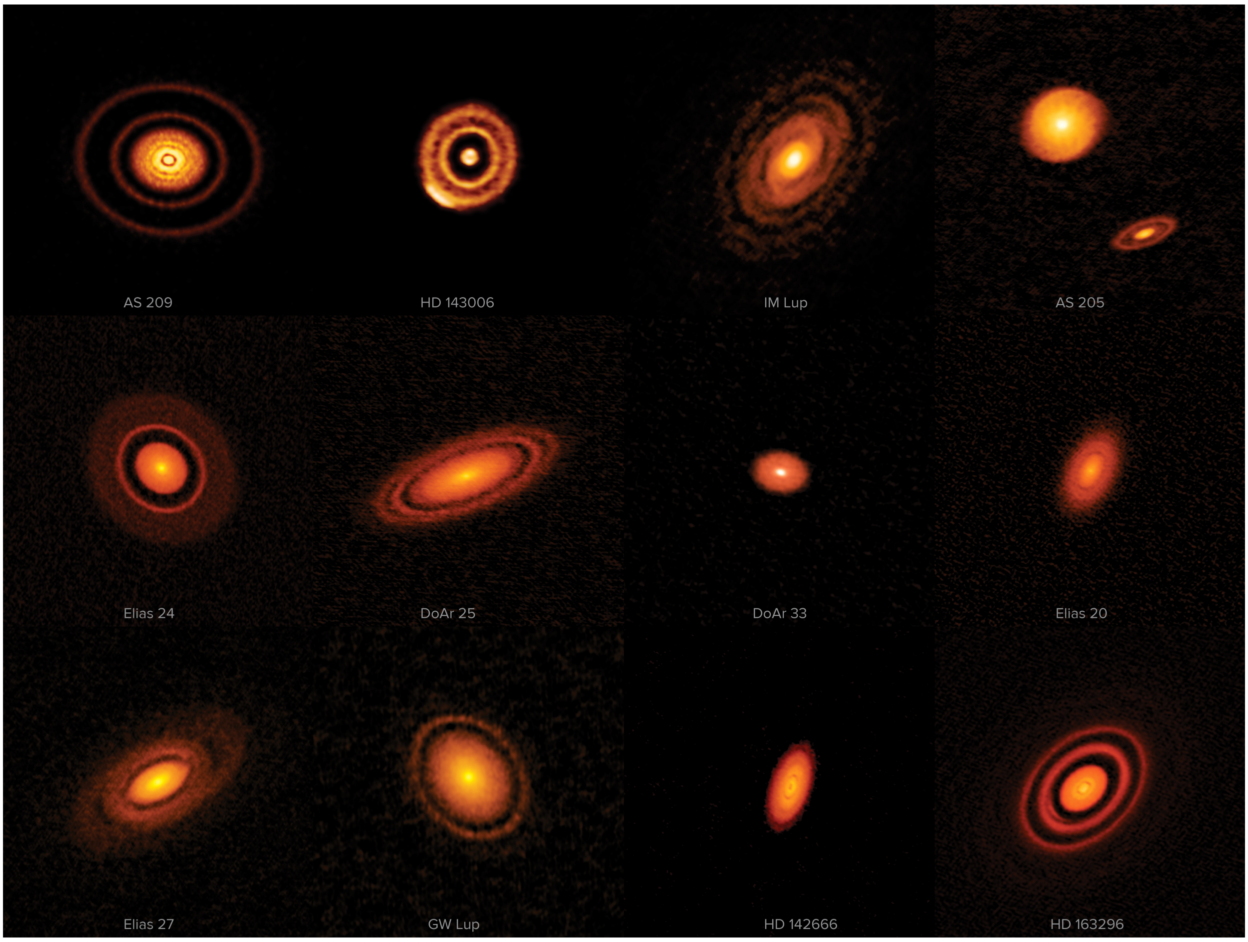


Figure 1: A collection of nearby protoplanetary disks imaged by ALMA during its first Large Program, known as the Disk Substructures at High Angular Resolution Project (DSHARP, Andrews et al. 2018). The name of each disk is reported under each image. The program unveiled remarkable substructures and spirals in almost all of the targeted disks, making a giant leap in our understanding of planet formation. Credit: ESO.[1]

In 1995 Mayor & Queloz made the first ever detection of a planet orbiting a main-sequence star other than the Sun. Since then, the search for exoplanets has been a prolific one, with well over 3500 having been observed with a variety of different methods. Protoplanetary disks are a fascinating window into the birthplace of planets and minor bodies, and their study plays a key role in the scientific effort to understand the formation not only of exoplanets, but also of our own solar system, the mechanisms of their dynamical and chemical evolution, and ultimately the origin of life. These disks of gas and dust form around young stars and become visible within the first 1 Myr of a star's life. The mass of protoplanetary disks will eventually aggregate to form the components

---

[1] `https://www.eso.org/public/images/potw1904a`

of a stellar system, with planet formation expected to be well underway within the first few Myr since the star's initial collapse.

After confident detections in star-forming regions and protostellar envelopes of several chemical species thought to be important for the emergence of life, such as a diverse set of Complex Organic Molecules (COMs), the chemical study of protoplanetary disks represents the new frontier of astrochemistry, with the aim to characterize the molecular budget available in disks, which will ultimately determine the composition of planets' surfaces and atmospheres, and possibly provide the building blocks of life. However, the detection of molecules in disks is made much harder by the fact that these systems are colder than protostellar envelopes, and a lot of their material is trapped in solid form, inaccessible to line observations.

With the advent of ALMA's very high resolution mm and sub-mm observations, the scientific community has investigated protoplanetary disks with unprecedented detail, leading to the discovery of substructures and complex dynamical interactions with the disk's surroundings (Andrews et al. 2018), as well as the first detection of methanol emission lines (Walsh et al. 2016), and even observations of actively forming planets embedded in disks (Teague et al. 2018). The upcoming Wideband Sensitivity Upgrade (WSU) planned for 2030 promises to ensure that ALMA will remain at the technological cutting edge of radio astronomy for the next decades, and is sure to bring even more revolutionary achievements.

In the context of the preparations for the WSU, ALMA has acquired a large demonstration dataset consisting of 32 high angular and spectral resolution observations of HL Tau, whose protoplanetary disk is one of the most extensively studied. Being a very young disk, it is still interacting with its ambient ISM, with accretion spirals and outflows, but the substructures observed within it are suggestive of already active planet formation. HL Tau has thus been a popular target for astrochemical studies, which have led to the discovery of molecules associated with slow impact events from late accretion (Garufi et al. 2022), and the detection of water (Facchini et al. 2024). Still, the budget of larger molecules and COMs remains elusive.

With this thesis I aim to investigate the demonstration dataset acquired by ALMA with the objective to detect and identify any molecular spectral line present in the dataset, analyze their properties, including spatial distribution and integrated line fluxes, derive column densities and temperatures when possible, and compare the results with findings in the literature to better constrain them. Additionally, I will take advantage of the large frequency range covered by the demonstration dataset to produce a wideband continuum image of HL Tau and study its spectral energy distribution, which can be used as a probe for its dust properties.

Before delving into the analysis of the dataset, Chapter 1 of this thesis will provide the necessary scientific framework by covering the mechanisms of star formation and the properties of protoplanetary disks, by outlining the state of the art of astrochemistry in star- and planet-forming environments, and by introducing HL Tau and its disk. Chapter 2 will touch on the theoretical basics of radiative transfer and molecular spectroscopy, as well as summarize concepts of radio astronomy relevant to this work. It will also discuss the tools and methodologies adopted throughout the project. Chapter 3 will list the steps taken for the reduction and imaging of the data, then present the results of the spectroscopic and continuum analysis of the HL Tau disk. Finally, in Chapter 4, the results will be summarized, and some future prospects will be given.

Chapter 1

Scientific Framework

1.1 Star formation and protoplanetary disks

1.1.1 The birth of stars and early stellar evolution

Figure 1.1: Image of the Sh 2-106 (S106) compact star forming region in the constellation Cygnus taken by the Hubble Space Telescope. A newly formed star at the center of the image is accreting dust from the surrounding molecular cloud. Credit: NASA & ESA.[1]

[2]Stars are born in giant molecular clouds, large accumulations of gas made up of mostly molecular hydrogen, whose cores ~ 0.1 pc in size reach densities as high as $n_H \sim 10^6$ cm^{-3} and temperatures as low as $T \sim 10$ K. In order to become stars, these cores must undergo a collapse that will bring their size and density to typical values of $\sim 10^9$ m in radius and $\sim 10^{30}$ cm^{-3} for a sun-like star.

In a thermally supported cloud mutual gravitational attraction acts to bring gas particles together in a denser configuration; particles are at the same time pulled apart by the outward pressure they exert as a consequence of their velocity, inherently resulting from the gas temperature. Such system

---

[1] `https://esahubble.org/images/heic1118a`

[2] Main references in Section 1.1.1: Williams 2021, Frediani 2023, Armitage 2020.

will become unstable and collapse if it reaches a large enough mass, such that the gravitational force overcomes the opposing pressure. The critical mass scales with the gas temperature and density as:

$$M_J \propto T^{3/2}\rho^{-1/2} \ , \tag{1.1}$$

thus a colder, denser cloud requires a lower amount of mass to begin to collapse.

This is what makes the cores of molecular clouds the ideal location for star formation: the cloud is opaque to ionizing stellar radiation from the interstellar radiation field, with most UV light being absorbed in the outermost layers, but the cloud is also optically thin to line emission from chemical species which act as coolants, allowing gas to efficiently radiate energy away. Note however that even deep in the interior of molecular clouds heating processes can occur, especially as a result of cosmic rays, which are not absorbed by the outer layers of the cloud, and which can heat the gas through direct collisions and secondary ionizing agents that result from them, like UV photons and electrons.

The subsequent 10^7 years after the initial collapse will see the pre-stellar core become first a protostar, then a main sequence star, while the envelope in which it is embedded will be cleared by stellar winds and jets, accretion onto the central object, and aggregation into planetesimals and minor bodies. Collectively, stars undergoing formation are known as Young Stellar Objects (YSOs), and are classified based on observational characteristics into four classes, each tracking a different stage in early stellar evolution.

- **Class 0** YSOs have a positive spectral index in the range 2–10 µm ($\alpha_{IR} > 0$, $S_{\nu[2-10\mu m]} = \nu^{\alpha_{IR}}$). These YSOs are in their initial collapse phase and their central object is obscured by the contracting envelope. Their SED is dominated by sub-mm emission from the infalling material. At this stage the material closer to the star is expected to begin forming a circumstellar disk (see Section 1.1.2).

- **Class I** YSOs maintain a spectral index $\alpha_{IR} > 0$. About 0.1 Myr after the initial collapse, the protostar, emits as a cold (~ 10 µm) black-body, while the surrounding material starts to be diffuse enough to distinguish the hot disk, visible as an infrared excess. At this stage outflows or jets also appear.

- **Class II** YSOs, also known as T Tauri stars, are defined by a spectral index $-1.5 < \alpha_{IR} < 0$. The central object, now ~ 1 Myr old, has entered the pre-main sequence, having reached a high enough temperature in its core to activate (but not yet sustain) hydrogen fusion. With most of the envelope accreted by the star or settled onto the disk, The SED features a stellar black-body at ~ 2 µm, while the infrared excess from the disk flattens the spectrum at longer wavelengths.

- **Class III** YSOs have almost completely cleared their envelope. The stellar black-body dominates the SED, with a possible infrared excess from the cooling disk, likely with planet formation at an advanced stage. The spectral index reaches $\alpha_{IR} < -1.5$. About 10 Myr have passed since the initial collapse.

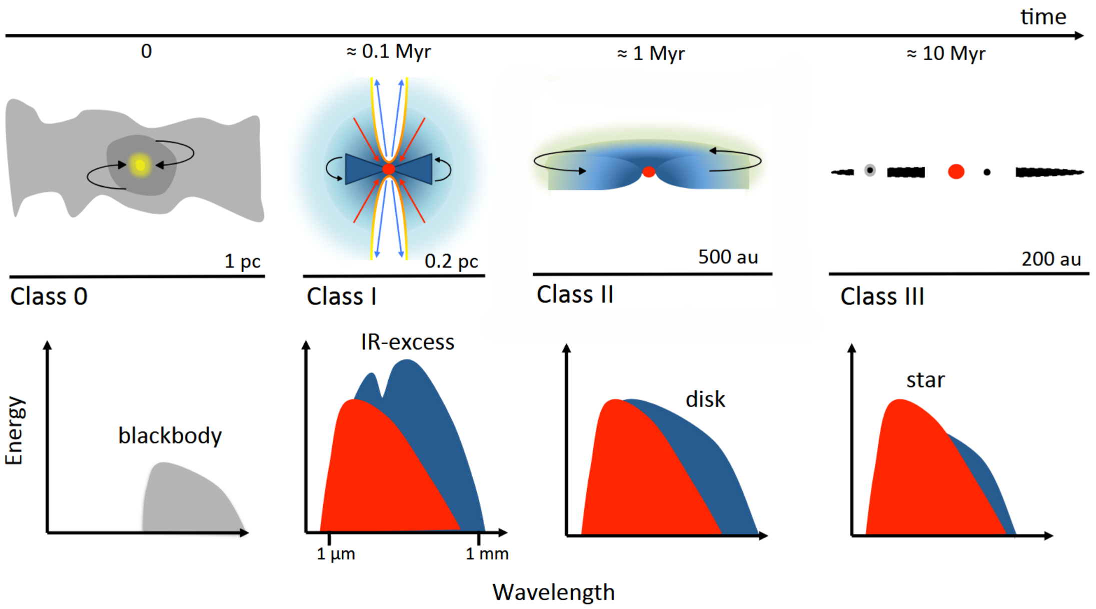


Figure 1.2: A sketch of the four classes of YSO evolution and their observed SED. Credit: Miotello (2018).

1.1.2 Protoplanetary disks: structure and properties

[3]Protoplanetary disks form around young stars as a direct consequence of the fact that the cores of molecular clouds are expected to have non-zero angular momentum. Indeed, a cloud cannot directly collapse to stellar densities, since the increase in rotational velocity resulting from such a jump in size scale would be such that the centrifugal force would overcome the gravitational pull of the star. Instead, after the central object has reached a certain size, the rest of the infalling material will arrange itself in a disk aligned with the rotational equator of the protostar, thus distributing the angular momentum along the extent of the disk. Angular momentum is further redistributed in the system by accretion of matter from the disk onto the star (see Section 1.1.3) and by the action of outflows and jets.

Vertical and radial structure

As discussed above protoplanetary disks are observed in a YSO SED as an infrared excess to the the main black-body emission of a star. In first approximation one can consider a disk to be non-evolving (static), and *passive*, meaning its only emission comes from the dust, responsible for absorbing and reprocessing radiation from the host star. A flat disk would then intercept progressively fewer amounts of radiation at increasingly large radii, due to its geometry, and the fact that radiation flux decreases as r^{-2}. Thus, dividing the disk into infinitesimal anuli and assuming a constant temperature within each of them, the SED of the protoplanetary disk can be modelled as (Williams 2021):

$$F_\nu = \frac{2\pi \cos i}{d^2} \int_{R_{min}}^{R_{max}} B_\nu(T)(1 - e^{-\tau_\nu}) r dr \quad , \tag{1.2}$$

where i is the inclination angle of the disk ($i = 0°$ corresponding to face-on) and $\cos i$ corrects for the inclination with respect to the line of sight, d is the distance of the disk, and $\tau_\nu = \kappa_\nu \Sigma(r)/\cos i$ is the optical depth of the emitting material, the product of $\Sigma(r)$, the surface density of the disk at each radius, and κ_ν the absorption coefficient.

---

[3] Main references in Section 1.1.2: Williams 2021, Armitage 2020, Miotello et al. 2023.

The above expression also requires the knowledge of a temperature profile, which can be assumed to be in the form $T(r) \propto r^{-q}$, with $q = 3/4$ for a flat disk. This model manages to reproduce the observed SED of disks faithfully in the near-infrared, but underestimates the flux at longer wavelengths, suggesting the disk's temperature does not decrease with radius as rapidly as expected from a flat disk model. Indeed in reality the temperature profile is much flatter, because the disk is *flared*, meaning that its vertical thickness increases with radius and thus the disk intercepts more starlight at large radii than it would if it were flat.

One can find the vertical density profile of gas in a disk $\rho(z)$ by solving the vertical component of the equation of hydrostatic equilibrium under the following assumptions: (*i*) a geometrically thin disk, such that the disk thickness $h(r)$ is much smaller than the radius r at all radii; (*ii*) a disk mass M_{disk} much smaller than the stellar mass M_* so that the disk is not self gravitating; and (*iii*) a vertically isothermal temperature gradient. Thus, in cylindrical coordinates:

$$\frac{dP}{dz} = -\rho g_z = -\rho \frac{GM_*}{(r^2+z^2)^{3/2}} z \ . \tag{1.3}$$

Where $P = c_s^2\rho$, $c_s = k_B T/\mu m_H$, k_B is the Boltzmann constant, and m_H the mass of the hydrogen atom. Since it follows from assumption (*i*) that $z \ll r$, and remembering that the keplerian angular velocity $\Omega_K = \sqrt{GM_*/r^3}$, one can write:

$$\frac{dP}{dz} = \frac{d(c_s^2\rho)}{dz} \simeq -\Omega_K^2 \rho z \ , \tag{1.4}$$

which can be integrated to yield:

$$\rho(z) = \rho_0 \exp\left(\frac{-z^2}{2h^2}\right) \ , \tag{1.5}$$

where ρ_0 is the density at $z = 0$ (at the midplane), and $h \equiv c_s/\Omega_K$.

A radial density profile cannot be derived from the assumption of a static disk, however the expression for the rotational (or azymuthal) velocity of the gas v_ϕ as a function of radius proves to be interesting. It can be derived from the fluid equation for momentum conservation:

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v} = -\frac{1}{\rho}\nabla P - \nabla\Phi \ , \tag{1.6}$$

whose azymuthal component can be written in the midplane as:

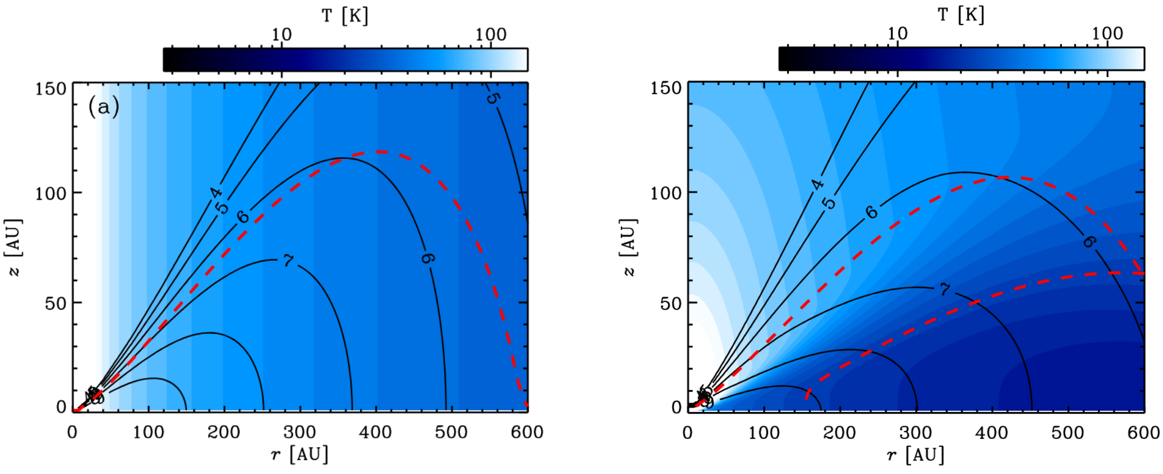
$$\frac{v_\phi^2}{r} = \frac{GM_*}{r^2} + \frac{1}{\rho}\frac{dP}{dr} \ . \tag{1.7}$$

From here v_ϕ can be rewritten in terms of the keplerian velocity $v_K = \sqrt{GM_*/r}$ as:

$$v_\phi^2 = v_K^2 + \frac{r}{\rho}\frac{dP}{dr} \ . \tag{1.8}$$

Equation 1.8 shows how the pressure gradient experienced by gas in a disk results in a gas velocity which deviates from a strict keplerian velocity. Indeed, since the gas is hotter and denser closer to the star $dP/dr < 0$, meaning the gas in a disk has sub-keplerian velocity. This deviation is not a large one, but has important consequences on the evolution of dust particles, as will be explained in Section 1.1.3.

The description made so far relies on the approximation of an isothermal vertical structure, which however does not hold true in real disks, because of the shielding from stellar radiation operated by the surface layers of the disk, and cooling from molecular emission. In general, a disk can be vertically divided into a hot surface layer, a warm molecular layer, and a cold midplane. This stratification has important consequences on the chemistry of the disk, as will become apparent in Section 1.2.3. A more accurate density profile can then be derived by defining a vertical temperature gradient. One such model was developed by Rosenfeld et al. (2013) to account for the observed asymmetry in the ^{12}CO $J = 2-1$ and $J = 3-2$ emission. Figure 1.3 provides a summary of this model.



(a) **(b)**

Figure 1.3: Density profile and temperature comparison between (a) a vertically isothermal model, and (b) a hydrodynamical model with a temperature profile constructed by Rosenfeld et al. (2013). Temperature is shown in color scale; density is in black contours denoting $\log n_{gas}$ in cm^{-3}. Credit: Rosenfeld et al. (2013).

Dust size and opacity

Opacity is an important probe of dust growth in protoplanetary disks. Indeed the optical depth of a grain population strongly depends on the maximum grain size a_{max}. Traditionally, dust size in YSOs has been determined through spectral index observations in the millimeter, assuming an optical depth dominated by absorption, and an optically thin emission from the dust. In the absorption-dominated case the optical depth can be defined as $\tau_\nu = \Sigma_{dust}\kappa_\nu$, where κ_ν is the absorption coefficient. It behaves spectrally at millimeter wavelengths as $\kappa_\nu \propto \nu^{\beta_\kappa}$, with the slope β_κ being dependent on a_{max}. With the additional assumption of optical thinness ($\tau_\nu \ll 1$), one can relate β_κ with the spectral index α of the dust SED F_ν through the following equation:

$$F_\nu \simeq F_0(\nu/\nu_0)^\alpha = F_0(\nu/\nu_0)^{2+\beta_\kappa} \quad , \tag{1.9}$$

valid in the Rayleigh-Jeans approximation. Thus with the measurement of $\alpha = 2+\beta_\kappa$ the maximum dust size a_{max} can be determined. With this approach dust grains up to cm-size have been detected in disks, with a radially increasing spectral index suggesting larger grains closer to the star.

However the conditions of optical thinness and absorption-dominated optical depth may not hold in protoplanetary disks, where optically thick ($\tau_\nu \gg 1$) or moderately optically thin ($\tau_\nu \leq 1$) emission is expected, and the effects of scattering cannot be ignored. Indeed, in the limit of optically thick dust emission the value of α in (1.9) approaches 2, which will cause an overestimation of dust sizes.

Without making assumptions on the optical thinness of the emission, and including scattering in the definition of τ_ν, Carrasco-González et al. (2019) have determined a maximum grain size in the HL Tau protoplanetary disk of a few millimeters.

1.1.3 Evolution of protoplanetary disks and planet formation

Viscous accretion

[4]In order for a star to continue to accrete material from its surroundings, angular momentum must be redistributed or removed from its disk. A way in which this is achieved is through the action of viscosity. Since the material rotates faster closer to the star, one can imagine collisions between particles in differentially rotating sections of the disk to slow down material at small radii, transferring angular momentum outwards. In reality, the cross-section of particle collisions in a disk is not large enough to lead to significant viscosity effects in a timescale comparable to the observed evolution timescale of protoplanetary disks. Instead, turbulence in the disk is thought to provide an *effective* or *turbulent* viscosity by mixing material at neighbouring radii. This viscosity can be described in terms of the sound speed in the material c_s and the disk scale height h through the so-called α *prescription*:

$$\nu = \alpha c_s h \ , \tag{1.10}$$

as introduced by Shakura & Sunyaev (1973), where α is a dimensionless parameter. The result of this redistribution of angular momentum to larger radii is a *stretching* of the disk.

Dust evolution

The behaviour of dust in protoplanetary disks is tied to its aerodynamical interaction with gas. When dust moves at a velocity v relative to the velocity of the gas it experiences drag. The drag force is proportional to the dust surface area presented to the colliding gas, and the resulting acceleration scales like force divided by the particle mass. It follows, then, that for a spherical particle of fixed density the acceleration experienced by the dust will become negligible once it has grown to a large enough size. Therefore it is important to study how small dust particles are affected by drag before they agglomerate into larger bodies.

The mechanism through which dust experiences drag depends on the size of the particles. For particles smaller than the mean free path of the gas, the experienced drag will depend on the frequency of collisions between dust and gas particles. This is known as *Epstein drag*. The interaction of particles whose size exceeds the mean free path of gas can instead be treated in classical fluid terms, through the so-called *Stokes drag*, without needing to take into account single particle collisions.

In the absence of turbulence dust both smaller and larger than the gas mean free path will undergo vertical settling onto the disk midplane. This happens as a result of the fact that gas is supported above the midplane by pressure, while dust is not. As a result, any dust above the midplane will experience the downward component of the gravitational force, and move toward the midplane. The presence of gas turbulence in real disks, however, results in the smaller dust, which is more tightly coupled with the gas, to be swept up and remain suspended above the midplane. In order to become strongly concentrated in the disk midplane, then, dust must grow beyond a certain size.

Dust also undergoes radial migration towards smaller radii. This happens again both under Epstein and Stokes drag, though through different mechanisms. As mentioned in Section 1.1.2, the azymuthal velocity of gas in a disk is sub-keplerian. Small dust which is coupled with the gas will be swept along with it and thus also have a sub-keplerian velocity. However, because dust is not supported by pressure, this velocity will be insufficient to balance gravity at a given radius, the dust will thus experience and inward drift towards the central star. Larger grains, on the other hand, are decoupled from the gas and their motion is keplerian. By moving through sub-keplerian

---

[4] Main references in Section 1.1.3: Armitage 2020.

gas, dust of this size feels a headwind that reduces its angular momentum, resulting again in an inward drift.

The radial drift of dust has important consequences on the formation of planetesimals. Indeed, the estimated size-dependent timescale for dust grains to drift all the way to the central star is much smaller than the estimated lifetime of a disk. In particular, at a distance on 1 au it is estimated that a dust grain will reach its maximum inward velocity at around 1 meter in size. This implies that in order for planet formation to occur, planetary seeds must grow beyond meter sizes very quickly, or their radial drift be otherwise slowed down or halted. Indeed, the inward drift is caused by the outwardly decreasing pressure gradient of the gas, but if local maxima in the pressure profile are created by turbulent motions dust can instead drift towards these positions and accumulate.

From dust to planetesimals to planets

Dust is expected to grow to at least mm - cm sizes through *coagulation* resulting from dust collisions. In order to continue growing through collisions, dust particles need to stick to each other without fragmenting or eroding. The outcome of a collision depends on its cross section, which ultimately depends on the velocity and size of the colliding grains. At intermediate scales growth mechanisms are the most uncertain, but it is clear they must occur quickly. One proposed way in which dust could rapidly grow to large scales is through *gravitational instability*, which suggests dust accumulated on a thin disk in the midplane, as a result of vertical settling, might become unstable and break down into clumps and filaments that would quickly collapse to planetesimal sizes.

After reaching km scales, planetesimal interaction becomes dominated by gravity. Terrestrial planets are expected to from in the inner disk and their growth is thought to continue through collisions of planetesimals of similar size and gravitational focusing of smaller bodies onto larger ones. Gas giants are expected to form in the outer disk, either through the initial formation of a solid core large enough to begin runaway accretion of gas, or through disk instability which would lead to the direct collapse of gas into a compact configuration.

1.2 Astrochemical evolution of YSOs

1.2.1 Chemical processes in YSOs

[5]The study of astrochemistry in star-forming and planet-forming environments aims at unveiling the chemical processes that take place at different stages of star formation and evolution, and that ultimately shape the chemical richness of stellar systems. The molecular species emerging from these processes make up the initial chemical budget of planets and moons, and form the basis that might lead (and has led in at least one instance) to the emergence of life. Indeed, the now ubiquitous detection of complex organic species in star-forming environments suggests that at least some of the chemical complexity necessary to life is inherited by planets from earlier stages of stellar evolution. It thus becomes crucial to investigate the evolution of astrochemistry from the very beginning of star formation, trying to understand at each step the dominant processes for the creation and destruction of molecular species, and how much of the chemical complexity inherited from earlier stages is preserved, destroyed, or reprocessed.

Chemical processes in all stages of YSO evolution are strongly dependent on environmental conditions, particularly temperature, and exposure to ionizing radiation, as these determine whether

---

[5] Main references in Sections 1.2.1 and 1.2.2: Frediani 2023, de Simone 2022, Williams 2021, Öberg et al. 2023.

different species undergo solid-phase or gas-phase reactions.

When a species reaches its *freeze-out* temperature, it adsorbs onto dust grains, exiting the gas-phase. Dust grains become coated in ices (primarily H_2O, CO, and CO_2 ices, followed by CH_3OH, NH_3, and CH_4) and within these icy mantles solid-phase chemistry can take place, as species migrate along the surface of the grain and react. High energy photons or cosmic rays often act as catalysts for these reactions, making the species mobile within the ice or dissociating them (photolysis and radiolysis). Products of solid-phase reactions can re-enrich the gas-phase if the temperature rises enough to sublimate the icy mantles, or as a result of non-thermal desorption (e.g. as a result of the energy released by a chemical reaction, or sputtering).

It is only when species are in the gas-phase that they can be detected through emission resulting from electronic, vibrational, and rotational transitions. Reactions in this phase are regulated by the exposure to ionizing radiation, which breaks molecular bonds through photodissociation. Thus, only ions generally survive in areas exposed to direct stellar or interstellar radiation.

1.2.2 Astrochemistry of pre- and proto-stellar environments

The first locations for the creation of complex molecules are pre-stellar environments (molecular clouds and cores undergoing stellar formation). These will be inherited and reprocessed in protostellar envelopes (in YSOs of Class 0 and I). These are the stages at which the detection of complex organic molecules has been most successful so far.

Molecular clouds

Plenty of molecular species, including complex organics, have been found already at this stage. Indeed, the cold and dense environment, and crucially the self-shielding against UV radiation, make the inner regions of molecular clouds the perfect location for the creation of complex molecules. Outside molecular clouds photodissociation from the interstellar radiation field prevents the formation of molecules. Thus, at the interface between molecular clouds and the diffuse interstellar medium only atoms and ions are present.

Moving inward, the cloud becomes partially shielded. Here the first simple molecules form through gas phase reactions, including *polycyclic aromatic hydrocarbons* (PAHs).

In the cold prestellar cores, a variety of molecules are present, mostly frozen-out in icy layers around dust grains. In these ices solid-phase reactions take place involving the addition of hydrogen and deuterium atoms to oxygen, nitrogen and carbon atoms (hydrogenation and deuteration), resulting in the production of H_2O, CH_4, and NH_3, and their deuterated counterparts. At temperatures below 10 K CO, the second most abundant volatile in molecular clouds after H_2, freezes out. This allows for its hydrogenation and deuteration in the solid phase, forming formaldehyde (H_2CO) and methanol (CH_3OH), along with their deuterated forms. But the depletion of CO from the gas phase also makes way for reactions between species that remain volatile which were not possible before. In particular, the production of H_2D^+ and H_2 from H_3^+ and HD (an exothermic reaction which becomes favoured with respect to its inverse thanks to the low temperatures) can only occur in an environment depleted of CO, which prevents competitive gas-phase reactions between CO and H_3^+.

Interstellar Complex Organic Molecules (iCOMs) have also been deteceted in prestellar cores. These are molecules with more than six atoms, including at least one carbon atom. They are the most complex species that can arise in the ISM and are thus extremely relevant to the study of the formation of complex pre-biotic molecules. Their detection through line emission suggests they populate prestellar cores at least partially in the gas phase, although it is not yet clear through what mechanisms.

Protostellar envelopes

As the protostar starts to produce luminosity through contraction, the temperature gradient in the surrounding envelope is inverted. In the outskirts of the envelope the conditions are similar to those found in prestellar cores, with molecules mostly frozen-out into ices and iCOMs detected in the gas phase.

Moving inward towards the *warm layer* of the envelope, the temperatures rise above 22 K, the sublimation temperature of CO. As the CO ices melt they release many species in the gas phase, including CH_4, which is responsible for rich gas-phase chemistry leading to the production of unsaturated carbon chains.

In the envelope layers closest to the central protostar, temperatures rise above the water sublimation temperature (100 K). Here all molecules are released into the gas phase, allowing for abundant detections. These regions have been found to be particularly enriched in a variety of iCOMs and/or unsaturated carbon chains. The former have been dubbed *hot corinos* to distinguish them from the more extended hot cores surrounding the protostar, where iCOMs were detected in a much lower abundance. iCOMs detected in teh gas phase within hot corinos include: formaldehyde H_2CO, methanol CH_3OH, formic acid HCOOH, acetaldehyde CH_3CHO, glycolaldehyde CH_2OHCHO, methyl formate CH_3OCHO, dimethyl ether CH_3OCH_3, acetic acid CH_3COOH, methyl cyanide CH_3CN, ethyl cyanide C_2H_5CN, formamide NH_2CHO, and propyne CH_3CCH (Cazaux et al. 2003, Frediani 2023, de Simone 2022). In the case of the hot corino hosted by IRAS 4A2, spatial segregation of some of these iCOMs was found, with methanol tracing the largest scales and glycolaldehyde the smallest, with emission from all examined iCOMs contained within 50 au, comparable to the size of the solar system (Frediani 2023).

At this stage, YSOs also begin to feature jets and outflows, which promote unique chemistry within their shocked regions. In these environments fast shocks ($\geq 10\,\mathrm{km\,s^{-1}}$) drive the enrichment of the gas phase through gas-grain (sputtering) and grain-grain collisions. The increase in temperature also triggers gas-phase chemistry unique to those environments.

Slow shocks ($\sim 1\,\mathrm{km\,s^{-1}}$) have also been observed in accreting YSOs at the *centrifugal barrier*, i.e. the interface between the infalling envelope and the forming proto-disk. These slow shocks are not energetic enough to cause grain-grain shattering, but can release species from the icy mantles of grains through sputtering (Garufi et al. 2022 and references therein).

1.2.3 Astrochemistry of protoplanetary disks

[6]Chemical processes become increasingly less straightforward as YSOs evolve into more complicated structures. Protoplanetary disks are indeed very complex systems both in their structure and their evolution, but in first approximation the main parameters that drive the chemistry within them remain temperature and exposure to ionizing radiation. Temperature is itself driven by exopsure to stellar and external radiation (and thus the optical thickness of the disk), as well as cosmic rays, and the distribution of cooling molecules. With this in mind, one can make an initial qualitative separation of the disk into layers, with increasingly colder layers further from the star and at higher vertical depth from the surface.

The hottest layers close to the star and on the entire surface of the disk are completely exposed to radiation and can reach temperatures of 1000 K. Here, all inherited chemistry is reset: molecules are dissociated and only atoms and ions survive. Underneath the surface layer volatiles should populate a warm molecular layer, where gas-phase chemistry can take place, reprocessing inherited species. Here *isotopic fractionation* through isotopologue-specific photodissociation can occur,

---

[6] Main references Section 1.2.3: Öberg et al. 2023, Frediani 2023, Miotello 2018, Soave 2024.

whereby the slight differences in the energy required to photodissociate different isotopologues of a certain species result in the enrichment of some isotopologues over others.

In the deeper, most shielded region, known as the midplane, ices are expected to survive, with at least a partial preservation of the molecules inherited from the ISM and prestellar stages. Here, chemistry is regulated by the snowlines (or snow-surfaces) of individual molecules, which will determine whether a molecule is found in the gas-phase or is frozen out. The resulting adsorption and desorption of certain molecules at different positions in the disk will trigger specific gas-phase and solid-phase chemistry, also as a result of the depletion of certain species which would otherwise block certain reactions. Another type of isotopic fractionation occurs here due to different adsorption energies between heavier and lighter isotopologues.

Disks, however, are not static systems, and snow-surfaces are expected to change in position and shape depending on variation in stellar activity and physical conditions of the disks. Molecules can also survive in their volatile form outward of their snowlines due to desorption caused by cosmic rays.

Furthermore, vertical and radial mixing of dust and gas is ubiquitous in the disk, and causes a constant exchange of volatiles and icy grains between different layers. The migration of dust is also relevant to the distribution of molecules in the disk. Larger grains tend to settle to the midplane, but dust also migrate inward, carrying molecules trapped within their ice layers to smaller radii. Nevertheless dust migration is not straightforward, as dust can become trapped in pressure minima and accumulate around snowlines.

As mentioned in Section 1.2.2, shocks can also provide conditions for unique chemistry. In young protoplanetary disks which have not yet cleared their surrounding envelope, shocks can originate as a consequence of late accretion events onto the disk, or as a result of winds impacting the surrounding medium. Emission from S-bearing species in particular was found in some sources, associated with an accretion streamer impacting on the disk surface (Garufi et al. 2022, see Section 1.3 for more details).

With all these considerations in mind it is clear that determining which molecules inherited from the ISM ultimately survive the protoplanetary disk phase trapped in ices, and which instead are reprocessed or created in-situ is difficult, and requires physical and chemical modelling of the disks.

CO and its isotopologues

Carbon monoxide (CO) is the second most abundant molecule in the ISM (and in disks), right after H_2, but differently from the latter it is readily observable through its rotational lines. CO is an important gas tracer, and has been used extensively for estimating the gas mass of disks. Its less abundant isotopologues $C^{18}O$ and $C^{17}O$ are particularly useful as their emission lines are often optically thin and thus are able to probe the innermost layers of the gaseous disk. As already established, CO is also a molecule of interest for pre-biotic and complex astrochemistry, as its ices host a variety of solid-phase reactions.

The reprocessing of CO that takes place in disks is thus of relevance, and particularly the processes of isotopic fractionation through isotope-specific photodissociation. Indeed, in the hot surface layer of disks CO is dissociated by UV photons through line absorption. This absorption can become optically thick, causing self-shielding of CO, which then survives in a molecular layer beneath the disk surface. Since shielding is stronger where the species column density is higher, the less abundant isotopologues of CO will have their photodissociation layer extended further down, closer to the disk midplane. As a result, a layer of the disk will form where the main CO isotopologue ($^{12}C^{16}O$) survives in gas form, while the less abundant ^{13}CO, $C^{18}O$, and $C^{17}O$

are still photodissociated. An isotope ratio computed in disks from CO line observations such as $C^{18}O/^{12}CO$ will then reflect this phenomenon and result to be much lower than the corresponding ratio in the ISM.

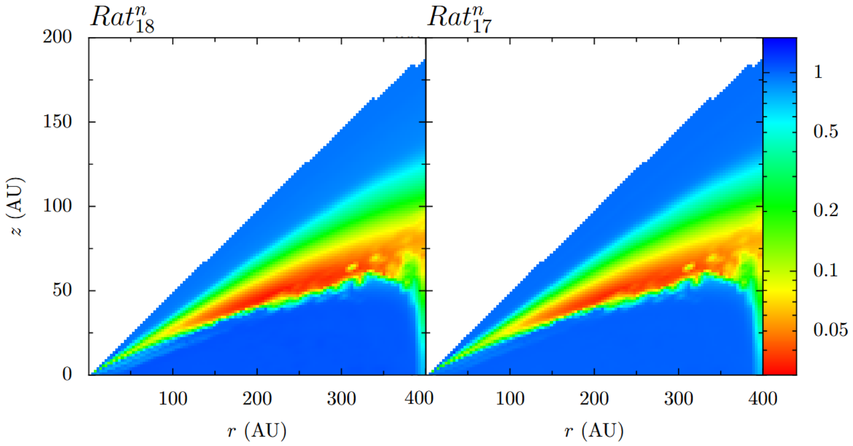


Figure 1.4: Graph of a disk simulation performed by taking into account the effect of isotope-specific photodissociation. It shows the relative abundance of $C^{18}O$ and $C^{17}O$ with respect to ^{12}CO, corrected for the ratio found in the ISM. These corrected abundance ratios are referred to as Rat^n_{18} and Rat^n_{17}, and they show how the ratios $C^{18}O/^{12}CO$ and $C^{17}O/^{12}CO$ strongly deviate form the expected isotope ratios in the ISM, within a specific layer of the disk where ^{12}CO is dissociated but the other isotopologues are not. Credit: Miotello (2018).

COMs in disks

While Complex Organic Molecules have been succesfuly observed and characterized in many protostellar environments (Class 0/I YSOs), their detection in protoplanetary disks has been much more challenging. This is mainly due to the fact that disks are much colder systems and in their deep layers much of the COM budget is thought to be locked within ices where they are impossible to observe. Figure 1.5 gives a summary of the COMs detected in disks so far. The amount of reprocessing that these molecules undergo before they end up onto larger bodies through dust accretion is still uncertain.

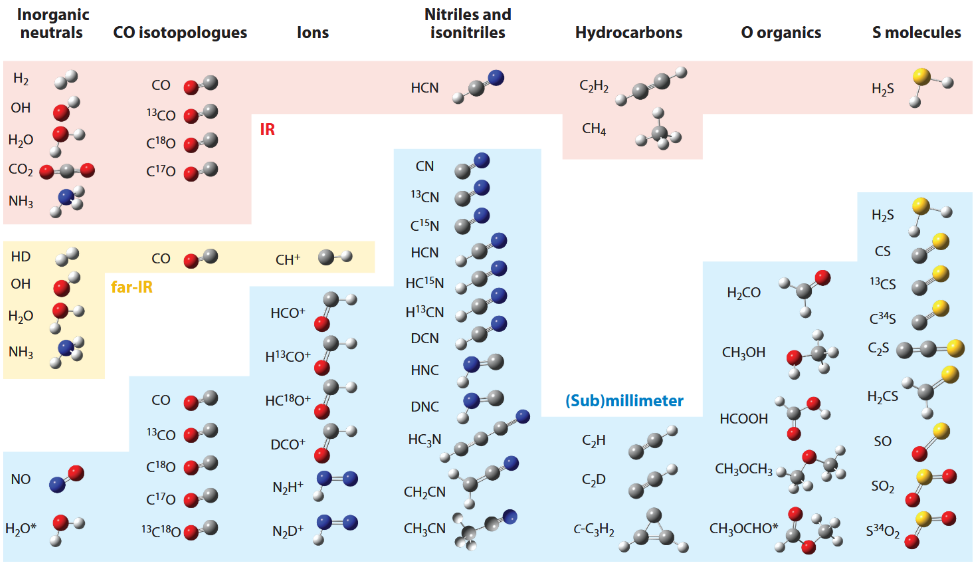


Figure 1.5: Summary of Complex Organic Molecules detected in disks as of 2023. Different colors indicate different wavelegths at which the detections were made: infrared (red), far-infrared (yellow) and (sub)millimeter (blue). Credit: Öberg et al. (2023).

1.3 Astrochemistry of HL Tau

HL Tauri (HL Tau) is a T Tauri protostar in the Taurus star-forming complex, located at a distance of ~ 140 pc and with an estimated mass of 2.1 $M_\odot$. It is classified as an intermediate Class I/II YSO, still partially embedded in its gaseous envelope, remnant of the molecular cloud from which it formed, with a protoplanetary disk clearly visible in the mm, and highly collimated outflows traced by molecular and atomic gas. Having just emerged from its protostar phase, its unique transitional nature has made it the ideal target to study the astrochemical link between early and late-stage star formation, as well as the formation of planetesimals and protoplanets.

The HL Tau protoplanetary disk has therefore been the subject of extensive studies, having drawn particular interest after ALMA observations at very high resolution during the 2014 Long Baseline Campaign revealed for the first time the presence of a substructure of concentric bright and dark rings in the disk, which correspond to gaps in the dust density and which are now known to be nearly ubiquitous in disks. Their first discovery in a disk as young as HL Tau suggested that planetary formation might begin much earlier than previously thought (ALMA Partnership 2015).

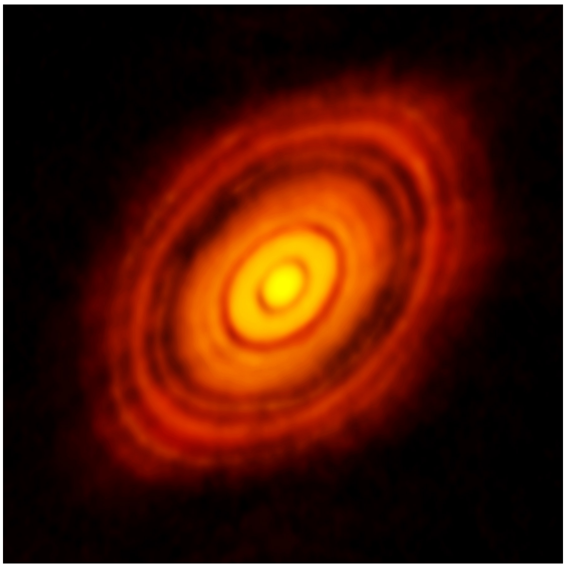


Figure 1.6: Continuum image of HL Tau at 1 mm taken by ALMA. HL Tau was the first protoplanetary disk where a ring substructure was observed. Credit: ALMA Partnership (2015).

Late accretion events and S-bearing species

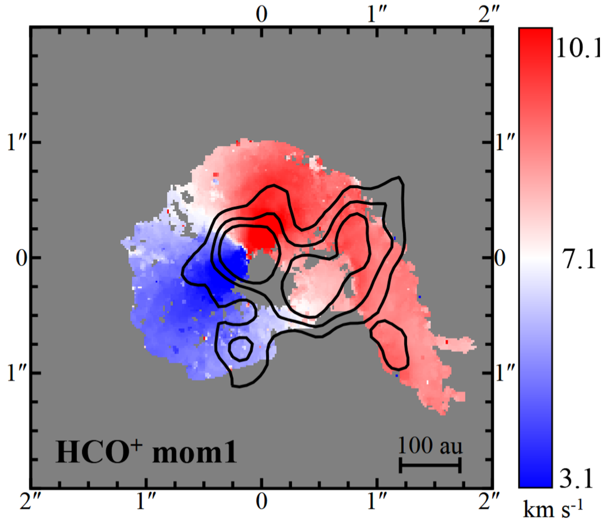


Figure 1.7: A velcity map of HCO^+ in HL Tau showing the streamer protrusion to the North-West (Yen et al. 2019a). Detections of SO_2 are overlaid in contours. Credit: Garufi et al. (2022).

There is evidence that the diffuse envelope arund HL Tau is still interacting with the protoplanetary disk, with late accretion episodes having been observed in the form of streamers traced by CO, HCO^+ and HC_3N (Yen et al. 2019a). Where these streamers impact the disk they produce slow shocks, creating the conditions for many molecular species to enrich the gas phase as a result of sputtering of the icy mantles of dust grains. Localized emission of S-bearing species (SO and SO_2 in particular) has been observed at the intersection of these streamers with the HL Tau disk, at velocities coherent with that of the streamer, as well as towards the center of the disk, with velocities comparable to the keplerian disk velocity. The association of these detections with the faster shocks of HL Tau's outflows is excluded, as they do not align spatially. It is also unlikely that the release of these molecules into the gas phase is tied to thermal desorption, as that would result in a more symmetric emission in the disk. Instead it is probable that the S-bearing species are being released in the shocks and later spiral towards the inner sections of the disk before any chemical reprocessing occurs (Garufi et al. 2022).

Water and methanol

HL Tau is also the first disk where water vapor emission lines have been spatially and spectrally resolved. Facchini et al. (2024) found three water lines within the inner regions of the disk, from which they were able to place an upper limit on the water snowline at 17 au.

The spatial resolution of gas-phase water in HL Tau sparked the search for complex organic molecules whose chemical evolution is tied to that of water. Methanol (CH_3OH) is one such molecule, whose formation in the ISM is known to occur only within ices through solid-state chemistry, and whose volatility is similar to water (sumblimation temperature around 100-150 K); it is thus expected that methanol should be detected in the gas-phase close to the water snowline. Methanol is also a particularly significant molecule for pre-biotic chemistry, as it provides a bridge for the formation of more complex COMs. Soave (2024) attempted to detect methanol in HL Tau with ALMA Band 7 observations resulting mostly in non-detections, with the possible exception of one line, whose attribution to methanol is however uncertain. Methanol is, on the other hand, detected in several more evolved disks such as TW Hydra (Walsh et al. 2016).

Chapter 2

Methodologies

2.1 Theoretical background

2.1.1 Specific intensity and flux density

[1]Consider a source radiating energy as a function of frequency E_ν. The energy dE_ν that reaches a detector within a solid angle $d\Omega$, and through a projected area $\cos\theta\ d\sigma$, in a time dt, and in a narrow frequency band $d\nu$ is given by:

$$dE_\nu = I_\nu\ \cos\theta\ d\sigma\ d\Omega\ dt\ d\nu \quad , \tag{2.1}$$

where I_ν is known as *specific intensity* or *spectral brightness*. The above expression can be rearranged considering the definition of power as energy per unit time $dP = dE/dt$ to yield:

$$I_\nu = \frac{dP}{(cos\theta\ d\sigma)\ d\Omega\ d\nu} \quad , \tag{2.2}$$

which gives a quantitative definition of I_ν as power received from a source per unit frequency, per unit area, and per unit solid angle. Its units are $\mathrm{W\,m^{-2}\,Hz^{-1}\,steradian^{-1}}$. Despite this operative definition given in terms of a measurement, specific intensity is a property intrinsic to the source, and crucially does not change with the distance between source an detector.

Because observations have a finite resolution, however, specific intensity is not measured directly, instead it is integrated over an angular area. The integral of specific intensity over a solid angle is known as the *flux density* of the source:

$$F_\nu = \int I_\nu\ \cos\theta\ d\Omega \quad . \tag{2.3}$$

Its units are $\mathrm{W\,m^{-2}\,Hz^{-1}}$, but astronomers make use of a more convenient unit, the Jansky (Jy): $1\ \mathrm{Jy} = 10^{-26}\,\mathrm{W\,m^{-2}\,Hz^{-1}}$. Differently from the specific intensity, flux density is a measurement that depends on the distance between source and observer, following the inverse square law: $F_\nu \propto d^{-2}$.

For most interferometers, each pixel of the final output image is assigned a value in units of $\mathrm{Jy\,beam^{-1}}$. This is a way of expressing specific intensity, using the number of *beams* (which represent the angular resolution of an interferometer, see Section 2.2) contained within an area. Thus over a sky area Ω the flux density in Jy is:

$$F_\nu\ [\mathrm{Jy}] = I_\nu\ \left[\frac{\mathrm{Jy}}{\mathrm{beam}}\right] \times \frac{\Omega}{\Omega_{beam}} \quad , \tag{2.4}$$

---

[1] Main references in Section 2.1: Williams 2021, Condon & Ransom 2016, Wilson et al. 2009.

where Ω_{beam} is the sky area of a beam, and so Ω/Ω_{beam} is the number of beams contained within Ω. However since an image contains discrete pixels, in order to extract the flux density in Jy within an arbitrary area, it is useful to first convert the image units from Jy beam^{-1} to Jy pixel^{-1} by dividing for the number of pixels in a beam (multiplying for the number of beams in a pixel):

$$I_\nu \left[\frac{\mathrm{Jy}}{\mathrm{pixel}}\right] = I_\nu \left[\frac{\mathrm{Jy}}{\mathrm{beam}}\right] \times \frac{\Omega_{pixel}}{\Omega_{beam}} \quad , \tag{2.5}$$

where Ω_{pixel} is the sky area of a pixel. The flux density in Jy will then be easily obtained by summing over each pixel in the area of interest.

2.1.2 Radiative transfer

Light emitted by a source interacts with the material that it encounters within the source itself and along the path between the source and the observer. This causes the specific intensity to be modified along the line of sight, through the action of emission and absorption of the traversed matter. The specific intensity I_ν of radiation going through a pathlength ds is modified by a positive term due to emission, and a negative term due to absorption, according to:

$$dI_\nu = j_\nu ds - \kappa_\nu I_\nu ds \quad , \tag{2.6}$$

where j_ν is the *emission coefficient*, and κ_ν is the *absorption coefficient*. Note that the negative term includes I_ν, since the number of absorbed photons depends on the number of photons that pass through the material.

One can define *optical depth* as:

$$d\tau_\nu = \kappa_\nu ds \tag{2.7}$$

$$\tau_\nu = \int \kappa_\nu \, ds \quad . \tag{2.8}$$

Optical depth gives a measurement of the transparency of the material to light, and thus of the amount of matter that light interacts with along a pathlength ds. Thus, (2.6) can be rewritten in terms of $d\tau_\nu$ as:

$$\frac{dI_\nu}{d\tau_\nu} = S_\nu - I_\nu \quad , \tag{2.9}$$

where $S_\nu = j_\nu/\kappa_\nu$ is known as the *source function*.

The radiative transfer equation has the general solution:

$$I_\nu(\tau_\nu) = I_\nu(0)\, e^{-\tau_\nu} + \int_0^{\tau_\nu} S_\nu \, e^{(\tau_\nu - \tau')} d\tau' \quad , \tag{2.10}$$

which shows how the specific intensity is the sum of an initial background intensity attenuated by τ_ν, and a contribution from the emission of the material along the line of sight, corrected for absorption.

The specific intensity assumes a specific form in the case of *thermodynamic equilibrium* (TE). A system is said to be in TE when matter and radiation are in equilibrium at the same temperature.

A collection of particles in TE at temperature T is such that the proportion of particles in different energy levels follows the *Boltzmann distribution*:

$$\frac{n_u}{n_l} = \frac{g_u}{g_l} e^{-(E_u - E_l)/k_B T} \quad , \tag{2.11}$$

where n and E represent the number density of particles and energy of two different levels, *up* (u) and *low* (l), while g is the statistical weight, or degeneracy, associated with each level. The radiation field that results from a system of particles in TE at a temperature T is given by the *Planck function*:

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/k_B T} - 1} \quad , \tag{2.12}$$

also known as *blackbody radiation*. In these conditions, the specific intensity is a blackbody everywhere: $I_\nu = B_\nu(T)$, and $dI_\nu/d\tau_\nu = 0$ resulting in $S_\nu = B_\nu$ and thus:

$$j_\nu = \kappa_\nu B_\nu(T) \quad , \tag{2.13}$$

which is known as *Kirchoff's law of thermal radiation*, and relates the emitting properties of matter with its absorbing properties.

Thermodynamic equilibrium is, however, rarely met in star forming environments. A sufficient approximation that results valid in most cases of interest to this work is *local thermodynamic equilibrium* (LTE), where matter is locally in thermal equilibrium at a well-defined kinetic temperature, but it is not in equilibrium with the radiation field. In LTE Kirchoff's law still applies, since κ_ν and j_ν are properties of the material independent of the surrounding radiation.

By assuming that S_ν is the same everywhere, it can be taken out of the integral in (2.10) so that the solution to the radiative transfer equation can be written as:

$$I_\nu(\tau_\nu) = I_\nu(0)\, e^{-\tau_\nu} + S_\nu\, (1 - e^{-\tau_\nu}) \quad . \tag{2.14}$$

Two regimes thus emerge: for $\tau_\nu \gg 1$ the emission is *optically thick*, and (2.14) reduces to $I_\nu = S_\nu = B_\nu$, as a result of Kirchoff's law. Thus as optical depth increases, the emission from a system in LTE approaches a blackbody. Indeed, a blackbody is defined as the emission from a perfectly absorbing material. On the other hand, for $\tau_\nu \ll 1$, the emission is *optically thin*, and I_ν becomes dominated by $I_\nu(0)$, with a small emission contribution as $1 - e^{-\tau_\nu} \sim \tau_\nu$.

When neither TE nor LTE is met, the population of the particles' energy levels can still be described by (2.11), with the substitution of the kinetic temperature T with the *excitation temperature* T_{ex} defined as:

$$T_{ex} = \frac{E_l - E_u}{k_B \ln{(g_u n_l / g_l n_u)}} \quad . \tag{2.15}$$

This temperature is not a physical one and can vary between different pairs of energy levels. The difference between T and T_{ex} parametrizes the departure of the system from LTE. Even in non-equilibrium Kirchoff's law remains valid, again with T_{ex} in place of T: $j_\nu = \kappa_\nu B_\nu(T_{ex})$.

2.1.3 Brightness temperature

For blackbody radiation such that $h\nu \ll k_B T$, which is the case in most contexts of radio astronomy, the Planck function simplifies to:

$$B_\nu(T) \approx \frac{2\nu^2}{c^2} k_B T \quad . \tag{2.16}$$

This is known as the *Rayleigh-Jeans approximation* (RJ). The direct proportionality that exists between the intensity of radiation and the temperature of its source in the RJ regime has led radio astronomers to use *brightness temperature* in place of intensity, or brightness.

Given a measure of intensity I_ν, brightness temperature is defined as the temperature T_b which would result in I_ν if inserted into the Rayleigh-Jeans law, i.e.:

$$T_b = \frac{c^2}{2k_B\nu^2} I_\nu \quad . \tag{2.17}$$

Of more use to this work is the brightness temperature expressed in terms of flux density:

$$T_b = \frac{c^2}{2k_B\nu^2} \frac{1}{\Omega} S_\nu \quad , \tag{2.18}$$

where Ω is the sky area solid angle, over which the flux density is collected, expressed in steradians.

Despite its definition stemming from the properties of a blackbody, T_b is used just as easily outside backbody conditions. Only in the case of blackbody emission, however, does the brightness temperature coincide with the thermodynamical temperature of the source.

2.1.4 Basics of spectral lines

Emission and absorption of photons occurs through the transition of particles between quantized energy levels. Photons that are emitted or absorbed in a given transition are such that their energy $h\nu$ matches the energy difference ΔE between levels in the transition. As a result, detectable spectral features are produced at specific frequencies known as emission or absorption lines.

The spectral lines appear broadened over a range of frequencies due to Heisenberg's uncertainty principle (natural broadening). Further broadening is caused by the Doppler effect, as a result of the internal motions of particles in the material. The Doppler effect also causes a shift in the observed frequency of the line resulting from bulk motion of the material. By comparing the observed frequency of a line with its rest frequency one can therefore infer the systematic bulk velocity of the emitting/absorbing gas.

The proportion of emitted or absorbed photons as a function of frequency is described by the *line profile* $\phi(\nu)$, normalized such that $\int \phi(\nu)\, d\nu = 1$.

In the case of electronic transitions of single atoms, an electron may absorb a photon and transition from a lower energy level l, to a higher energy level u. This occurs at a rate $B_{lu}\,\bar{u}$, where $\bar{u}$ is the profile-weighted energy density of the radiation filed, averaged over direction. An electron may also spontaneously decay from level u to level l at a rate A_{ul}, emitting a photon in the process. Finally, the radiation field can stimulate the emission of a photon from an electron in level u at a rate B_{ul}, causing it to transition to level l. This latter phenomenon is known as *stimulated emission*, or *negative absorption*. B_{lu}, A_{ul}, and B_{ul} are known as the *Einstein coefficients*. They can be shown to be related to each other by:

$$g_l B_{lu} = g_u B_{ul} \quad , \qquad B_{ul} = \frac{c^2}{2h\nu^3} A_{ul} \quad . \tag{2.19}$$

Equations 2.19 connect the phenomena of absorption and emission just like Kirchoff's law, but they do so at a microscopic level. It follows then that the Einstein coefficients should be related to the macroscopic absorption and emission coefficients. Indeed, one can find that the emission coefficient j_ν is directly related to $n_u A_{ul}$:

$$j_\nu = \frac{h\nu}{4\pi} n_u A_{ul}\, \phi(\nu) \quad , \tag{2.20}$$

while κ_ν depends on the difference between normal and negative absorption:

$$\kappa_\nu = \frac{h\nu}{4\pi} \left(n_l B_{lu} - n_u B_{ul} \right) \phi(\nu) = \frac{h\nu}{4\pi} \left(1 - e^{-h\nu/k_B T_{ex}} \right) n_l \, B_{lu} \, \phi(\nu) \quad . \tag{2.21}$$

In reality, in addition to radiative excitation and de-excitation, electrons can also be excited or de-excited through collisions between atoms, which cause particles to exchange between them the energy associated with the transitions, instead of it being radiated away or absorbed via photons.

As the density of the gas increases, so do the collisions between particles, leading to collisional excitations and de-excitations growing prevalent over their radiative counterparts. The density for which the collisional excitation rate C_{lu} becomes comparable to the radiative excitation rate A_{lu} is known as the *critical density* $n^* \propto A_{ul}$. For densities above n^* collisions dominate and the gas is thermalized at a common kinetic temperature, satisfying the conditions for LTE. The energy level population will thus be distributed according to the Boltzmann distribution (2.11).

2.1.5 Molecular spectroscopy

When atoms are paired together their outer electrons lie in shared quantized orbitals, and continue to absorb and emit photons while undergoing transitions. A molecule will however possess additional degrees of freedom in terms of vibrational and rotational motions, each with its set of quantized energy levels. A molecule will thus be able to produce spectral lines through electronic, vibrational, or rotational transitions. The energy jumps associated with each type of transition are very different, such that each is observed at different wavelength ranges. Typical energies are 10, 0.1, 0.001 eV, for electronic, vibrational, and rotational transitions respectively, corresponding to 100 nm, 10 µm, and 1 mm in wavelength.

Computing the energy levels of molecular transitions means solving the Schrödinger equation of the molecule. This is made easier if one applies the Born-Oppenheimer approximation, which works under the assumption that electronic and nuclear motions and rotations can be treated independently. With this approximation the wavefunction of a molecule can be separated into a nuclear and electronic part, with the nuclear part further separated into a rotational and a vibrational component. Thus the total energy of the system is simply given by the sum of an electronic, a vibrational, and a rotational contribution.

$$E_{tot} = E_{el} + E_{vib} + E_{rot} \quad . \tag{2.22}$$

Rovibrational transitions of linear molecules

For the purposes if this work it is sufficient to consider only vibrational and rotational transitions of linear diatomic molecules. In this case the vibrational part of the Schrödinger equation describes an harmonic oscillator, with linearly spaced energy levels quantized by the integer v:

$$E_{vib} = \left(v + \frac{1}{2} \right) h\omega_{vib} \quad , \tag{2.23}$$

where ω_{vib} is the vibrational frequency. Rotational energy levels are instead described by the quantum integer number J. They assume a quadratic form:

$$E_{rot} = h \, B \, J(J+1) \quad , \tag{2.24}$$

where B is known as the rotational constant and has units of Hz. A rotational ladder of energy levels is associated to each vibrational state, thus transitions occur between a level described by the pair (v, J) to a level described by the pair (v', J'). Transitions are more likely to occur between neighboring energy levels, with selection rules $\Delta v = \pm 1$, $\Delta J = \pm 1$. Purely rotational transitions, involving only the quantum number J with $\Delta J = \pm 1$, can take place in diatomic molecules only if they possess a permanent electric dipole (dipole transitions). It is therefore not possible to observe purely rotational transitions in diatomic homonuclear molecules such as H_2, which renders it a famously difficult to observe molecule despite it being the most abundant in the universe.

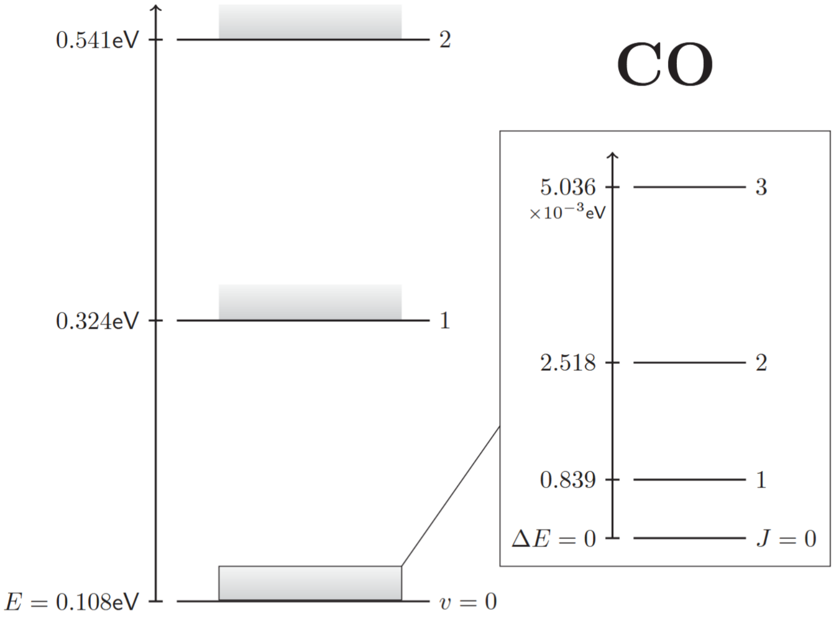


Figure 2.1: A representation of the vibrational and rotational levels of CO. The rotational levels are much closer together than the vibrational levels (lower associated energy jump) and are represented in the left panel as a gray shading. The right panel is a zoom-in showing a few individual rotational lines. Credit: Williams (2021).

Hyperfine transitions

The interaction between nuclei and electrons can be complicated by the presence of an electric quadrupole or magnetic dipole in nuclei. This leads to the splitting of molecular states into degenerate levels, giving rise to a hyperfine structure which requires additional quantum numbers to be described.

Catalogs may list different quantum numbers to identify a transition, but the most common in the case of linear molecules are: N, the the rotational angular momentum excluding electron and nuclear spins; $J = N + S$, the rotational angular momentum including electron spin; and $F = J + I$, the total angular momentum, which includes I, the nuclear spin. When it comes to selection rules in hyperfine transitions only $\Delta F = \pm 1, 0$ is strictly obeyed.

Linear molecules with a hyperfine structure include HCN and HNC, as a result of the nuclear spin of ^{14}N which induces a quadrupole moment. ^{13}C and ^{17}O also induce hyperfine splitting due to their magnetic dipole moments.

2.1.6 Rotational diagrams

Column density and temperature from line observations

The observation of the strength of molecular spectral lines can be used to derive the abundance and temperature of the emitting gas, given that certain conditions are met.

The abundance of an emitting gas is described through its *column density* N (not to be confused with the quantum number mentioned in the previous paragraph).

$$N = \int n \, ds \quad [\mathrm{cm}^{-2}] \quad , \tag{2.25}$$

where n is the number density of the gas in cm^{-3}, and the integral is computed along the line of sight. Starting from the solution to the radiative transfer equation (2.14), assuming no background sources ($I_\nu(0) = 0$), and remembering Kirchoff's law from which $S_\nu = B_\nu(T_{ex})$, one finds that the intensity from a source with optical depth τ_ν is described by:

$$I_\nu = B_\nu(T_{ex})(1 - e^{-\tau_\nu}) = \frac{2h\nu^3}{c^2}(e^{h\nu/k_B T_{ex}} - 1)^{-1} \left(\frac{1 - e^{-\tau_\nu}}{\tau_\nu}\right) \tau_\nu \quad . \tag{2.26}$$

Using (2.17), the intensity above can be expressed as a brightness temperature:

$$T_b = \frac{h\nu}{k_B}(e^{h\nu/k_B T_{ex}} - 1)^{-1} \left(\frac{1 - e^{-\tau_\nu}}{\tau_\nu}\right) \tau_\nu \quad . \tag{2.27}$$

Rather than T_b, however, the more significant and measurable quantity when observing spectral lines is its integral W over the line profile:

$$W = \int T_b \, dv \quad , \tag{2.28}$$

a quantity which directly relates to the *integrated line intensity* (or the integrated line *flux*), by taking the integral of both sides of (2.17) (or of (2.18)). Equation 2.28 uses $dv = c \, d\nu/\nu$. Indeed, the integration is often computed over velocity instead of frequency.

By integrating (2.27) over velocity, W can finally be related to the column density of the emitting material, through the dependence on $\int \tau_\nu \, dv$. To do this τ_ν is re-written using (2.20) and (2.11), with $E_u - E_l = h\nu$ for a line transition:

$$\tau_\nu = \int \kappa_\nu \, ds = \frac{8\pi\nu^2}{A_{ul} c^2}(e^{h\nu/k_B T_{ex}} - 1)^{-1} \, \phi(\nu) \int n_u \, ds \quad , \tag{2.29}$$

where $\int n_u \, ds = N_u$ is the column density of the upper level in the transition. The above equation can then be integrated over frequency (recall $\int \phi(\nu) \, d\nu = 1$) and, approximating ν and $(1 - e^{-\tau_\nu})/\tau_\nu$ as constants within the line, (2.29) can be combined with (2.27) and rearranged to yield:

$$N_u = \frac{8\pi k_B \nu^2}{h c^3 A_{ul}} W \left(\frac{\tau_\nu}{1 - e^{-\tau_\nu}}\right) \quad , \tag{2.30}$$

where a correcting factor ν/c is added to convert the frequency integration to velocity. Two important considerations follow from this result: (*i*) that no knowledge of the excitation temperature T_{ex} is required to derive the upper level column density, and (*ii*) that the upper level column density depends on the optical depth through the factor $\tau_\nu/(1 - e^{-\tau_\nu})$, which leads to the following distinction between the two optical depth regimes.

- For **optically thick** lines $\tau_\nu/(1-e^{-\tau_\nu}) \sim \tau_\nu$, and thus the column density cannot be derived without the knowledge of τ_ν in the source. Optically thick lines *hide* the contribution from the deeper layers of gas and thus the column density of the upper level population cannot be inferred solely from line observations. On the other hand, optically thick lines can be used to estimate the temperature of the source, as their intensity approaches a blackbody $B_\nu(T_{ex})$.
- For **optically thin** lines $\tau_\nu/(1-e^{-\tau_\nu}) \sim 1$, which means the upper level column density can be derived simply from a single line observation.

N_u is only half of the picture, as it only describes the population of molecules in the upper state of a given transition, but the assumption of optically thin lines is the first step to derive the total column density. To obtain N_{tot} from N_u the further assumption that the system is in LTE is required. In this case the levels are thermalized, and the entire population can be described by a single excitation temperature T, equal to the gas kinetic temperature. Thus it follows from the Boltzmann distribution that:

$$N_{tot} = \sum_i N_i = N_0 \sum g_i e^{-E_i/k_BT} = N_0\, Q(T) \quad , \tag{2.31}$$

where N_0 is the column density of the ground level, and $Q(T) \equiv g_i e^{-E_i/k_BT}$ is known as the *partition function*, and through it the measurement of the upper level column density of any transition can be converted to the total column density:

$$N_{tot} = \frac{e^{E_u/k_BT}}{g_u} N_u\, Q(T) \quad . \tag{2.32}$$

The partition function depends on T, and so, even under LTE conditions, when only one optically thin line is observed, deriving N_{tot} requires the assumption of the gas temperature. Fortunately, both total column density and temperature of a molecular gas can be derived if two or more optically thin lines from distinct transitions are observed.

The rotational diagram

By rearranging (2.32) and taking the natural logarithm of both sides one finds a linear relationship between $\ln(N_u/g_u)$ and E_u under the conditions of LTE and optically thin lines:

$$\ln\frac{N_u}{g_u} = \ln\frac{N_{tot}}{Q(T)} - \frac{E_u}{k_BT} \quad . \tag{2.33}$$

If line intensities from two or more different transitions are available, a linear fit can be performed on this relation, where the only unknowns are N_{tot} and T. Specifically, the slope m and intercept q of the fit would yield $-1/k_BT$ and $\ln(N_{tot}/Q(T))$ respectively. Therefore, if E_u is expressed in units of K (E [K] $= E$ [J]$/k_B$) then:

$$T = -\frac{1}{m} \quad , \tag{2.34}$$

which can be put into $Q(T)$ to obtain:

$$N_{tot} = e^q\, Q(T) \quad . \tag{2.35}$$

The value of $Q(T)$ at different temperatures can be found for different molecules in spectroscopical databases.

A plot showing $\ln(N_u/g_u)$ and E_u for each observed transition and the relative linear fit is known as a *rotational diagram*, or *population diagram*. A rotational diagram is not only useful for determining column densities and temperatures, but also to test the assumptions of LTE and optical thinness.

Indeed, if the population cannot be described by a single temperature (LTE does not apply), then the rotational diagram would not follow a straight line, although it is worth noting, as described by Goldsmith & Langer (1999), that for linear molecules non-LTE rotational diagrams may appear misleadingly straight even within large energy ranges.

If LTE applies but the lines are optically thick, and not enough points at different energies are available as to reveal a non-linear trend, then a rotational diagram fit will underestimate the total column density, while the temperature can be both underestimated in the lower energy range, and overestimated in the higher energy range.

In general, for non-LTE optically thick lines, disentangling the contribution from either phenomenon is not straightforward.

2.2 Radio interferometry

2.2.1 Radio antennas

[2]Antennas are the fundamental component of a radio interferometer. In their simplest form, they are dipoles that convert electromagnetic radiation into voltage. Antennas used for radio astronomy need a large collecting area, and are thus fitted with large parabolic reflectors that focus the electromagnetic waves in phase onto waveguide horns where the actual antenna is located.

The response of a receiving antenna can be described by its *power pattern*, or *beam pattern*, defined as the power received by an antenna as a function of direction per unit solid angle. In order for an antenna to be functional for radio astronomy it should not be isotropic, meaning it should not receive power equally from all directions. Rather, its power pattern should feature a large primary beam collecting most of the power from a given observing direction, and small secondary lobes (see Figure 2.2).

The power pattern response of an antenna observing a point source at infinity is known as the *Point Spread Function* (PSF). For an antenna with a circular aperture the PSF is given by the Airy function, and the width of its main lobe at half maximum (Full Width Half Maximum, FWHM; also known as Half Power Beam Width, HPBW) determines the angular resolution of the antenna, i.e. the minimum sky angle that can be resolved by the instrument. The angular resolution of an antenna depends on the diameter of its aperture D and the observed wavelength λ according to:

$$\theta_{FWHM} \approx \frac{\lambda}{D} \quad , \tag{2.36}$$

Two sources are resolved if they are separated by an angle $\theta \geq \theta_{FWHM}$. Given the dependency on D, in trying to improve the angular resolution of an instrument (decrease θ_{FWHM}) one soon runs into structural limits associated with building large dishes; one must then turn to radio interferometry to achieve higher resolutions.

---

[2] Main references in Sections 2.2.1 –2.2.4: Wilson et al. 2009, Taylor et al. 1999.

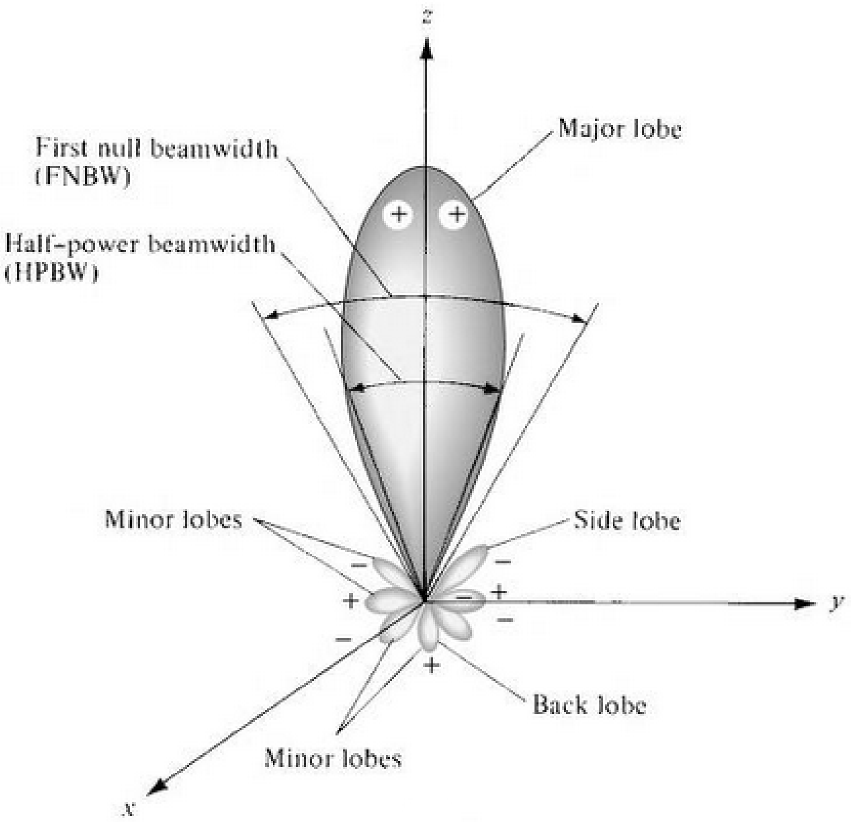


Figure 2.2: Power pattern of an antenna. Credit: Balanis (2016).

2.2.2 Two- and multi-element interferometer

The simplest radio interferometer consists of two identical dishes separated by a distance known as the baseline b. The antennas have output voltages $V_1 = V\cos[\omega(t-\tau_g)]$ and $V_2 = V\cos(\omega t)$, where $\tau_g = \frac{b\cos\theta}{c}$ is known as the geometric delay and accounts for the extra distance the signal must travel to reach antenna 1, and θ is the inclination angle of the antenna with respect to the baseline. A correlator multiplies V_1 and V_2 to produce the correlator response R, which varies sinusoidally over time:

$$R = \langle V_1 V_2 \rangle = \frac{V^2}{2}\cos(\omega\tau_g) \quad . \tag{2.37}$$

The correlator response is the product of two functions. The first is given by $\frac{V^2}{2} \propto S(A_1 A_2)^{\frac{1}{2}}$, where A_1, A_2 are the effective collecting areas of the two antennas, proportional to their power pattern. The second one, $\cos(\omega\tau_g)$ is a sinusoidal function representing the phase shift between the signals received by the two antennas. Although real antennas have two-dimensional apertures it is convenient to simplify the description to one dimension. It can be demonstrated that the power pattern of a uniformly illuminated 1-D aperture observing a point source at infinity is described by a *sinc*^{2} function. Thus, ultimately the correlator response of a one-dimensional interferometer will be a cosine modulated by a *sinc*^{2}.

The cosine period is regulated by the baseline length b, and determines the angular resolution (*beam width*) of the interferometer. The diameter D of each antenna instead defines the sinc^2 modulation and determines the Field Of View (FOV) of the instrument. Both can be expressed as:

$$\theta_{BW} \sim \frac{\lambda}{b\sin\theta} \quad , \quad \theta_{FOV} \sim \frac{\lambda}{D} \quad . \tag{2.38}$$

Equations 2.38 highlight the advantage of radio interferometers: by using antennas with the same diameter D, a higher angular resolution can be achieved by simply increasing the distance between them.

As can be seen from Figure 2.3 , however, the beam pattern of a two-element interferometer is not ideal. Instead of being concentrated in a main lobe, the signal is spread throughout the prominent sidelobes. To improve the shape of the beam pattern it is necessary to add more antennas (and thus more baselines) to the instrument. A multi-element interferometer extends the two-element system to N antennas which produce $N(N-1)/2$ unique responses given by each baseline. Each response from every baseline is averaged, resulting in a final synthesized beam which approaches a gaussian as N increases.

Longer baselines are more sensitive to compact objects and provide a higher angular resolution. Shorter baselines allow the interferometer to detect more extended objects, though with lower angular resolution. The maximum baseline b_{max} of the interferometer limits the angular resolution of the instrument, thus determining a minimum resolvable size. However an interferometer is also limited by its minimum baseline b_{min} which determines the maximum recoverable scale of the instrument. An interferometer is thus only sensitive to a range of angular sizes defined by:

$$\frac{\lambda}{b_{\text{max}}} < \theta < \frac{\lambda}{b_{\text{min}}} \quad . \tag{2.39}$$

Indeed, since b_{min} cannot be infinitely small, due to the finite size of the antenna dishes, no interferometer is able to observe isotropic sources, such as the Cosmic Microwave Background (although interferometers can observe CMB anisotropies), which can only be observed with a single antenna.

2.2.3 Complex visibilities and imaging

The response obtained in (2.37) corresponds to the output of the *cosine* correlator; however, to fully capture the power of the source, it is also necessary to compute the odd sine component of the signal, given by:

$$R_s = \frac{V^2}{2} \sin(\omega \tau_g) \quad . \tag{2.40}$$

From the two components of the response we can define the amplitude A and the phase ϕ :

$$A = ((R_c)^2 + (R_s)^2)^{1/2} \quad , \tag{2.41}$$

$$\phi = \tan^{-1}\left(\frac{R_s}{R_c}\right) \quad . \tag{2.42}$$

The correlator computes both the sine and cosine component simultaneously. This is called complex correlator. The response to any source of the interferometer with a complex correlator is called complex visibility:

$$\mathcal{V} = R_c - iR_s = Ae^{-i\phi} \quad . \tag{2.43}$$

Visibilities are the final output of an interferometer. One can define a cartesian coordinate system on a plane perpendicular to the direction of observation, with axes $u = x/\lambda$ and $v = y/\lambda$, with λ being the wavelength of the observed radiation. Through this coordinate system it is possible to visually represent the complex visibilities, as each visibility corresponds to one point on the (u,v) plane. Visibilities are related to the source sky brightness distribution through a 2-D Fourier Transform:

$$\mathcal{V}(u,v) = \int\int I_\nu(l,m)\, e^{-2\pi i(ul+vm)}\, dl\, dm \quad , \tag{2.44}$$

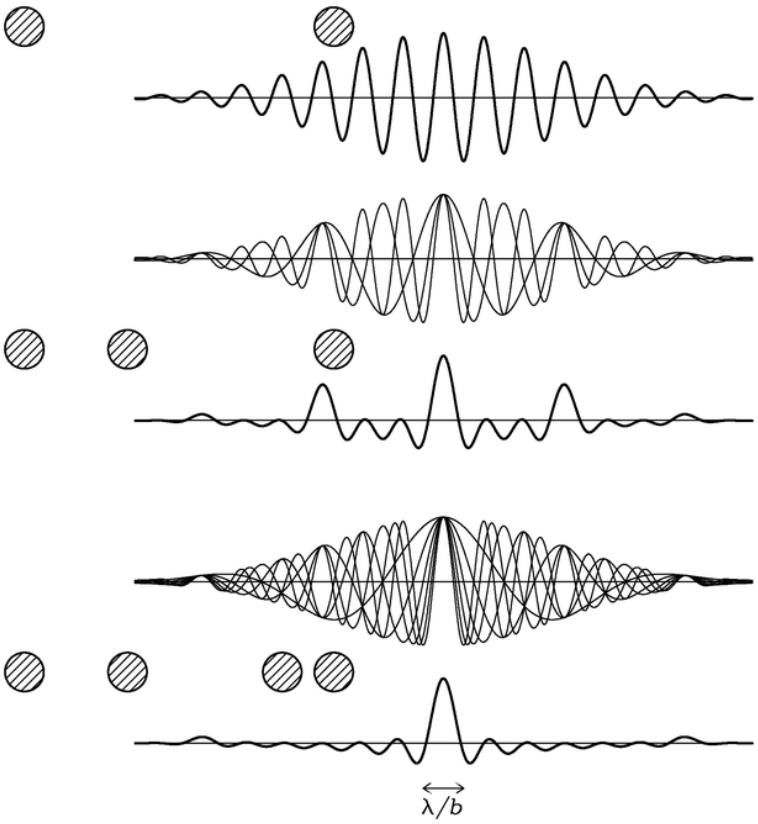


Figure 2.3: The power pattern of an interferometer approaches a gaussian as the number of baselines increases from one (two-element) to six (four-element). Credit: Condon & Ransom (2016).

where (l, m) are the cosines of the angular coordinates of the source in the sky. The source brightness distribution could then be obtained from the 2-D anti-Fourier Transform of the visibility function:

$$I_\nu(l,m) = \int\int \mathcal{V}(u,v)\, e^{-2\pi i(ul+vm)}\, du\, dv \quad . \tag{2.45}$$

However, an interferometer does not measure the full set of visibilities on the (u,v) plane, rather it samples a finite number of points. The more points on the (u,v) plane are sampled, the more signal from the source is collected. Each baseline samples one point (one visibility) per polarization, channel, and integration time. The (u,v) coverage is further increased by the effect of Earth's rotation, which changes the projected baseline lengths and therefore allows the interferometer to cover more points on the plane.

The sampled visibilities can be described as $\mathcal{V}(u,v)S(u,v)$, where $S(u,v)$ is the *sampling function*. From them a *dirty image* can be reconstructed as:

$$I_\nu^D(l,m) = \int\int \mathcal{V}(u,v)S(u,v)\, e^{-2\pi i(ul+vm)}\, du\, dv \quad , \tag{2.46}$$

Making use of the convolution theorem for Fourier transforms, the dirty image I_ν^D is related to the true source brightness I_ν through the following convolution:

$$I_\nu^D(l,m) = I_\nu * B \quad , \tag{2.47}$$

where B is the Fourier transform of $S(u,v)$, and corresponds to the PSF of the interferometer often referred to also as *dirty beam*.

From the above equations it is clear how visibilities and source brightness, as well as the PSF and the sampling function, are tied as Fourier pairs. The image output of an interferometer can never be the true source brightness because of the fact that the source is imaged through the PSF, which is not a perfect gaussian, and inevitably features sidelobes. The result is a dirty image, given by the convolution of the true source brightness distribution with the PSF. In Fourier space, this translates as a finite number of visibilities in the (u,v) plane, thus to the dirty image corresponds a sampled visibility function, resulting from the product of the complete visibilities with the sampling function.

In order to extract I_ν, one must deconvolve (2.47). This is usually done numerically through deconvolution algorithms such as the CLEAN algorithm, the basics of which will be discussed in Section 2.2.5.

2.2.4 Calibration

Interferometers are not perfect systems and they do not operate in perfect conditions. The sampled visibilities that an interferometer observes, therefore, will always differ from the *true* visibilities. These alterations can be modeled by a quantity G such that

$$\mathcal{V}^{obs} = G \cdot \mathcal{V}^{true} \quad . \tag{2.48}$$

G are known as the total complex gains, and they include both instrumental and atmospheric effects that inevitably corrupt the true visibilities. Calibration serves to correct for these gains by comparing the interferometric data with *primary* and *secondary calibrators*. The former are bright, featureless sources, usually quasars, within 30 deg of the target, observed only once for each observing session. They are used to determine the antenna response to a source of known flux density, in order to set a flux density scale for the target. The latter are sources observed frequently and within 4 deg of the target. They track time variable properties due to changing conditions of the instrument and the environment, most crucially atmospheric conditions.

From the calibrators, calibration tables are produced, which are then applied to the data to obtain a calibrated dataset.

It is worth noting that the calibration process is subject to intrinsic errors which affect the resulting dataset. Notably, the fluxes of calibrators come with their own uncertainty, and unaccounted for atmospheric events can disrupt their observation. Generally, this accounts for a $5-10\%$ uncertainty on the calibrated fluxes which should be taken into consideration.

2.2.5 Imaging with CASA

The Common Astronomy Software Applications package (CASA, Bean et al. 2022) is the software primarily used for handling, reduction, and imaging of data from ALMA and the Karl G. Jansky Very Large Array (VLA). CASA was developed by an international collaboration guided by the National Radio Astronomy Observatory (NRAO) in the United States.

In this work, CASA was used mainly to perform interferometric imaging. The tool provided by CASA for this purpose is the `tclean` task.

As mentioned in Section 2.2.3, the process of imaging in radio astronomy, i.e. converting sampled visibilities into a sky brightness distribution, requires the deconvolution of (2.47) through numerical means. The `tclean` task allows for various algorithms to be used for this process. The most popular, and the one used in this work, is based on the CLEAN algorithm developed by Högbom (1974) and Clark (1980).

The CLEAN algorithm as implemented by tclean

The basic assumption behind the CLEAN algorithm is that an extended source can be approximated as a sum of individual point sources. The process can be summarized with the following two equations:

$$I^R(x, y) = I^D(x, y) - \gamma B(x, y) * I^C(x, y) \ , \tag{2.49}$$

$$I(x, y) = I^C(x, y) * B^I(x, y) + I^R(x, y) \ . \tag{2.50}$$

Equation 2.49 is performed iteratively. In each iteration, the algorithm identifies the brightest point in the dirty image I^D, assuming it corresponds to a point source. This brightest point is called *clean component* I^C. The contribution of this clean component to the dirty image scaled by the γ factor (loop gain) is subtracted using the dirty beam, which is the response of the interferometer to a point source at that location. The algorithm then updates the residual image I^R which is used in the next iteration to identify the next brightest point. The process repeats until a stopping criterion is met, such as when the intensity of the residuals falls below a certain threshold, or after a fixed number of iterations.

Once the clean components have been identified and subtracted, through Equation 2.50 the algorithm restores the image. This is done by convolving the identified clean components with a clean beam B^I, which is typically a gaussian approximation of the dirty beam's central lobe, and adding them to the residual image. The result is a clean image I representing the true source brightness.

Masking

For more accurate cleaning, a region can be specified over which clean components are extracted. This is known as *masking* and can be done by providing the `tclean` task with a previously defined region file, by manually identifying regions at each iteration cycle using `tclean`'s interactive mode, or automatically by using auto-masking.

In order to automatically generate regions with auto-masking, `tclean` needs to be provided with threshold parameters to identify prominent emission around which to generate a region or *mask*. The current auto-masking algorithm available in `tclean` is called `auto-multithresh`. Its implementation is controlled by a set of sub-parameters, most notably:

- `noisethreshold` indicates the minimum signal-to-noise value that should be masked;
- `sidelobethreshold` indicates the minimum sidelobe level that should be masked, at each iteration the greater value between `sidelobethreshold` and `noisethreshold` is used as threshold for masking based on the rms noise and sidelobe level in the residual image;
- `lownoisethreshold` indicates a threshold for expanding a mask into low signal-to-noise regions, useful to capture fainter emission;
- `minbeamfrac` indicates a minimum region size in terms of beam fraction, any region smaller than this value is removed, or *pruned* from the final mask;
- `negativethreshold` indicates a threshold for negative features in the image, useful for detecting absorption.

Multi-Frequency Synthesis

When doing spectral line imaging, each channel is imaged separately and different masks for different channels can be applied. This produces an output in the form a datacube (see Section 2.2.6). However, the (u,v) data from an interferometer can also be combined during imaging to produce a single image of the source in *combined light*. The process is known as Multi-Frequency Synthesis (MFS), and can be used to obtain a total continuum image of the source. A more accurate Multi-Term MFS can be used for wideband imaging, which takes into account the effects of varying frequency on the (u,v) samples by fitting the source sky brightness as a function of frequency with a Taylor polynomial. MT-MFS then returns $N + 1$ images which act act as coefficients of the N^{th} order Taylor polynomial. The zeroth component image represents the sky brightness at the reference frequency (by default set as the central frequency of the observations), while the ratio between the zeroth and the first component images produce a map of the spectral index of the source.

Data weighting

During imaging it is possible to assign weights to different visibilities according to certain weighting schemes. This allows for the instrument's natural response to be altered in order to meet different observational needs and exploit an interferometer's capabilities to their fullest.

The simplest weighting options are *natural* weighting and *uniform* weighting. Natural weighting gives equal weight to all samples. Since visibilities corresponding to shorter baselines are sampled more, this results in a lower angular resolution and better sensitivity to extended emission. By contrast, uniform weighting weights each spatial frequency equally, thus shorter baselines become down-weighted. This results in a higher angular resolution and better sensitivity to compact emission, but a higher rms noise.

2.2.6 Products of interferometry: datacubes

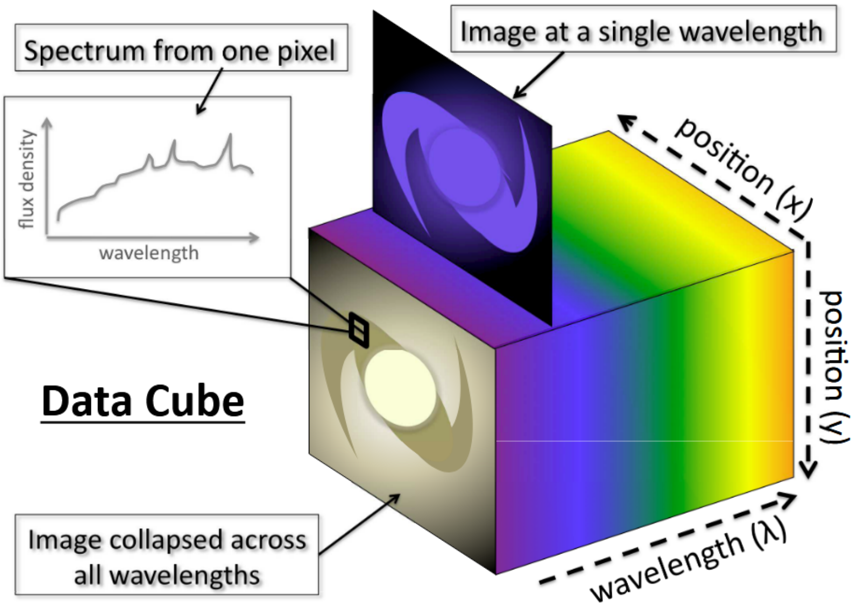


Figure 2.4: Schematic representation of a spectral cube and its components. Credit: Harrison (2014).

The final output of a spectral interferometric observation is a *datacube*, or *spectral cube* with an x and y axis that correspond to the sky coordinates Right Ascension (RA) and Declination (Dec), and a frequency or velocity axis divided into individual channels. Thus, a slice of a datacube taken parallel to the (x,y) plane reveals the image, called a *channel map* of the source (its sky brightness distribution) at a specific frequency channel, while the spectral information of each pixel is stored along the frequency axis.

For spectral line analysis datacubes are often continuum-subtracted, along each channel. Examining the continuum subtracted cube along the frequency axis reveals spectral features, and by collapsing the cube along that axis the spatial distribution of line-emitting regions can be studied. Depending on how pixels from each slice of the cube are combined, one can produce different output images, known as moment maps (see Section 2.3.2), which, in addition to information about the intensity distribution, can reveal details about the velocity and velocity dispersion of the emitting material.

2.2.7 Sensitivity and rms noise

The sensitivity of an instrument refers to the weakest detectable signal, i.e. the weakest signal that can be differentiated from the noise. It follows then that the sensitivity of an instrument depends on the sources of noise.

Noise contribution is usually described in terms of equivalent temperature of a matched resistor on the receiver. The power P_N received from any noise-like source can be expressed as a temperature:

$$T_N = \frac{P_N}{k_B \Delta\nu} \quad . \tag{2.51}$$

In these terms, a *system noise temperature* T_S can be defined as the sum of all noise power contributions, including the noise contributed by the science target. The latter is very small compared with the other contributions, most notably the atmospheric noise and the intrinsic instrument noise.

The root mean square fluctuations of the instrument's output can be related to the system noise temperature. For a single antenna (Condon & Ransom 2016):

$$\sigma_{rms} \approx \frac{T_S}{\sqrt{\Delta t \Delta\nu}} \quad , \tag{2.52}$$

where Δt is the observing time. For a interferometer with N antennas:

$$\sigma_{rms} = \frac{2k_B T_S}{A_e\sqrt{N(N-1)\Delta t \Delta\nu}} \quad , \tag{2.53}$$

where A_e is the effective collecting area. In order to be detectable, then, a signal must be larger than σ_{rms}. In other words, σ_{rms} describes the sensitivity of the instrument.

Crucial is the dependence of σ_{rms} on $\Delta\nu$. When doing continuum observations (or when doing MFS, combining together channels of a spectral cube) this means an increased sensitivity (smaller σ_{rms}) will result from a larger total bandwidth. For spectral observations $\Delta\nu$ refers to the channel width. An increased sensitivity can thus be obtained with channel averaging, at the cost of a lower spectral resolution.

2.2.8 Bandwidth and time-average smearing

Bandwidth smearing and time average smearing are distortions which affect an interferometer's final image. They persist after deconvolution with CLEAN and cannot be corrected with calibration. It is thus important to characterize the extent of their impact and minimize it.

Bandwidth smearing

Bandwidth smearing occurs as a result of the finite width of receiver passbands, whose visibilities are however treated as if corresponding to a single central frequency. This effect is further amplified when applying channel averaging to the data, which rebins all data within a given frequency band $\Delta\nu$ to a single channel of central frequency ν_0. The effects of smearing will be larger for a larger $\Delta\nu$, it is then useful to derive the relationship between $\Delta\nu$ and the reduction in peak response to a point source at the edge of the primary beam (Bridle & Schwab 1999):

$$R = I/I_0 = \frac{1}{\sqrt{1+\beta^2}} \quad , \quad \text{where} \quad \beta = \frac{\Delta\nu}{\nu_0}\frac{b_{max}}{2\sqrt{\ln 2}\, D} \quad ; \tag{2.54}$$

with R being the ratio between measured peak intensity and peak intensity without the effects of smearing, b_{max} being the maximum baseline, and D being the antenna diameter. The above equation is valid for a gaussian bandpass with a circular gaussian beam taper. Figure 2.5 shows how R varies as a function of β, in specific bandpass shapes and tapers.

Using (2.54) a maximum value can be computed for $\Delta\nu$ when doing channel averaging, corresponding to a given decrease in R that is deemed acceptable.

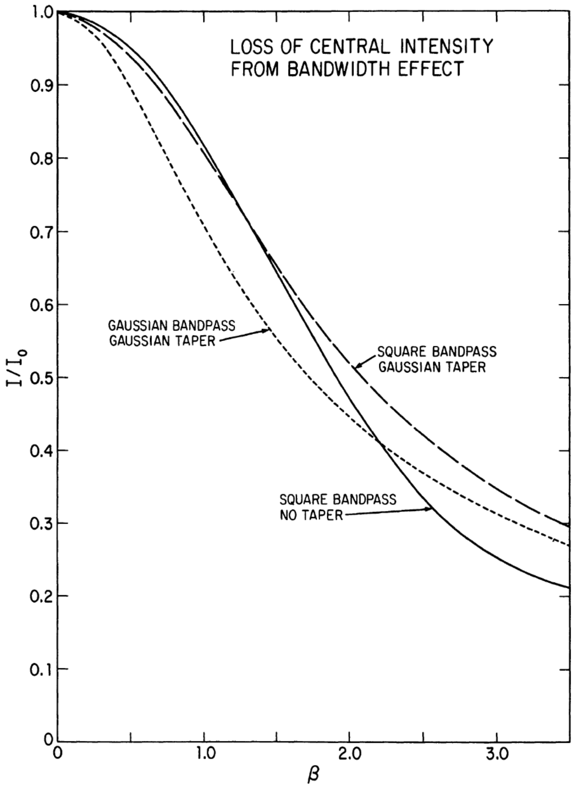


Figure 2.5: Loss of central intensity from bandwidth smearing for different bandpass shapes and tapers. Credit: Bridle & Schwab (1999).

Time-average smearing

Similarly to badwidth smearing, when a time average is performed visibilities are reassigned to the center of the time averaged interval, causing larger smearing for larger time average intervals. The average loss of intensity within a 12-hour period $\langle R_\tau \rangle = \langle (I/I_0)_\tau \rangle$ can be estimated as (Bridle & Schwab 1999):

$$\langle R_\tau \rangle = 1 - 1.22 \times 10^{-9} \left(\frac{\theta}{\theta_{HPBW}} \right)^2 \tau^2 \quad , \tag{2.55}$$

valid for circular coverage with gaussian tapering, where θ is the angular separation from the phase center, θ_{HPBW} is the Half Power Beam Width of the primary beam, and τ the time average interval in seconds. At the edge of the primary beam $\theta/\theta_{HPBW} = 0.5$.

2.2.9 ALMA and the WSU

The Atacama Large Millimeter/sub-millimeter Array (ALMA) is an interferometer located in the Chajnantor plateau of the Atacama desert in Chile. It is funded by a partnership comprising: the member states of the European Southern Observatory (ESO), the USA, Canada, Japan, Taiwan, and South Korea; in collaboration with the Republic of Chile.

ALMA consists of a large array of 12-meter antennas, and a compact array of 12-meter and 7-meter antennas. ALMA operates in a frequency range between 35 and 950 GHz ($0.3 - 8.5$ mm) divided into 10 receiver bands. Band 1 was made available in March 2024 on the 12-meter array only, while Band 2 is currently under development. The antennas on the large array are placed on pads and can be rearranged in different configurations with two specialized transporters. In their most extended configuration the antennas have a maximum baseline of about 16 km, resulting in an angular resolution of 20 mas at 230 GHz.

ALMA is currently undergoing a major upgrade of its hardware and software dubbed the Wideband Sensitivity Upgrade (WSU), with full implementation expected by 2030. The main goal of the WSU is to increase the available instantaneous bandwidth, i.e. the range of frequencies that can be observed simultaneously. This will significantly increase continuum sensitivity as per an increase of $\Delta\nu$ in (2.53), but it will also improve the signal-to-noise ratio (SNR) of spectral lines, since a single large frequency range can be covered at once instead of splitting the available observing time into n small-bandwidth observations. This increases Δt for the same channel $\Delta\nu$ in (2.53), for a improvement in σ_{rms} of a factor $\sqrt{n}$.

To achieve these goals ALMA is being fitted with new broadband receivers, as well as upgraded electronics, and a more powerful correlator in order to handle the increase in data size. New software for data handling and processing is also under development. Band 2, currently undergoing installation and testing, will be the first ALMA band belonging to the WSU. It will cover the frequency range of Band 3 and extend it down to 67 GHz.

Figure 2.6: Four ALMA antennas under a view of the Milky Way. Credit: J.F. Salgado (ESO).[3]

---

[3] `https://www.almaobservatory.org/en/four-alma-antennas-on-the-chajnantor-plain/`

2.3 Tools for spectral analysis of protoplanetary disks

2.3.1 Shifting and stacking spectra with *GoFish*

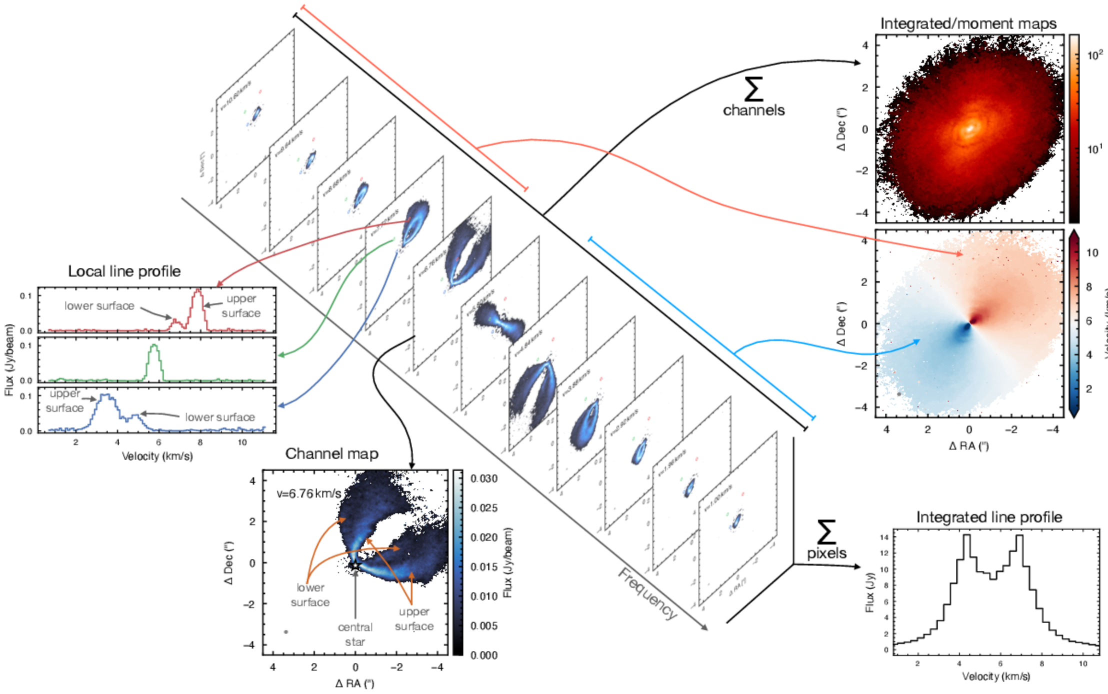


Figure 2.7: A typical protoplanetary disk molecular data cube, here using the ^{12}CO $J = 2-1$ ALMA observations of HD 163296. The figure shows how channel maps can be extracted at a given velocity (notice the *butterfly pattern* taking shape in each channel map); how local and global line profiles can be obtained by integrating spatially the flux over a selected region; how moment maps are computed by integrating along the velocity axis. Credit: Pinte et al. (2022).

Figure 2.7 shows a typical continuum-subtracted datacube containing molecular emission in a protoplanetary disk. Since protoplanetary disks are rotating systems, any emission coming from them will be spread along the frequency axis, due to the different velocity of the gas with respect to the systematic velocity of the object at different radii and and azimuthal angles. This is best visualized when viewing channel maps of the emission at different frequencies. As frequency increases the emission appears first on the redshifted side of the disk and gradually moves towards the blueshifted side. This results into a characteristic *butterfly pattern*. On a spectrum, the resulting line appears broadened and split into a redshifted and blueshifted peak.

The spread-out emission that results from the (near) keplerian rotation of the gas can translate into weak lines and a low SNR, which is why a crucial step in the analysis of protoplanetary disk spectra is the process of shifting and stacking. By exploiting the known keplerian profile of a rotating disk one can gather all emission pertaining to the same line, shift it to a common line center, and stack it to obtain a single-peaked narrower line centered at the systematic velocity of the disk. This technique was first described by Yen et al. (2016) and it serves a dual purpose: (*i*) to greatly increase the signal-to-noise ratio of detected lines, and (*ii*) to detect otherwise invisible lines, whose emission is too faint and too dispersed along the velocity axis to emerge above the noise.

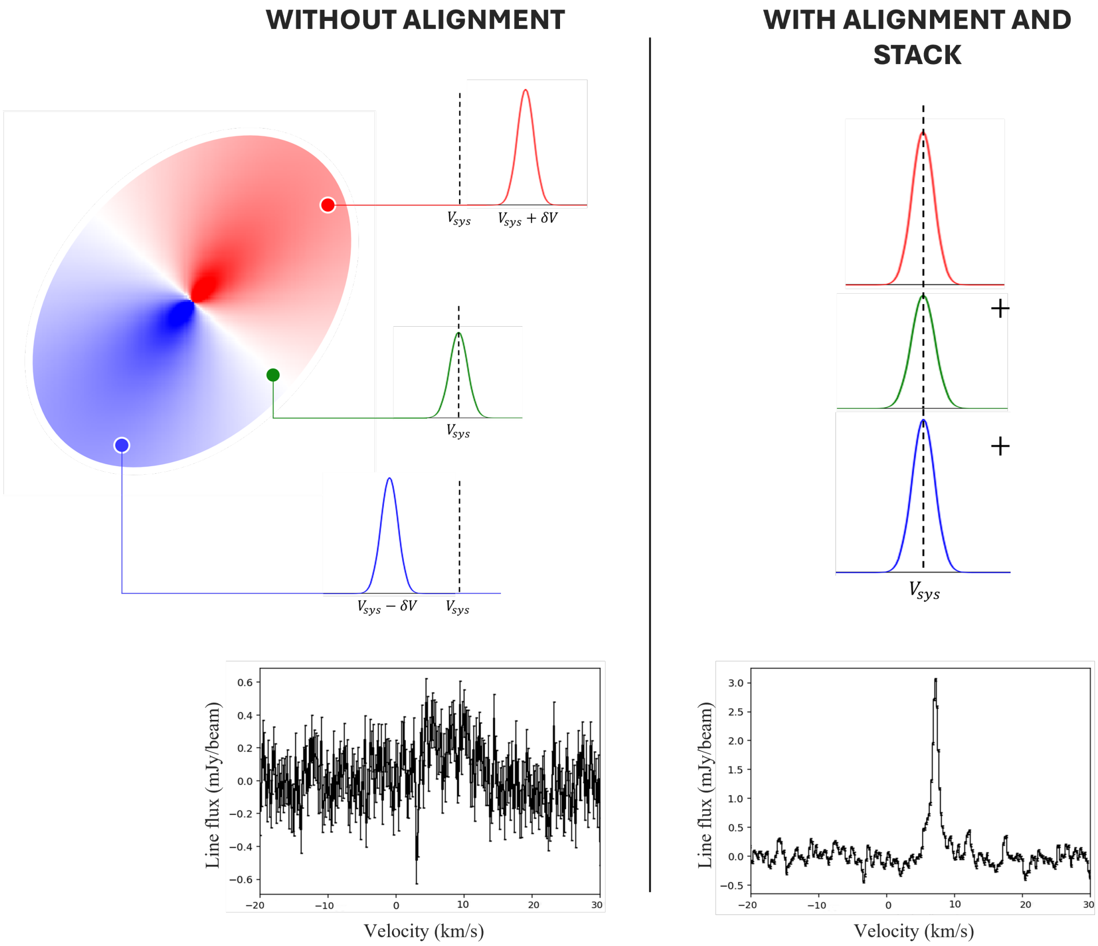


Figure 2.8: Sketch outlining the basic idea behind *shift-and-stacking*. A molecular line will appear red- or blueshifted from the systematic velocity depending on where on the keplerian disk the spectrum is extracted. This leads to dispersed, double peaked spectrum when averaging over the full disk (left panel). If the lines are shifted to a common velocity center and stacked before averaging, the resulting spectrum features a narrow line with boosted SNR (right panel). Figure recreated based on Fig. 1 in Yen et al. (2016). The example spectra are from this work ($^{12}C^{18}O$, see Chapter 3).

For the purpose of easily and reliably applying this technique, Teague (2019) developed a Python module called *GoFish*. It contains tools for shifting and stacking spectra, as well as other functionalities useful for spectral analysis. The two functions from the package prominently used in this work are:

- `average_spectrum` – this function takes the mass M_* of the host star, the radius of the gaseous disk over which to perform the operation, the inclination angle i of the disk, the distance of the object, and the position angle (PA) of the disk. The latter is defined as the angle from north to the red-shifted major axis of the disk in an easterly (anti-clockwise) direction. Using these parameters it reconstructs the keplerian velocity field of the disk as:

$$v_0(r,\phi) = \sqrt{\frac{GM_* r^2}{(r^2+z^2)^{3/2}}}\,\cos\phi\ \sin i\ , \tag{2.56}$$

  and accordingly shifts the spectra in concentric annuli. The spectra are then combined, weighted by the area of each annulus, returning a boosted and single-peaked line in units of intensity (Jy beam^{-1}).

- `integrated_spectrum` – this function takes the same parameters as `average_spectrum`, and performs the same keplerian correction, but additionally integrates spatially the shifted spectra over the area of the disk, thus returning a spectrum in units of flux density (Jy). Alternatively, if no parameters for the determination of the keplerian correction are provided, the function performs a simple spatial integration over the disk area.

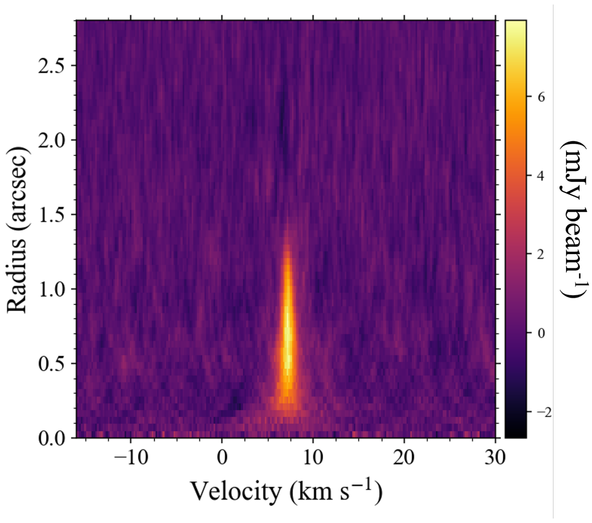


Figure 2.9: An example of a teardrop plot, showing emission form a $^{13}C^{16}O$ line (this work, see Chapter 3). The plot shows how the bulk of the emission is contained within $\sim 1.5''$, and how the velocity spread is larger close to the center of the disk (faster rotating gas).

The above functions assume by default a razor-flat disk, however GoFish also allows for a 3D disk model to be implemented, though this functionality was not used in this work.

GoFish also contains tools for extracting spectra from a pixel, and for constructing radial intensity profiles. A useful way to visualize keplerian emission in the disk is through a so-called *teardrop plot* (see Figure 2.9). A teardrop plot shows intensity as a function of radius and velocity. If the velocity spread of the line is indeed keplerian, the teardrop plot will be well defined within a velocity interval and a maximum radius. A noisy teardrop plot indicates that the emission is not behaving as expected from a keplerian velocity distribution, which might indicate that the emitting gas has a non-keplerian motion.

2.3.2 Moment maps using *Bettermoments*

Figure 2.7 also shows how a cube can be collapsed into moment maps. Different orders of moment maps can be produced, dubbed moment 0, moment 1, moment 2, etc., however for the purposes of this work only the first two moments are of interest:

- **Moment 0 maps**, or intensity integrated maps, are produced by integrating the intensity in each pixel along the spectral axis of the cube. Numerically, the resulting integrated intensity in each pixel is calculated as:

$$M_0 = \sum_i^N I_i \, \Delta v_i \quad , \tag{2.57}$$

  where N is the number of channels that are being collapsed together, and I_i, Δv_i are the intensity and spectral width of the i^{th} channel.

- **Moment 1 maps** show the average velocity of emission weighted by the intensity. Numerically:

$$M_1 = \frac{\sum_i^N I_i \, v_i}{\sum_i^N I_i} \quad , \tag{2.58}$$

  where N is the number of channels that are being collapsed together, and I_i, v_i are the intensity and velocity of the i^{th} channel.

Moment maps are affected by the noise in each channel, which is why it is often necessary to apply channel selection, summing only over the channels containing useful emission; or threshold clipping, where pixels below a given threshold are removed from the map.

To produce moment maps, the Python package *Bettermoments* was used in this work. It was developed by Teague & Foreman-Mackey (2018) and allows the user to create moment maps with simple commands. It includes functionalities such as spectral smoothing of the map, channel selection, threshold clipping, and the possibility to apply custom masks.

2.3.3 Additional tools

CARTA

The Cube Analysis and Rendering Tool for Astronomy (CARTA, Comrie et al. 2024) is a software for visualizing and analyzing astronomical images, designed to handle the large datacubes produced by ALMA, VLA, and the Square Kilometer Array (SKA).

CARTA was used in this work for the visual analysis of datacubes, moment maps, and spectra, as well as for defining regions for spectral extraction.

Supporting Python packages

The majority of python scripts used for this work were written and run on the open-source software Spyder IDE[4] running Python 3. Aside from the standard *NumPy* math package and *Matplotlib* for creating plots, specific astronomy packages, namely *Astropy* and *Spectral-cube*, were used for preliminary handling of datacubes and spectral extraction. Finally, the *lmfit* package was used for gaussian fitting of spectral lines and fitting of linear relationships.

---

[4] `https://www.spyder-ide.org/`

Chapter 3

Results and Analysis

3.1 Data reduction and imaging

3.1.1 The data

This thesis takes advantage of the largest dataset produced to date of any protoplanetary disk. The observations were carried out with the Acatama Large Millimiter/submillimiter Array (ALMA) under project code 2022.A.00036.S between 12/09/2023 and 27/09/2023, with antenna configuration C-8 corresponding to a maximum baseline of 8.5 km and a minimum baseline of 110 m. The science target is the protoplanetary disk of HL Tau.

The dataset was produced as a demonstration dataset meant to reproduce the standard bandwidth size expected after the Wideband Sensitivity Upgrade. It spans a total of 14.5 GHz in ALMA's Band 3, from 86.37 to 93.63 GHz, and from 108.37 to 115.63 GHz, for a central observing frequency of 101 GHz (3 mm), with an average spectral resolution of 0.18 km s^{-1}, and an angular resolution between $0''.09$ and $0''.20$. The dataset is split into 32 Observing Unit Sets spanning 0.46 GHz each with some overlap. Each Observing Unit Set is further split into four spectral windows 0.12 GHz wide.

A preliminary check of the uncalibrated data available in the ALMA archive[1] was performed through CARTA, and although some line emission could already be detected at this stage, it was noted that the low signal-to-noise ratio of the uncalibrated dataset, resulting from the uniform weighting, was limiting the analysis.

It was thus decided to request the calibrated measurement sets through the ALMA Regional Center, and re-image the data with CASA's `tclean` task, using natural weighting. The products of the calibration pipeline are made available for download by the ALMA Reginal Center, both before and after continuum subtraction. Both versions were downloaded for this work, in order to conduct an analysis of both spectral and continuum emission.

Of the 32 measurement sets, 30 were analyzed. In fact, two failed to download, and filing a new request to the ALMA Regional Center was not deemed necessary, seeing as the detections made from the downloaded sets were considered sufficient for this work, and no bright lines were expected to be contained in the missing datasets. Additionally, further data reduction and analysis would not have been productive under the project's time constraints. Therefore the final range of frequencies surveyed in this work is: 86.37 – 90 GHz, 90.46 – 91.36 GHz, 91.82 – 93.62 GHz, and 108.38 – 115.62 GHz.

---

[1] `https://almascience.eso.org/`

3.1.2 Line imaging

The 30 continuum-subtracted measurement sets were time averaged with CASA's `split` task in order to reduce their size and allow for easier handling and faster computational times. Keeping in mind the effects of smearing (see Section 2.2.8), a time bin of 30 s was chosen, corresponding to a loss of intensity at the edge of the primary beam much less than 0.001%.

The time-averaged datasets were run through the `tclean` task for imaging, with natural weighting and non-interactive auto-masking using `auto-multithresh`. The sub-parameters used for `auto-multithresh` are listed in Table 3.1. Their choice was based on the values suggested by the CASA Guides tutorial[2], with some modifications applied through trial and error in interactive mode to ensure all prominent emission was being masked. A global stopping criterion was set with the parameter `nsigma` set to 2.0. This sets a lower flux threshold for the cleaning algorithm at $2\sigma_{rms}$ where σ_{rms} is the average rms noise of the image.

Through the `tclean` task the datacubes were also trimmed to an area of 300×300 px^2 ($6.3'' \times 6.3''$) centered on the continuum peak of HL Tau, and rebinned in the spectral axis to a velocity resolution of 0.2 km s^{-1}.

Parameter		Value
`usemask`	=	`'auto-multithresh'`
`noisethreshold`	=	`4.0`
`lownoisethreshold`	=	`1.5`
`sidelobethreshold`	=	`2.0`
`minbeamfrac`	=	`0.2`
`negativethreshold`	=	`7.0`

Table 3.1: The `auto-multithresh` sub-parameters used for auto-masking in `tclean`.

3.1.3 Continuum imaging

The 30 non-continuum-subtracted measurement sets were channel-averaged with a bin width of 480 channels, and time averaged with a time bin of 30 s. The estimated intensity loss from smearing due to channel averaging is 1%.

The time- and channel-averaged datasets were run through `tclean` using multi-frequency synthesis mode, which results in an single channel output image. A single image was produced for each dataset by combining all four spectral windows. The same cleaning and auto-masking parameters, as well as image size, were used as those adopted for line imaging (Section 3.1.2).

---

[2] `https://casaguides.nrao.edu/index.php/Automasking_Guide_CASA_6.6.1`

3.2 Continuum emission of HL Tau

3.2.1 Correcting for flux calibration errors

Before proceeding to continuum and spectral analysis it is important to check, and if necessary correct, for errors in the flux density calibration process. As discussed in Section 2.2.4, these errors are expected to contribute to an overall uncertainty in the measured flux densities of about 5 – 10%. Large deviations well above this value are however possible and must be corrected for if present.

From each of the 30 continuum images the total flux density in Jy was extracted from an elliptical region centered on the peak of the continuum, with a semi-major axis of $0''.80$ and a semi-minor axis of $0''.56$. These values were chosen on account of the radius of the continuum reported by Yen et al. (2019a) at 1.1 mm and assuming an inclination angle of 45.5°. With the flux densities thus extracted, the continuum SED can be plotted as flux versus the central frequency of each continuum image. The SED is expected to follow a power law ν^{α}, which can be fitted with a straight line in logarithmic scale.

The SED and linear fit are shown in Figure 3.1. The majority of continuum fluxes lie within $\sim 5\%$ of the best fit line, however three points deviate significantly, with the associated fluxes being underestimated by more than 14%.

The deviating points were thus excluded from the fit as they were skewing it to lower values, and the fit repeated (this final fit is the one shown in Figure 3.1). A correction factor was obtained for each point as the ratio $F_{\nu,\,fit}/F_{\nu}$, where $F_{\nu,\,fit}$ is the expected flux obtained from the linear fit at a frequency ν, while F_{ν} is the observed flux at the same frequency. The correction factors can then be applied to each datacube to even out the discrepancies. The correction was only applied for some specific analysis that required it the most, which will be mentioned in later sections.

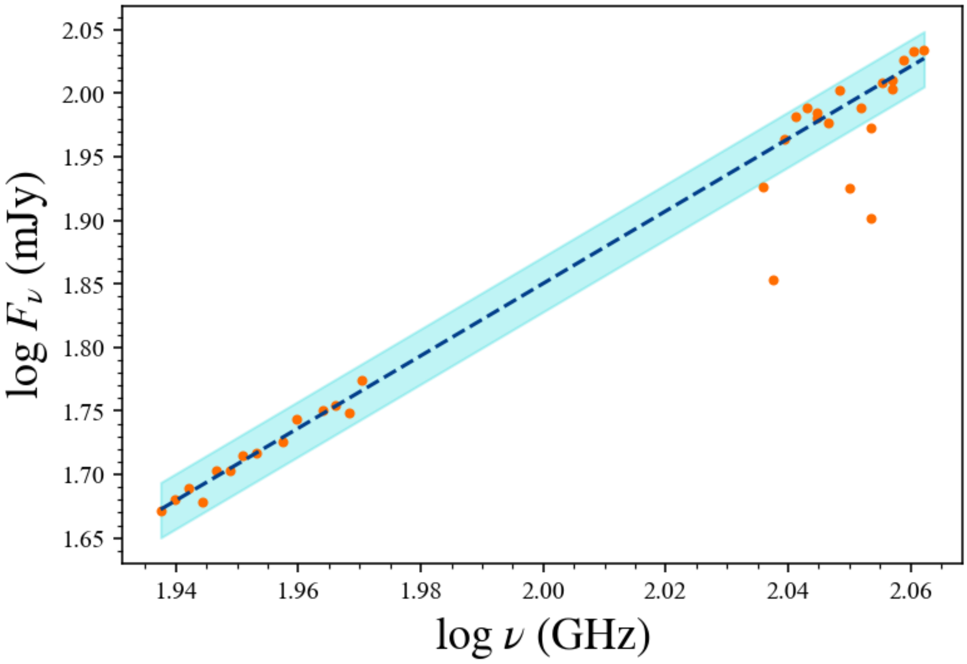


Figure 3.1: SED of the continuum emission of HL Tau in the frequency range 86 – 116 GHz. *Orange dots*: total flux density of the continuum as a function of frequency, collected within an ellipse centered on the continuum peak with a semi-major axis of $0''.80$ and a semi-minor axis of $0''.56$. *Blue dotted line*: power law fit excluding the three most deviating points. *Blue shading*: area within 5% of the best fit line.

3.2.2 Wideband imaging and spectral index map

Using Multi-Term Multi-Frequency Synthesis in CASA (`tclean` task, with `specmode='mfs'` and `deconvolver='mtmfs'`) an image of the continuum was produced from the full set of visibilities. It was obtained by combining all of the time-averaged and channel-averaged continuum measurement sets, excluding those corresponding to the three most deviating points in the SED fit (Figure 3.1). The MT-MFS was performed with a first order Taylor polynomial model of the sky spectrum (`nterms=2`). The final image (zeroth component of the Taylor polynomial) is shown in Figure 3.2.

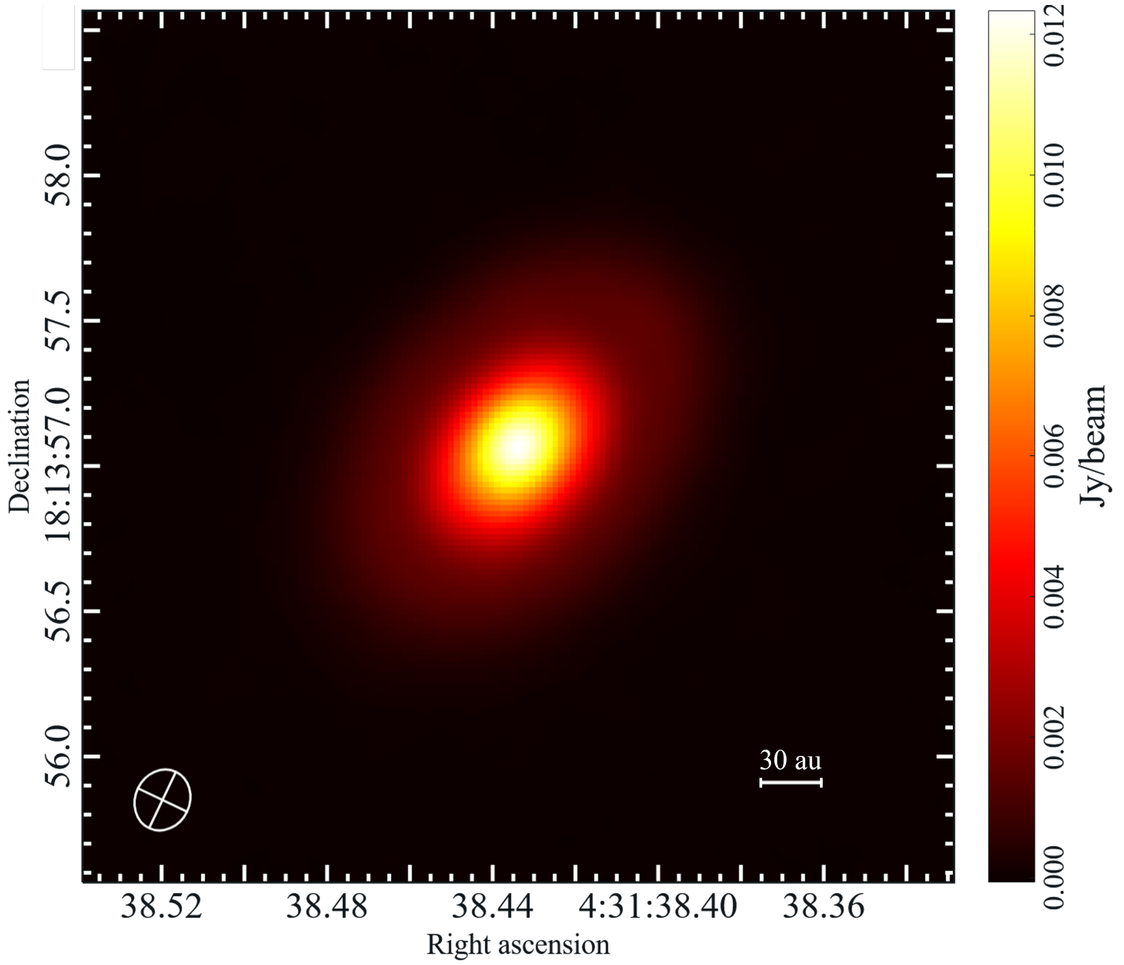
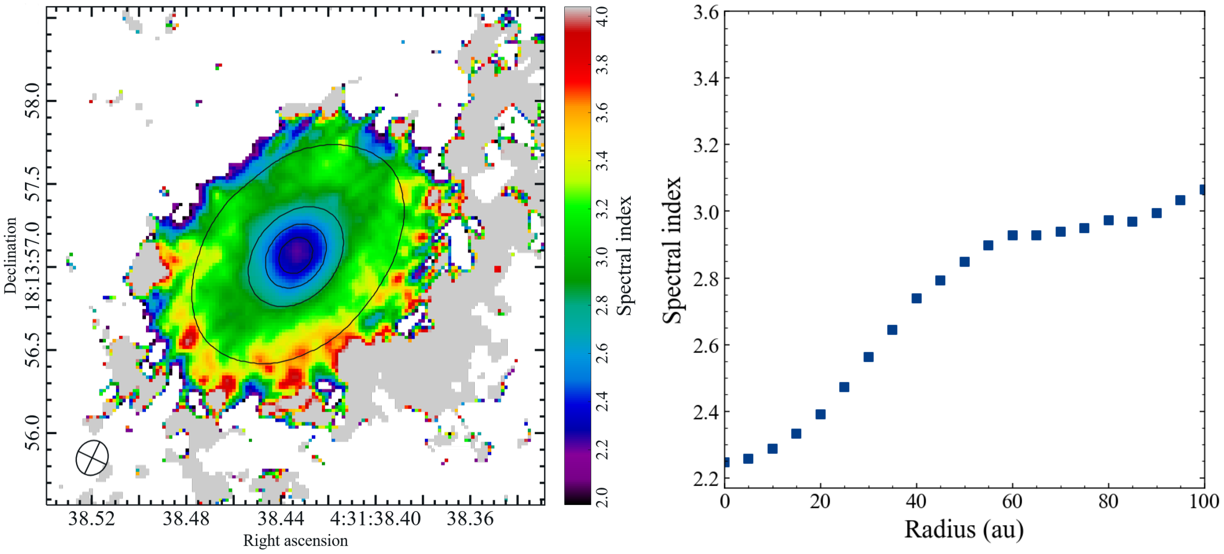


Figure 3.2: Total continuum image of HL Tau made from visibilities in the range 86 – 94 GHz and 108 – 116 GHz (central wavelength of 3 mm). σ_{rms} = 15 μJy. RA and Dec are J2000.

MT-MFS also produces a map of the spectral index from the ratio of the two coefficient images in the Taylor polynomial. This map is shown in Figure 3.3 along with the spectral index radial profile. Note that the latter was obtained from a limited number of pixels in the spectral index map, only considering those contained in a narrow (3 px wide) strip running along the major axis of the disk. This results in a very rough approximation of the radial profile of the spectral index, but allows for projection effects to be ignored. Additionally, the values of spectral index were averaged in radial bins of 5 au.



(a) **(b)**

Figure 3.3: Spectral index of the HL Tau continuum in the range 86 – 116 GHz (3.5 – 2.8 mm). (a) Spectral index map obtained with MT-MFS. Black contours and beam refer to the continuum image (Figure 3.2). Contour levels are: 50, 250, 450, and 650 times the continuum rms of 15 μJy. The semi-major axis of the largest contour line corresponds to a disk radius of ~ 109 au, that of the smallest contour line to a radius of ~ 16 au. RA and Dec are J2000. (b) Radial profile of the spectral index, obtained from a 3 px wide strip along the major axis of the spectral index map, and with values averaged in radial bins of 5 au.

As expected the spectral index reveals a dust continuum emission which is optically thick ($\alpha \sim 2$) in the densest inner radii of the disk, and moderately optically thin ($\alpha \sim 3$) at larger radii.

The result is consistent with the VLA analysis by Carrasco-González et al. (2016, Fig. 3d). Despite being a very quantitative estimate, the radial profile in this work's Figure 3.3b computed between 3.5 and 2.9 mm fits well enough between the radial curves Carrasco-González et al. found in the ranges 7.0 – 2.9 mm and 2.9 – 1.3 mm.

3.3 Molecular lines in the HL Tau system

3.3.1 Detection and spatial distribution of the lines

A total of sixteen rotational lines were confidently detected in the imaged data, pertaining to nine different linear molecules, namely: $^{12}C^{16}O$, $^{13}C^{16}O$, $^{12}C^{18}O$, $^{12}C^{17}O$, HCO^{+}, CN, HCN, HNC, and SO. Table 3.2 summarizes the detections.

Of the detected lines, those belonging to CN, HCN, and HNC are absorption lines, the rest are emission lines, though a few of the latter feature partial absorption (this will be further discussed in the following sections). All lines are from rotational transitions at the ground vibrational state ($v = 0$). Multiple lines from CN and HCN were detected from their hyperfine ladder.

Intensity integrated maps (moment-zero) and velocity maps (moment-one) from the reimaged datacubes were produced to investigate the spatial distribution of the detected molecules. They were obtained by combining only the channels that contained the line, and a clipping of $3\times\sigma_{rms}$ was applied to each.

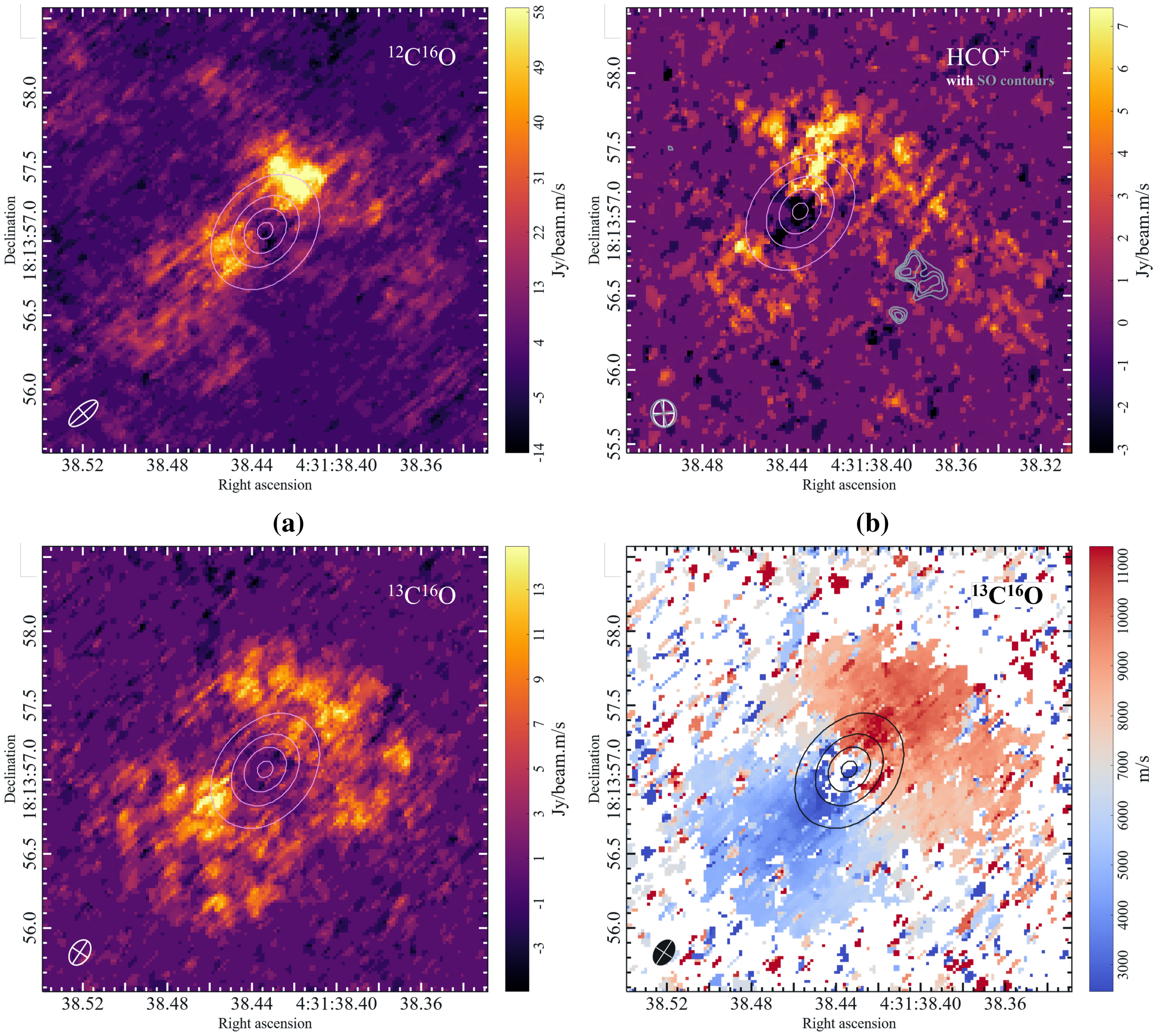
Molecule	Transition	ν_{rest} (GHz)
	Emission	
$^{12}C^{16}O^{*}$	1 – 0	115.27120
$^{13}C^{16}O^{*}$	1 – 0	110.20135
$^{12}C^{18}O$	1 – 0	109.78217
$^{12}C^{17}O$	1 – 0	112.35928
HCO^{+*}	1 – 0	89.18852
SO	3 2 – 2 1	109.25222
	Absorption	
CN	1 1 1 – 0 1 2	113.14416
CN	1 1 2 – 0 1 1	113.17049
CN	1 1 2 – 0 1 2	113.19128
CN	1 2 2 – 0 1 1	113.48812
CN	1 2 3 – 0 1 2	113.49097
CN	1 2 2 – 0 1 2	113.50890
HCN	1 1 – 0 1	88.63042
HCN	1 2 – 0 1	88.63185
HCN	1 0 – 0 1	88.63394
HNC	1 – 0	90.66357

Table 3.2: Summary of the detected lines, associated species, and rest frequencies ν_{rest}. Asterisks denote partially absorbed emission lines. All transitions are at the ground vibrational state. Quantum numbers are reported in the following format: $N^{up}\ J^{up}\ F^{up}\ -\ N^{low}\ J^{low}\ F^{low}$, where $J = N + S$ and $F = J + I$, noting that half-integers are rounded up to the next integer. When $J = N$, J is reported rather than N. Values are from the CDMS catalogue (Endres et al. 2016).

CO and HCO^{+}

Figure 3.4 shows the intensity integrated maps of the two most abundant CO isotopologues: $^{12}C^{16}O$ and $^{13}C^{16}O$, as well as that of HCO^{+}, and the velocity map of $^{13}C^{16}O$. From a visual analysis of the maps, the bulk of the emission for HCO^{+} and the isotopologues of CO appears to come from the keplerian disk. This is most clear in the velocity maps such as that in Figure 3.4d, and is further demonstrated by an incomplete but recognizable *butterfly pattern* visible when scrolling through the datacube channels in CARTA.

However, from the intensity maps three features become apparent: (*i*) the presence of a protrusion of emitting material connecting to the North-West of the disk, best observed in the HCO^{+} map, compatible with the streamer first described by Yen et al. (2019a) and mentioned in Section 1.3; (*ii*) a brighter emission in the maps of $^{12}C^{16}O$ and HCO^{+} on the redshifted (North-Western) side of the disk, which could be explained by an increase in gas density as a consequence of the merging of the streamer with the rotating disk; and (*iii*) the presence, of absorbed, or fainter regions localized within the central areas of the disk and along its minor axis, resulting in the emission appearing as two lobes on either side of the disk separated by an absorbed stripe.



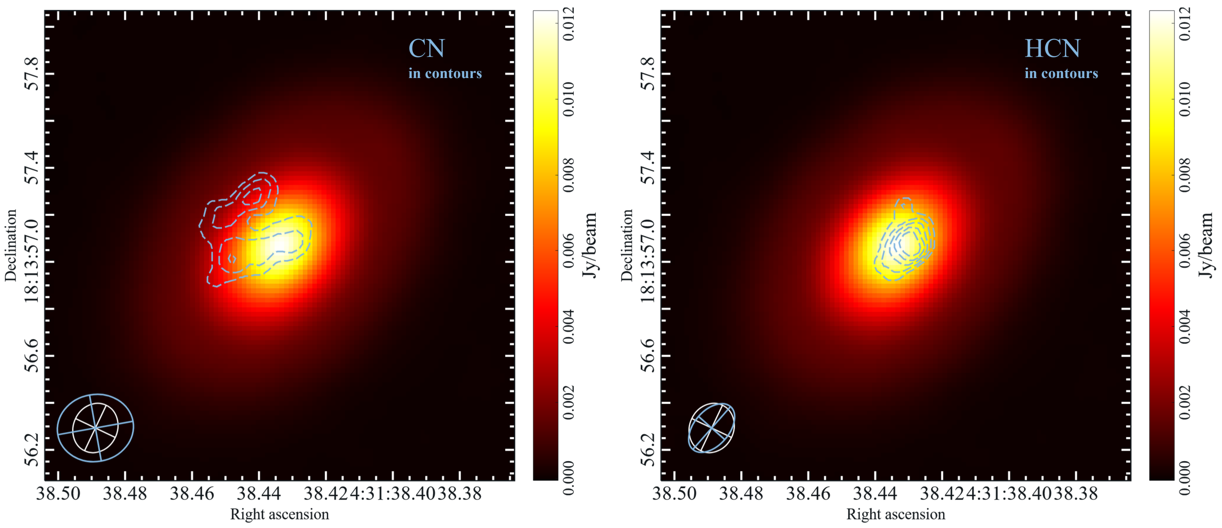
(a) **(b)** **(c)** **(d)**

Figure 3.4: Moment maps of molecular emission in HL Tau. (a) Moment-zero map of $^{12}C^{16}O$, $\sigma_{rms} = 3.44\,\mathrm{Jy\,beam^{-1}}$. (b) Moment-zero map of HCO^+ (color, $\sigma_{rms} = 0.70\,\mathrm{Jy\,beam^{-1}}$) and SO (green contours at 4, 5, 6, and $7 \times \sigma_{rms}$, where $\sigma_{rms} = 0.60\,\mathrm{Jy\,beam^{-1}}$). (c) Moment-zero map of $^{13}C^{16}O$, $\sigma_{rms} = 0.95\,\mathrm{Jy\,beam^{-1}}$. (d) Moment-one velocity map of $^{13}C^{16}O$. In each map, light blue contours represent the continuum emission at 150, 350, 550, and 750 times the continuum rms of $15\,\mu\mathrm{Jy\,beam^{-1}}$. The semi-major axis of the largest contour line corresponds to a disk radius of ∼ 59 au, that of the smallest contour line to a radius of ∼ 8 au. A $3 \times \sigma_{rms}$ threshold clip was applied to each image. RA and Dec are J2000.

CN, HCN, and HNC

Figure 3.5 shows the moment-zero maps of the combined absorption from three lines of CN, and three lines of HCN, in dashed contours overlaid on an image of the continuum. These molecules, together with HNC, are detected only in absorption against the backdrop of the disk continuum. The absorption only obscures the very center of the disk, where continuum emission is brightest.

As mentioned in Section 1.3, The disk of HL Tau is still embedded in a diffuse envelope, remnant of the pre-stellar core from which the system formed. Though not dense enough to hide the disk, molecules in the envelope can be responsible for the observed absorption.



(a) **(b)**

Figure 3.5: Moment maps of molecular absorption in HL Tau. (a) Moment-zero map of the combined absorption of three CN lines (the three at lower frequencies), in light blue dashed contours at −3, −4, and $-5 \times \sigma_{rms}$, with $\sigma_{rms} = 2.66\,\mathrm{Jy\,beam^{-1}}$. (b) Moment-zero map of the combined absorption of the three detected HCN lines, in light blue dashed contours at −5, −7, −9, −11, and $-13 \times \sigma_{rms}$, with $\sigma_{rms} = 0.50\,\mathrm{Jy\,beam^{-1}}$. Both are overlaid on the continuum image at 3 mm. A $3 \times \sigma_{rms}$ threshold clip was applied to each moment-zero map. RA and Dec are J2000.

The case of SO

Figure 3.4b shows the integrated intensity of SO in contours overlaid on the intensity map of HCO^+. SO is the only molecule whose emission is entirely localized within a small region to the West of the disk. As discussed in Section 1.3, Garufi et al. (2022) found similarly located SO and SO_2 emission at higher frequencies, which they associate with the streamer structure, and in particular with the slow shocks predicted to arise at the intersection between streamer and disk.

Disk, envelope, streamer

From the spatial analysis of the detected molecules three distinct regions of the HL Tau system can be identified, from which different molecules are observed to be emitting or absorbing: the keplerian disk, the surrounding envelope, and the infalling streamer.

While the molecules detected solely in absorption are easier to spatially separate from the rest, as belonging to the envelope, the same cannot be said for $^{12}C^{16}O$, $^{13}C^{16}O$, and HCO^+ which, while mostly emitting from the disk, are seen to be also tracing parts of the streamer. Additionally, the observed absorption of these species confined along the semi-minor axis of the disk could be explained by the same species absorbing from the envelope. Indeed, this absorption would occur only at the envelope systematic velocity, expected to correspond to the velocity of the keplerian gas along the minor axis of the disk, where its emission is neither redishfted, nor blueshifted by rotation. This leaves out the prominent absorption these species feature at the very center of the disk, where high velocity gas should not be absorbed by the envelope at the systematic velocity. Nevertheless the central absorption is explained by of the optically thick continuum (see Section 3.2.2)

The only detected molecule that can be associated with the streamer uniquely is thus SO, while the contributions of CO and HCO^+ from streamer and envelope cannot be isolated.

3.3.2 Extraction of the spectra

A spectrum for each of the detected lines was extracted. The areas of extraction were chosen differently for disk, envelope, and streamer molecules. Every spectrum had its x-axis converted to velocity, and was integrated over the area of extraction, so that the final spectra are in units of Jy, instead of Jy beam^{-1}.

Disk spectra

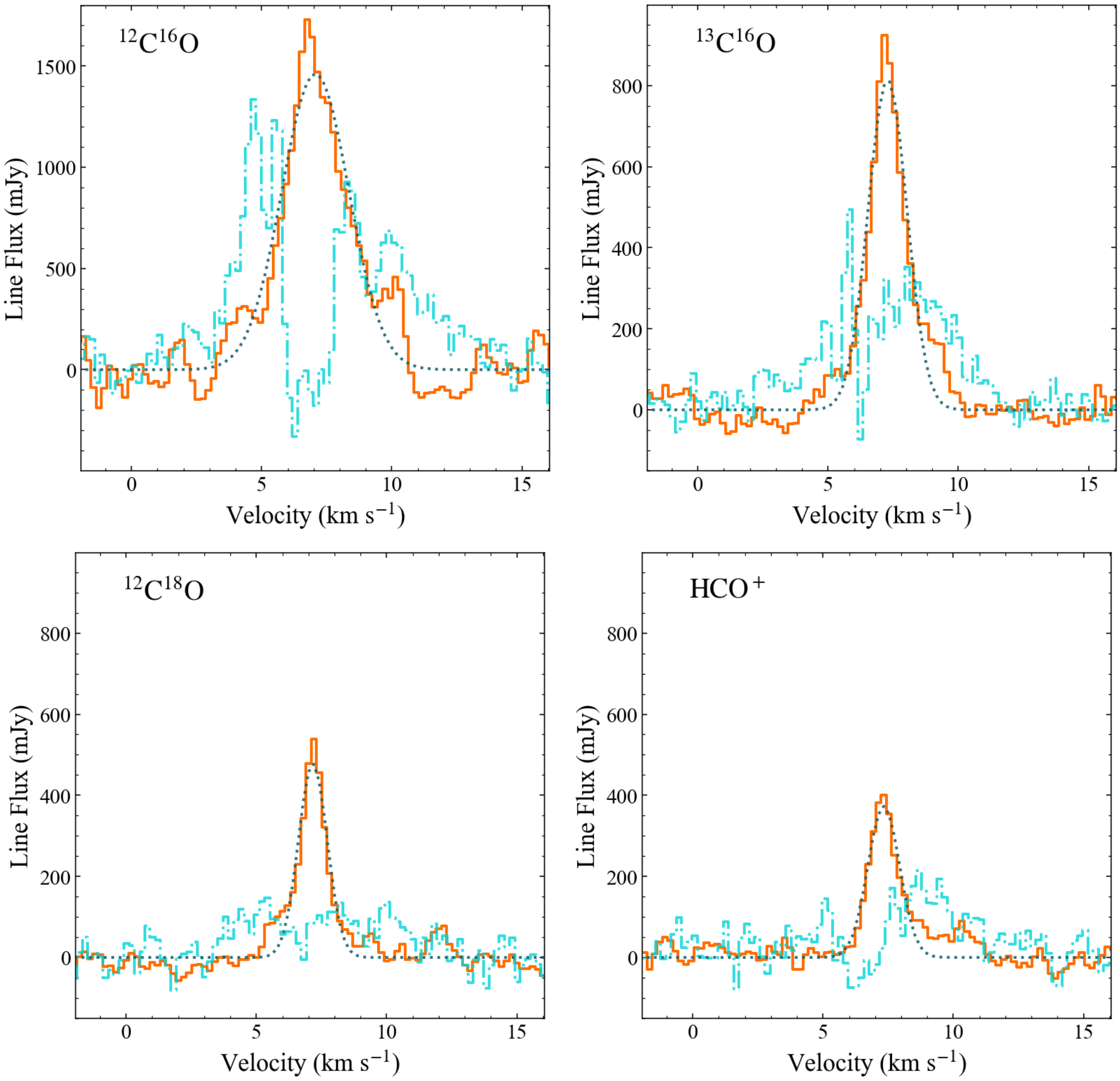


Figure 3.6: Spectra of the (1 – 0) transiton of $^{12}C^{16}O$, $^{13}C^{16}O$, $^{12}C^{18}O$, and HCO$^+$ in the HL Tau disk, integrated over an elliptical area centered on the peak of the continuum, with semi-major axis 1″.60 and semi-minor axis 1″.12, and PA = 318°. *Orange solid line*: integrated spectrum after shift-and-stack. *Light blue dots-and-dashes*: raw integrated spectrum before shift-and-stack. *Cyan dotted line*: gaussian fit. Parameters used for the shift-and-stack: $M_* = 2.1\,M_\odot$, $i = 45.5°$, $d = 140\,pc$.

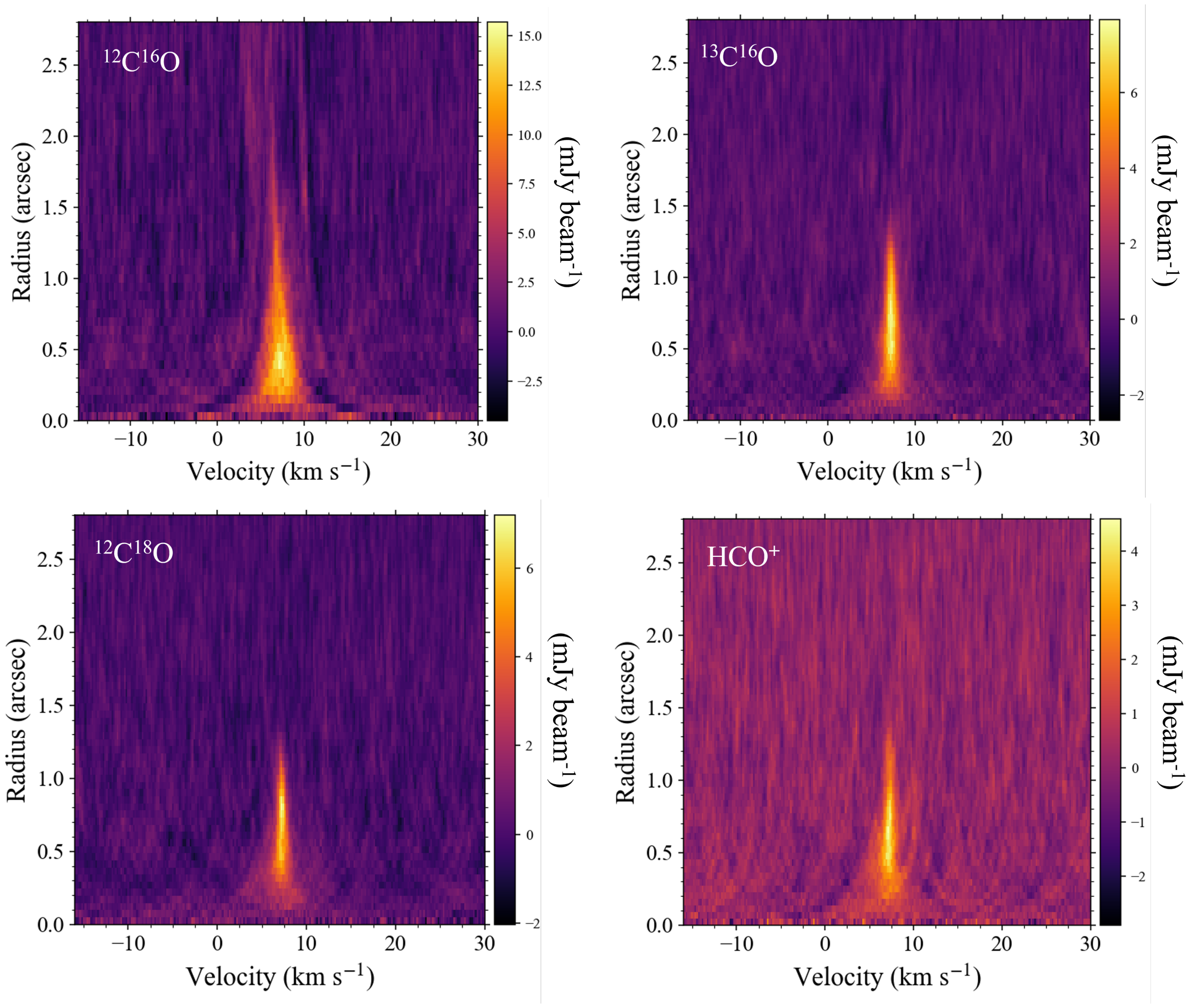


Figure 3.7: So-called *teardrop plots* of the (1 – 0) transiton of $^{12}C^{16}O$, $^{13}C^{16}O$, $^{12}C^{18}O$, and HCO^+ in the HL Tau disk. They show the radial profile and velocity spread of the disk emission after shift-and-stack. The emission was averaged in concentric radial bins within an elliptical area centered on the peak of the continuum, with semi-major axis $1''.60$ and semi-minor axis $1''.12$, and PA = 318°. Parameters used for the shift-and-stack: $M_* = 2.1\,M_\odot$, $i = 45.5°$, $d = 140\,pc$.

The spectra of $^{12}C^{16}O$, $^{13}C^{16}O$, $^{12}C^{18}O$, and HCO^+ were extracted from, and integrated over, an elliptical region centered on the peak of the continuum, with PA = 318° and with semi-major and semi-minor axes measuring $1''.60$ and $1''.12$ respectively, i.e. twice the value reported for the continuum by Yen et al. (2019a). These values were chosen so as to include the entirety of the gaseous disk, which is expected to extend to about twice the radius of the dust traced by the continuum. For these molecules only, two different spectra were produced: one before, and one after applying the shift-and-stack. Both are shown in Figure 3.6 for each molecule, together with the gaussian fit of the lines. So-called *teardrop plots* were also produced for each of these molecules and shown in Figure 3.7.

The shift-and-stack was performed with GoFish using the `integrated_spectrum` function over the region described above and with the following parameters: stellar mass $M_* = 2.1\ M_\odot$, disk radius $r_{max} = 1.6''$, distance $d = 140$ pc (Kenyon et al. 1994), inclination $i = 45.5°$, position angle PA = 318°.

The parameter reported with the most uncertainty in the literature is the inclination i of the disk, with the ALMA Partnership (2015) finding different values for each component of the ringed structure of HL Tau. The value of $i = 45.5°$ was adopted by varying an initial value of 45° by

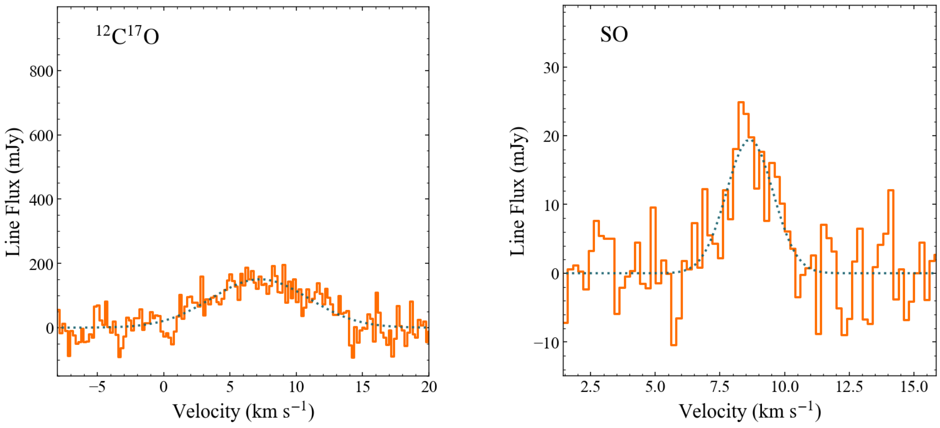
±0.5° until 9 total values were obtained, and by choosing the one which maximized the $^{12}C^{16}O$ integrated flux after shift-and-stack, and which produced the least noisy teardrop plot.

The spectra before shift-and-stack, shown in Figure 3.6 as light blue dots-and-dashes, reveal the spectral counterpart to the absorption visible in the spatial intensity maps. This is especially visible in $^{12}C^{16}O$, $^{13}C^{16}O$, and HCO^+ where the lines appear truncated by asymmetric absorption. Despite the truncated lines, GoFish is able to collect the keplerian emission either side of the absorption to a common velocity center (tough the absorption is likely to negatively influence the precision of this calculation, see Section 3.3.3) and stack it, resulting in the boosted signal shown as an orange solid line in Figure 3.6, and further confirming the disk origin of the bulk of the emission from these species. Notably, especially in the cases of $^{12}C^{16}O$ and HCO^+, some emission appears to have been left out of the shift-and-stack, appearing as smaller peaks on the wings of the lines. Yen et al. (2019a) observed similar wing emission in HCO^+ (3 – 2), which they attribute to contamination from ambient gas. The teardrop plots in Figure 3.7 help visualize the points made thus far. They function as checks for the goodness of the shift-and-stack, showing how well the program was able to associate the observed emission with a keplerian model. Non-keplerian contamination in $^{12}C^{16}O$ and HCO^+ is evident from these plots as well, in the form of emission around the main shape or protruding out of it.

As for $^{12}C^{17}O$, its spectrum is shown in Figure 3.8a with no shift-and-stack. In this case GoFish was not able to shift all of the emission to a common velocity. Instead the processed spectrum featured a double peak which is likely to be caused by the molecule's hyperfine structure, which is unresolved in the unprocessed spectrum but emerges when part of the emission is gathered and shifted to a different velocity. Because of the uncertainty in whether GoFish is able to handle a hyperfine structure, it was decided to proceed with the analysis of this molecule using its unprocessed spectrum.

SO spectrum

The spectrum of SO shown in Figure 3.8b was extracted from a circular region with a radius of 0″.2 centered on the brightest pixel in the moment-zero map, returning the integrated flux in Jy. The nature of this molecule's emission will be discussed in more detail in Section 3.4.



(a) **(b)**

Figure 3.8: (a) Spectrum of $^{12}C^{17}O$ (1 – 0) in the HL Tau disk, integrated over an elliptical area centered on the peak of the continuum, with semi-major axis 1″.60 and semi-minor axis 1″.12, and PA = 318°. (b) Spectrum of SO (3 2 – 2 1) in HL Tau, integrated over a circular area of radius 0″.2, centered on the brightest pixel in the SO intensity map. In both, the orange solid line indicates the integrated spectrum, the cyan dotted line indicates the gaussian fit.

Envelope spectra

The absorption spectra of CN, HCN, and HNC were extracted from a pixel corresponding to the continuum peak, where the absorption is expected to have the largest SNR, and converted to units of Jy. Their lines are shown in Figure 3.9 along with the gaussian fits.

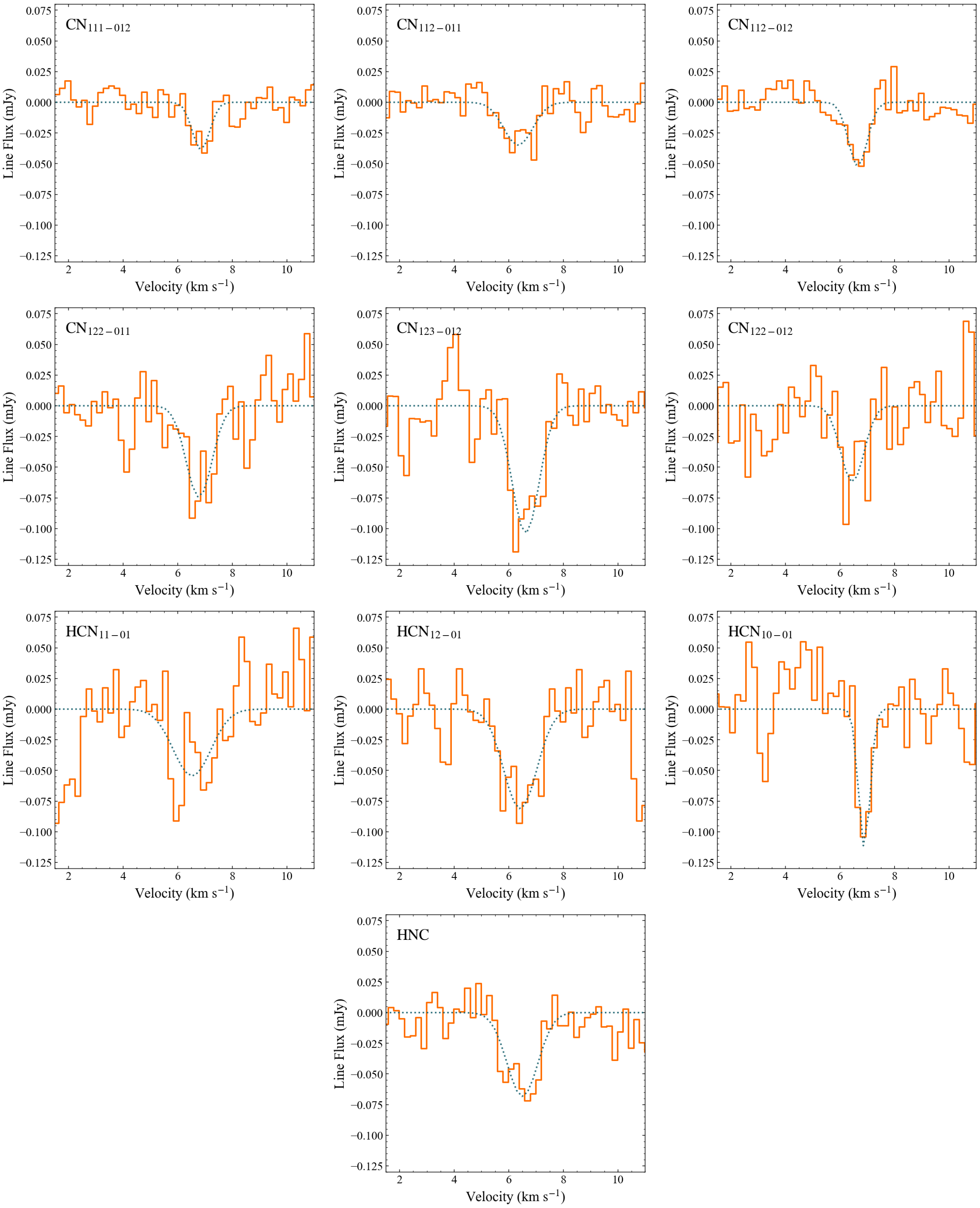


Figure 3.9: Absorption spectra of CN, HCN, and HNC in HL Tau, extracted from a pixel on the peak of the continuum. The cyan dotted line represents the gaussian fit.

A peculiar line among these is that corresponding to HCN (1 1 – 0 1), which appears to be split. No additional hyperfine transition is reported for this molecule, and the possibility of a split caused by a velocity structure is unlikely since the feature is absent in the other two lines of the same species. No other molecules with transitions at similar frequencies could be identified to be realistically responsible for the double line. Further observations with a higher SNR are needed, which should soon be achievable with ALMA Band 2, in order to either confirm the presence of the feature or discard it.

3.3.3 Line velocities

The systematic velocities obtained from the gaussian fits of each line are reported in Table 3.3, along with associated uncertainties. A general separation is evident between the average of these values in disk, envelope and streamer, as would be expected from three spatially distinct regions of the system. However, a large fluctuation of line velocities within envelope and disk which exceeds the estimated uncertainty is also noted. In the case of the disk there is indeed an additional source of error unaccounted for in the fit uncertainties, stemming from asymmetric line absorption hindering the correct identification of a common velocity center by GoFish during the shift-and-stack, because the program expects symmetric emission either side of the line velocity.

As for the envelope, the ambient medium around HL Tau has had a long history of studies attempting to untangle its complex kinematics comprising of infall, outflows, and rotation (Yen et al. 2017, Yen et al. 2019b). The lines detected in this work, tracing only the envelope right in front of the disk, have velocities blueshifted with respect to the systematic velocity of the disk, but still consistent with velocities found in the medium between 6 and 7 $\mathrm{km\,s^{-1}}$.

Finally, a velocity of 8.5 – 9 $\mathrm{km\,s{-}1}$ for SO is consistent both with higher-energy transitions of the same molecule (Garufi et al. 2022) and with the velocities of the streamer traced by $\mathrm{HCO^+}$ (3 – 2) (Yen et al. 2019a, Fig. 4).

3.3.4 Integrated line fluxes

Every line in the spatially integrated spectra was integrated over velocity to obtain integrated line fluxes in $\mathrm{Jy\,km\,s^{-1}}$ reported in Table 3.3. The integration window ΔV for each line was chosen to cover $3 \times std.$ either side of the gaussian fit center, where $std.$ is the standard deviation of the gaussian fit. The uncertainty on the integrated flux was calculated with the following method for each line. Six noisy spectra were extracted from six line-free regions around HL Tau, whose areas equaled the area of extraction for the line. The flux from these spectra was then integrated over 14 non-overlapping windows, of width equal to the integration width ΔV of the line. A total of $6 \times 14 = 84$ different integrated fluxes over noisy windows were thus obtained. The uncertainty was then obtained as the standard deviation of these integrated fluxes.

Investigating the values obtained for the line fluxes, the one corresponding to $^{12}\mathrm{C}^{18}\mathrm{O}$ stands out as severely underestimated, especially when compared to $^{12}\mathrm{C}^{17}\mathrm{O}$ which has a lower natural occurrence. Indeed, even if isotopic fractionation were taking place (see Section 1.2.3), it would not cause the inversion of the abundances of the two species. The reason for this inconsistency is likely to be attributed to a flux loss particularly evident in $^{12}\mathrm{C}^{18}\mathrm{O}$ but present in general in every observed disk emission, attributable to the particular ALMA configuration used in the observation, lacking short baselines and thus leading to a loss in received flux from extended emission. This, in light of the fact that $^{12}\mathrm{C}^{18}\mathrm{O}$ is expected to trace a more extended portion of the disk than $^{12}\mathrm{C}^{17}\mathrm{O}$, and combined with the additional loss of flux from the observed absorption, should explain the apparent inversion in the flux rations of the two species.

Molecule	Transition	σ_{rms} (mJy)	Line peak (σ_{rms})	V_{sys} (km s^{-1})	ΔV (km s^{-1})	Line Flux (mJy km s^{-1})
Disk						
^{12}C^{16}O	1 – 0	69.3	25.0	7.06 (0.03)	7.8	4999 (419)
^{13}C^{16}O	1 – 0	30.2	30.6	7.28 (0.02)	4.6	1603 (42)
^{12}C^{18}O	1 – 0	23.7	22.7	7.17 (0.02)	3.2	687 (35)
^{12}C^{17}O	1 – 0	45.9	4.3	7.26 (0.27)	22.0	1319 (132)
HCO$^+$	1 – 0	25.9	15.5	7.33 (0.02)	3.6	578 (39)
Envelope						
CN	1 1 1 – 0 1 2	0.012	3.5	6.85 (0.08)	1.8	−0.029 (0.008)
CN	1 1 2 – 0 1 1	0.012	4.0	6.35 (0.12)	3.2	−0.040 (0.010)
CN	1 1 2 – 0 1 2	0.012	4.4	6.65 (0.06)	2.2	−0.045 (0.008)
CN	1 2 2 – 0 1 1	0.025	3.7	6.78 (0.11)	2.8	−0.091 (0.022)
CN	1 2 3 – 0 1 2	0.025	4.8	6.62 (0.08)	3.0	−0.122 (0.023)
CN	1 2 2 – 0 1 2	0.025	3.9	6.45 (0.13)	2.6	−0.064 (0.022)
HCN	1 1 – 0 1	0.023	3.9	6.53 (0.18)	4.0	−0.080 (0.026)
HCN	1 2 – 0 1	0.023	4.0	6.42 (0.11)	3.6	−0.120 (0.025)
HCN	1 0 – 0 1	0.023	4.4	6.89 (0.03)	1.2	−0.054 (0.013)
HNC	1 – 0	0.017	4.2	6.50 (0.09)	3.4	−0.090 (0.017)
Streamer						
SO	3 2 – 2 1	4.8	5.2	8.65 (0.11)	5.4	44.5 (6.6)

Table 3.3: Summary of the line analysis. The columns show: molecular species; line transition; the rms noise σ_{rms} of each spectrum; the peak line flux as a multiple of the rms noise (when shift-and-stack was performed this number refers to the line peak *after* SNR boosting); the systematic velocity V_{sys} corresponding the center of the gaussian fit, with its uncertainty; the window of integration ΔV, defined as $3 \times$ std either side of the line, with std being the gaussian standard deviation; and the integrated line flux with its uncertainty.

Reports of line fluxes with which to compare these ones, for CO isotopologues and HCO$^+$ in the $(1-0)$ transition, could not be found in the literature. In order to verify the hypothesis of flux loss, then, additional observations would be needed, with shorter baseline configurations. Reports of the absorption line fluxes of CN, HCN, and HNC were also not found in the literature.

Unfortunately, ^{12}C^{18}O and ^{12}C^{17}O are also the only detected CO isotopologues with a low enough abundance to have optically thin lines, thus their unreliable line flux values prevents any analysis on the column densities and temperatures of these species and the derivation of any meaningful isotope ratio, which when compared to the values found in the ISM might have provided insight on chemical reprocessing in the disk and isotope fractionation.

3.4 Constraining the physical conditions of SO in the streamer

In their paper, Garufi et al. (2022) detected three SO lines in Band 9. Their detected transitions and related rest frequencies are reported in Table 3.4. They found emission located both in the inner disk, and within an outer region spatially compatible with the one traced by SO (3 2 – 2 1) in this work, and likely to coincide with the location of impact between streamer and disk. However, due to the energies of the transitions being very close to each other, Garufi et al. could not derive a reliable rotational diagram fit from the three SO detections. Instead, they used their SO_2 detections to derive an estimate of the gas temperature in the impact region, assuming it to be the same for the co-spatial SO.

Molecule	Transition	ν_{rest} (GHz)
SO	15 14 – 14 13	644.37892
SO	15 15 – 14 14	645.25493
SO	15 16 – 14 15	645.87592

Table 3.4: SO detections made by Garufi et al. (2022) in HL Tau with ALMA Band 9. Dataset from project id: 2017.1.01562.S. The table reports the transition and rest frequency of each detection. Quantum numbers are reported in the following format: $N^{up}\,J^{up}\; - \;N^{low}\,J^{low}$. Values are from the CDMS catalogue (Endres et al. 2016).

Using the lower energy transition detected in this work it is possible to complement the detections of Garufi et al. and derive an independent estimate of the temperature of the gas in the streamer impact region through SO rather than SO_2, as well as a measurement of the total column density of the molecule.

3.4.1 Obtaining upper column densities and other parameters

The estimate of column density and temperature was carried out making use of the rotational diagram method, under the assumptions of optically thin lines and LTE. Recall from Section 2.1.6 the rotational diagram equation (2.33) and the expression for the upper column density (2.30) in the case of optically thin emission:

$$\ln \frac{N_u}{g_u} = \ln \frac{N_{tot}}{Q(T)} - \frac{E_u}{k_B T} \quad , \tag{2.33}$$

$$N_u = \frac{8\pi k_B \nu^2}{hc^3 A_{ul}} W \quad . \qquad \text{(2.30, with } \tau_\nu \ll 1\text{)}$$

Computing the upper level column density N_u for each transition involved is the first step to construct a rotational diagram. It is derived from $W = \int T_b\, dv$, and can be obtained from the measurements of each line's integrated flux by taking the integral in dv on both sides of (2.18):

$$\int T_b\, dv = \frac{c^2}{2k_B\, \nu^2\, \Omega} \int S_\nu\, dv \quad , \tag{3.1}$$

where $\int S_\nu\, dv$ is the line integrated flux with appropriate unit corrections, such that in cgs the units on both sides of (3.1) are $\mathrm{K\,cm\,s^{-1}}$. Note that ν^2 could be taken out of the integral by approximating it as constant within the width of a line.

The integrated line flux for SO (3 2 – 2 1) as reported in Table 3.3 was used after applying the appropriate unit conversion. Additionally, in order to derive an accurate rotational diagram, the flux was corrected for the effects of calibration error, according to the results detailed in Section 3.2.1. As for the other transitions, since maintaining the same area of extraction is essential, the datacubes used by Garufi et al. were acquired and integrated spectra were extracted from the same circular region defined in Section 3.3.2. The lines were then integrated over velocity in the same way that was described in Section 3.3.4. The sky area of extraction Ω can be computed as πr^2, where $r = 0''.2$ converted to radians.

N_u additionally requires the knowledge of the Einstein coefficient A_{ul}, which were obtained for each transition, along with the value of E_u, from the Leiden Atomic and Molecular Database (Schöier et al. 2005, accessed 09/2025), while the upper state degeneracies g_u were obtained from the Cologne Database for Molecular Spectroscopy (Endres et al. 2016). These parameters are reported for each SO transition in in Table 3.5, along with the computed values of W and N_u. Computations were done using cgs units in order to obtain column densities in cm^{-2}.

Transition	$\log A_{ul}$ (s^{-1})	g_u	E_u/k_B (K)	$\log W$ (K cm s^{-1})	$\log N_u$ (cm^{-2})
$\mathrm{SO}_{3\,2-2\,1}$	−4.967	5	21.1	6.7172	14.049
$\mathrm{SO}_{15\,14-14\,13}$	−2.452	29	254.2	5.9457	12.305
$\mathrm{SO}_{15\,15-14\,14}$	−2.450	31	260.9	5.8952	12.387
$\mathrm{SO}_{15\,16-14\,15}$	−2.446	33	252.6	6.0288	12.251

Table 3.5: Parameters used for the rotational diagram of SO. Columns are: SO transition; base-10 log of the Enistein A_{ul} coefficient; upper level degeneracy; energy of the upper level in K, base-10 log of the integrated temperature brightness of the line; base-10 log of the upper level column density. A_{ul} and E_u are from LAMDA (Schöier et al. 2005), g_u is from CDMS (Endres et al. 2016).

3.4.2 Column density and temperature from the rotational diagram

Using the parameters listed in Table 3.5, the quantity $\ln(N_u/g_u)$ was derived and plotted for each SO transition versus E_u. Errors were propagated from the uncertainties on the line integrated flux, itself computed with the method outlined in Section 3.3.4, but with a number of integration windows lowered at 11 (intead of 14) to account for the smaller overall size of the spectra in the datacubes used by Garufi et al. An additional flux relative uncertainty of 10% for the SO in this work, and 15% for the SO in Garufi et al. was added in quadrature to account for the uncertainty in calibration for Band 3 and Band 9 respectively. The uncertainties on the physical constants h, c, k_B, as well as the parameters g_u, A_{ul}, and E_u, were assumed to be negligible.

A linear fit was performed from which, according to (2.33), the slope $m = -1/T$ and intercept $q = \ln(N_{tot}/Q(T))$ were derived. The rotational diagram is shown in Figure 3.10 . Because the three SO transitions from Garufi et al. lie in a very narrow energy range, the fit is almost entirely determined by SO (3 2 – 2 1) (square marker in Figure 3.10). As a result it is useful to determine a maximum and minimum slope and intercept associated with the maximum and minimum value in the uncertainty interval of $\ln(N_u/g_u)_{\mathrm{SO}_{3\,2-2\,1}}$. The maximum and minimum slope lines are shown in Figure 3.10 as dashed light blue lines.

A temperature estimate for the gas, assuming LTE and optically thin lines, is thus derived from the slope of the best fit line, an uncertainty interval can be limited by the maximum and minimum slopes described above. The values obtained are:

$$T_{\mathrm{SO}} = (40.4 \pm 1.2)\,\mathrm{K} \quad . \tag{3.2}$$

This result is indeed consistent with the value of 58 ± 19 K found by Garufi et al. using SO_2.

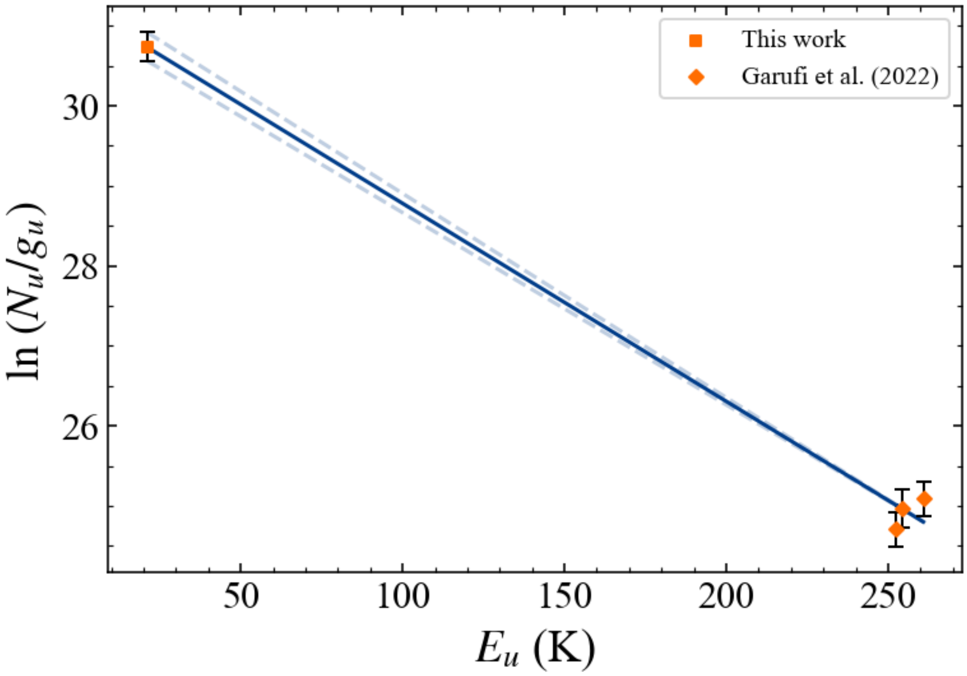


Figure 3.10: Rotational diagram of SO using the (3 2 – 2 1) transition detected in this work (orange square), and the (15 14 – 14 13), (15 15 – 14 14), and (15 16 – 14 15) transitions detected by Garufi et al. (2022) (orange diamonds). The solid blue line is the best fit line; the dashed light blue lines are the minimum- and maximum-slope lines based on the uncertainty interval of $\ln{(N_u/g_u)}_{SO_{3\,2-2\,1}}$.

As for the total column density of the gas N_{tot}, its derivation requires the knowledge of the partition function $Q(T)$ whose tabulated values can be found for a wide range of temperatures in various spectroscopy databases. The two temperature values closest to the obtained T_{SO} for which $Q(T)$ is reported by the CDMS are $T = 75$ K and $T = 37.5$ K. A linear interpolation in log-log space between the two was performed to obtain $Q(T_{SO})$, which was used with the rotational diagram intercept to finally compute the total column density $N_{tot} = e^q\, Q(T)$:

$$N_{tot,\,SO} = 3.6^{+0.9}_{-0.7} \times 10^{15}\ \text{cm}^{-2} \quad , \tag{3.3}$$

which, as expected, is consistent with the result of Garufi et al. ($(0.2-2)\times10^{16}$ cm^{-2}), which they found having assumed the same temperature as SO_2.

3.4.3 The origin of SO in relation to the streamer

The derived temperature of 40.4 K for the SO gas supports the idea that the observed enrichment of the gas phase in S-bearing species originates not from thermal, but rather from mechanical desorption, as a result of sputtering of the icy mantles of dust grains, driven by slow shocks in the region of impact between streamer and disk. Indeed, a temperature of 40 K is far below the sublimation temperatures of SO and SO_2, and if thermal desorption were taking place it should result in a uniform emission within the disk, instead of the observed localized areas of increased column density close to the streamer.

Chapter 4

Summary and Conclusions

4.1 Summary of results

The following is a summary of the results presented in Chapter 3.

1. The continuum SED of HL Tau in the frequency range 86 – 116 GHz was fitted with a linear relation, and three measurement sets were found to be deviating by more than 14%. Correcting factors were derived to compensate for these discrepancies. Their nature is likely to be attributed to errors during the calibration process, perhaps as a result of unaccounted for variable atmospheric effects. A follow-up investigation on the calibrators used for each measurement set and the time and date of observations would be needed to determine with more certainty the causes of the deviations.

2. A combined wideband continuum image of HL Tau was produced, at a central wavelength of 3 mm. A map of the continuum spectral index was also made, along with a plot of the spectral index radial profile. They show that the emission from the dust continuum is optically thick in the innermost $\sim$ 40 au, becoming moderately optically thin at larger radii, with a spectral index growing from 2 to 3 and plateauing at around 60 au, similarly to what was found in the literature at neighboring wavelength ranges.

3. Sixteen rotational lines were detected in HL Tau, belonging to nine different molecular species and originating from three distinct regions, namely: $^{12}C^{16}O$, $^{13}C^{16}O$, $^{12}C^{18}O$, $^{12}C^{17}O$, and HCO^+, emitting from the disk; CN, HCN, and HNC, absorbing from the diffuse envelope of HL Tau; and SO emitting from a localized region to the West of the disk, identified as part of the impact region between the spiral streamer feeding HL Tau and the disk itself. Intensity and velocity maps were produced which show the streamer traced by HCO^+, and heavy absorption in the innermost regions of the disk, likely a combination of the effects of the optically thick continuum and the absorbing envelope.

4. The spectra of each species were extracted, boosting when possible their SNR with shifting and stacking of the lines. The disk molecules feature lines which are truncated asymmetrically by absorption, and have some wing emission, symptom of contamination from the ambient medium. One absorption line of HCN features a double peak which is hard to explain, and which, if it is not an artifact of noise, might belong to an as-of-yet unidentified molecule.

   The line velocities reveal the difference in kinematics between disk, streamer and envelope. Their values are consistent with the literature, although their errors, derived from the uncertainties of the gaussian fits, are likely underestimated, as they fail to take into account asymmetric absorption affecting the shift-and-stack, and complex gas kinematics in the envelope.

Integrated line fluxes were derived, but flux loss due to the lack of short baselines and absorption caused their underestimation, especially in the case of $^{12}C^{18}O$ whose line flux results to be lower than the less abundant $^{12}C^{17}O$. Further observation are required with short baselines configurations to complement these results and gather more reliable line fluxes.

5. The detection of SO was complemented with three higher-frequency detections in the literature to produce a rotational diagram and derive more accurately the total column density and temperature of the molecule in the region associated with the streamer impact. The results are $T_{\mathrm{SO}} = (40.4 \pm 1.2)$ K, and $N_{tot,\mathrm{SO}} = 3.6^{+0.9}_{-0.7} \times 10^{15}$ cm^{-2}, comparable with measurements in the literature obtained from SO_2 in the same region, and consistent with a mechanically desorbed gas stemming from slow shocks in late accretion events.

4.2 General limitations

The dataset investigated with this thesis was acquired with the intention of reproducing the size that an average ALMA measurement set will have after the WSU, mainly for the purposes of testing the new data-handling infrastructure and pipelines. Therefore it was not produced with science objectives in mind, and this resulted in the observations having with very high angular and spectral resolution, which however greatly decreased the overall SNR. This caused many molecules that have been observed in HL Tau to not be detected in this work, and the lack of short baselines caused loss of extended emission.

Despite all this, the dataset was still worth investigating, as it is the first observation covering such a large frequency range available for any protoplanetary disk, and provides a demonstration of the limitations that will be overcome with the WSU.

4.3 The question of COMs in HL Tau

The detection and characterization of Complex Organic Molecules in a very young disk such as HL Tau would shed new light on the the chemical inheritance between early and late stage YSOs. The now established presence of dynamical events, such as late accretion, resulting in the release of species trapped in ices well below their sublimation temperatures, together with the detection of spatially resolved water lines, seem to point at HL Tau as the perfect candidate for COM detection.

Still, COMs remain elusive in HL Tau, and this work finds none in the frequency range covered by the analyzed dataset. Indeed, as outlined above, the high angular resolution of this dataset is the main contributor to the non-detection of faint lines. The new age of ALMA observations that the WSU will inaugurate might finally reveal COM lines too weak for the current available sensitivity, but if COMs remain undetected it might mean that, despite its young age, HL Tau is already too cold of a system, with the majority of its complex chemical budget already trapped in ices within its deeper layers.

4.4 Achievements and future prospects

The major takeaways of this thesis can be summarized as follows:

1. As for the advancements that WSU will bring to continuum observations, the preview dataset allowed, with its large frequency coverage, a derivation of the intra-band spectral index within the full Band 3. This measurement requires a large amount of observing times (here, it made use of 30 separate ones), but with the WSU the same-size dataset will be available with a single observation.

2. As for the role of late accretion events in the detection of otherwise frozen-out molecules, the assumptions made by Garufi et al. (2022) were confirmed, in regards to both (*i*) the origin of SO and SO_2, released from sputtering of icy grains caused by the slow shocks in the impact location between accretion streamer and disk, and (*ii*) the co-spatial nature of the two species that results in their gas temperatures being comparable.

3. As for the progress towards ever more complete COM detection in HL Tau and in disks in general, the high angular resolution of this preview dataset is not compatible with the sensitivity required for the detection of lines as weak as COM lines in disks. Even after the implementation of the WSU, compromises on the angular resolution will need to be made if any detection of COMs in HL Tau is to be attempted.

With the advent of the WSU, datasets the size of the one used in this work will become the norm, but with a much higher overall SNR. Already with the soon to be available Band 2, a large-bandwith observation of HL Tau can be repeated with higher sensitivity, which is likely to surpass the limitations of this work and provide a more complete survey of the chemical budget of HL Tau at millimeter wavelengths, granted that a lower angular resolution might still be necessary for any COM detection to be made.

An important point requiring further investigation is the characterization of CO isotopic ratios in the disk, to understand if phenomena of isotopic fractionation, such as isotope-specific photodissociation, are taking place at this stage of disk evolution. If more reliable line fluxes can be obtained, the column densities and temperatures of the optically thin isotopologues of CO, can be derived using rotational diagrams and higher energy transitions, widely detected in the literature (e.g. Booth & Ilee 2020, Zhang et al. 2020).

More targeted observations might be able to detect less abundant species that similarly to SO and SO_2 have been sputtered out of the solid phase by the streamer impact, beginning with other S-bearing species which have been observed in faster outflow shocks in multiple YSOs. Indeed there are a variety of molecules, including simple organics, that are thought to make up the sulfur reservoir in interstellar ices (e.g. CS, HCS, HSO, H_2S; Laas & Caselli 2019).

In general, more observations targeting disks of different age and different stellar mass are needed, especially those where conditions seem favorable for the release of molecular species from icy mantles, in order to compare the availability of COMs and their conditions, and contribute to draw a more complete picture of their evolution from the interstellar medium to planets.

References

ALMA Partnership, 2015, “The 2014 Alma Long Baseline Campaign: First Results From High Angular Resolution Observations Toward the Hl Tau Region”, *The Astrophysical Journal Letters* 808, L3

Andrews, S. M. et al., 2018, “The Disk Substructures at High Angular Resolution Project (DSHARP). I. Motivation, Sample, Calibration, and Overview”, *The Astrophysical Journal Letters* 869, L41

Armitage, P. J., 2020, Astrophysics of planet formation, Cambridge University Press

Balanis, C. A., 2016, Antenna theory: analysis and design, John wiley & sons

Bean, B. et al., 2022, “CASA, the Common Astronomy Software Applications for radio astronomy”, *Publications of the Astronomical Society of the Pacific* 134, 114501

Booth, A. S. & Ilee, J. D., 2020, “13C17O suggests gravitational instability in the HL Tau disc”, *Monthly Notices of the Royal Astronomical Society: Letters* 493, L108

Bridle, A. H. & Schwab, F. R., 1999, “Bandwidth and Time-Average Smearing”, *Synthesis Imaging in Radio Astronomy II*, vol. 180, Astronomical Society of the Pacific Conference Series

Carrasco-González, C. et al., 2016, “THE VLA VIEW OF THE HL TAU DISK: DISK MASS, GRAIN EVOLUTION, AND EARLY PLANET FORMATION”, *The Astrophysical Journal Letters* 821, L16

Carrasco-González, C. et al., 2019, “The Radial Distribution of Dust Particles in the HL Tau Disk from ALMA and VLA Observations”, *The Astrophysical Journal* 883, 71

Cazaux, S. et al., 2003, “The hot core around the low-mass protostar IRAS 16293–2422: Scoundrels rule!”, *The Astrophysical Journal* 593, L51

Clark, B., 1980, “An efficient implementation of the algorithm’CLEAN’”, *Astronomy and Astrophysics, vol. 89, no. 3, Sept. 1980, p. 377, 378.* 89, 377

Comrie, A. et al., 2024, *CARTA: The Cube Analysis and Rendering Tool for Astronomy*, version 4.1.0

Condon, J. J. & Ransom, S. M., 2016, Essential radio astronomy, Princeton University Press

de Simone, M., 2022, “Hot corinos : the early organic molecular enrichment of the planet formation zones”, PhD thesis, Université Grenoble Alpes

Endres, C. P. et al., 2016, “The Cologne Database for Molecular Spectroscopy, CDMS, in the Virtual Atomic and Molecular Data Centre, VAMDC”, *Journal of Molecular Spectroscopy* 327, New Visions of Spectroscopic Databases, Volume II, 95

Facchini, S. et al., 2024, “Resolved ALMA observations of water in the inner astronomical units of the HL Tau disk”, *Nature Astronomy* 8, 587

Frediani, J., 2023, “Constraining the Chemical Conditions of Iras 4A2 Protostar at Planet-Forming Scales”, MA thesis, Alma Mater Studiorum - Università di Bologna

Garufi, A. et al., 2022, “ALMA chemical survey of disk-outflow sources in Taurus (ALMA-DOT): VI. Accretion shocks in the disk of DG Tau and HL Tau”, *Astronomy & Astrophysics* 658, A104

Goldsmith, P. F. & Langer, W. D., 1999, “Population Diagram Analysis of Molecular Line Emission”, *The Astrophysical Journal* 517, 209

Harrison Christopher, M., 2014, “Observational constraints on the influence of active galactic nuclei on the evolution of galaxies”, PhD thesis, Durham University

Högbom, J., 1974, “Aperture synthesis with a non-regular distribution of interferometer baselines”, *Astronomy and Astrophysics Supplement, Vol. 15, p. 417* 15, 417

Kenyon, S. et al., 1994, “A new optical extinction law and distance estimate for the Taurus-Auriga molecular cloud”, *The Astronomical Journal (ISSN 0004-6256), vol. 108, no. 5, p. 1872-1880* 108, 1872

Laas, J. C. & Caselli, P., 2019, “Modeling sulfur depletion in interstellar clouds”, *Astronomy & Astrophysics* 624, A108

Mayor, M. & Queloz, D., 1995, “A Jupiter-mass companion to a solar-type star”, *nature* 378, 355

Miotello, A. et al., 2023, “Setting the stage for planet formation: measurements and implications of the fundamental disk properties”, *Protostars and Planets VII* 534, 501

Miotello, A., 2018, “The puzzle of protoplanetary disk masses”, PhD thesis, Universiteit Leiden

Öberg, K. I. et al., 2023, “Protoplanetary disk chemistry”, *Annual Review of Astronomy and Astrophysics* 61, 287

Pinte, C. et al., 2022, “Kinematic Structures in Planet-Forming Disks”, *Protostars and Planets VII*, Astronomical Society of the Pacific

Rosenfeld, K. A. et al., 2013, “A spatially resolved vertical temperature gradient in the HD 163296 disk”, *The Astrophysical Journal* 774, 16

Schöier, F. L. et al., 2005, “An atomic and molecular database for analysis of submillimetre line observations”, *Astronomy & Astrophysics* 432

Shakura, N. I. & Sunyaev, R. A., 1973, “Black holes in binary systems. Observational appearance.” *Astronomy and Astrophysics, Vol. 24, p. 337-355* 24, 337

Soave, A., 2024, “Hunting for methanol in a water rich, planet forming disk”, MA thesis, Università degli studi di Milano

Taylor, G. B. et al., 1999, “Synthesis imaging in radio astronomy II”, *Synthesis Imaging in Radio Astronomy II* 180

Teague, R., 2019, “GoFish: Fishing for Line Observations in Protoplanetary Disks”, *The Journal of Open Source Software* 4, 1632

Teague, R. & Foreman-Mackey, D., 2018, “A Robust Method to Measure Centroids of Spectral Lines”, *Research Notes of the American Astronomical Society* 2, 173, 173

Teague, R. et al., 2018, “A Kinematical Detection of Two Embedded Jupiter-mass Planets in HD 163296”, *The Astrophysical Journal Letters* 860, L12

Walsh, C. et al., 2016, “First Detection of Gas-phase Methanol in a Protoplanetary Disk”, *The Astrophysical Journal Letters* 823

Williams, J. P., 2021, Introduction to the interstellar medium, Cambridge University Press

Wilson, T. L. et al., 2009, Tools of radio astronomy, vol. 5, Springer

Yen, H.-W. et al., 2016, “Stacking spectra in protoplanetary disks: detecting intensity profiles from hidden molecular lines in HD 163296”, *The Astrophysical Journal* 832, 204

Yen, H.-W. et al., 2017, “1000 au exterior arcs connected to the protoplanetary disk around HL Tauri”, *Astronomy & Astrophysics* 608, A134

Yen, H.-W. et al., 2019a, “HL Tau disk in HCO+ (3–2) and (1–0) with ALMA: gas density, temperature, gap, and one-arm spiral”, *The Astrophysical Journal* 880, 69

Yen, H.-W. et al., 2019b, “Signs of outflow feedback from a nearby young stellar object on the protostellar envelope around HL Tauri”, *Astronomy & Astrophysics* 623, A96

Zhang, K. et al., 2020, “Rapid Evolution of Volatile CO from the Protostellar Disk Stage to the Protoplanetary Disk Stage”, *The Astrophysical Journal Letters* 891, L17

Ringraziamenti

Giunto alla fine di questo faticoso progetto è il momento di alcuni doverosi (ma non per questo meno sinceri) ringraziamenti verso coloro che sono stati coinvolti nella sua realizzazione.

Ringrazio il mio relatore prof. Leonardo Testi, una persona di inesauribile gentilezza, e la cui contagiosa passione per lo studio dell'Universo ho percepito e ammirato dalla prima lezione. Ringrazio poi la dott.ssa Anna Miotello, per essersi resa disponibile come mio supervisore allo European Southern Observatory, per avermi calorosamente accolto, e per aver condiviso con me la sua conoscenza, esperienza, ed il suo entusiasmo per la ricerca. Entrambi mi hanno non solo guidato nel lavoro di tesi, ma anche supportato nella preparazione di svariate domande di dottorato. Con entrambi spero di lavorare ancora in futuro.

Ci tengo anche a ringraziare Alessandro Soave per il suo essenziale sostegno, specialmente nelle fasi iniziali del progetto.

Ringrazio sentitamente lo European Southern Observatory e la sua amministrazione, per avermi concesso la possibilità di lavorare nel cuore della ricerca astronomica in Europa, e tutti gli studenti e ricercatori che ho avuto il piacere di conoscere durante la mia, seppur breve, permanenza. Mi hanno fatto sentire parte della loro meravigliosa comunità e tutti loro ho salutato augurandomi di rivederli prima o poi in qualche seminario o conferenza.

Ci sono poi ringraziamenti molto più difficili. Come si dimostra, in qualche paragrafo, gratitudine verso le persone più importanti della propria vita? Persone il cui diretto contributo a questo libriccino accademico sarà stato anche inesistente, ma la cui presenza ha caratterizzato così profondamente quel percorso che si conclude con queste pagine. Non lo so. Dovranno accontentarsi di questo misero tentativo.

Anna e Leonardo condividono il nome con i miei relatori, e con loro condividono anche una certa importanza in questi ringraziamenti. Quando Leonardo si è laureato e abbiamo smesso di condividere una stanza, ero rassegnato al fatto che non l'avrei mai più rivisto. Quando ho conosciuto Anna, ho pensato che sarei stato fortunato anche solo a instaurare con lei un superficiale rapporto di amicizia. Oggi sono le persone più presenti nella mia vita (e nella mia testa). L'amore che provo per entrambi non sono abbastanza bravo con le parole per poterlo concretizzare su carta. A loro si aggiunge un largo gruppo di bellissime persone a cui voglio un mondo di bene, e che mi perdoneranno se decido di non elencare qui nella loro totalità, ma tra le quali mi sembrerebbe sbagliato non menzionare, in nessun particolare ordine di importanza: Enrico e Sara, i quali si sono resi conto troppo tardi della loro pessima decisione di farmi partecipare al ranking delle anime opening; Federica, con cui finalmente, dopo sei mesi passati *nel paese più stabile e normale* *del mondo*, posso tornare a giocare a overcooked; Davide, violatore delle convenzioni di Ginevra (mi fa il solletico); Dav- ehm, volevo dire, Giacomo; Dima; Alice.

Alex ed io siamo amici, nemesi, *siblings*, colleghi in un fittizio studio artistico, e decine di altre cose. Con nessun'altra persona ho condiviso storie e costruito mondi per così tanto tempo. Ispira da sempre la mia creatività, e chiunque vorrà un ricordo della mia laurea si porterà a casa un po' della sua arte.

Edoardo e Giampaolo sono ormai praticamente dei familiari. Nonostante gli anni d'oro della nostra amicizia siano passati, c'è qualcosa di confortante nel vederli ancora popolare la mia mansarda dopo quasi 11 anni.

Di tutte le persone che ho conosciuto a Bologna, Lorenzo e Giulia meritano particolare riconoscimento, l'uno per aver patito con me le peggiori pene che la magistrale in Astrophysics and Cosmology ha da offrire (Multiwavelength Lab), e l'altra per averci sopportato entrambi mentre perdevamo lentamente la nostra sanità mentale. Con Leonardo (un altro ancora) ho condiviso buona parte del mio ultimo anno, insieme ci siamo probabilmente fregiati del titolo di frequentatori più assidui dell'ufficio del prof. Testi. Molti altri compagni astrofisici hanno riempito le mie giornate, molti continuerò a frequentare durante i prossimi tre anni, molti nuovi avrò occasione di conoscere.

E se queste erano le persone che fanno parte della famiglia che mi sono scelto, non mi rimane che menzionare infine la mia famiglia vera e propria. Ai miei genitori, a Michela, grazie per avermi cresciuto, sostenuto, amato.