universität wien

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Early Science with JWST/NIRSpec: Studying the protostellar Outflow of HH211

verfasst von | submitted by

Anita Isabel Schwaiger BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 861
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Astronomie
Betreut von | Supervisor:	Assoz. Prof. Dr. Alvaro Hacar Gonzalez MA

Abstract

The James Webb Space Telescope (JWST) and its Near InfraRed Spectrograph (NIRSpec) offer an unprecedented opportunity for detailed spectroscopic studies of young, deeply embedded protostars and their surroundings. Given the tight link between accretion and ejection of material, the study of protostellar jets and outflows and their interaction with the interstellar medium is particularly interesting.

This thesis studies the terminal bow shocks of Herbig-Haro 211 (HH211), a protostellar outflow driven by the very young protostar HH211 mm. I investigate the extinction, the excitation conditions and the ortho-to-para ratios (OPRs) of the system. Estimating and correcting the extinction accurately is essential for obtaining reliable results. Studying the excitation conditions provides insight into the excitation mechanism and, in this case, the shock properties. The OPR contains important information about the thermal history of the system.

For this project, I developed and implemented a set of custom python routines to carry out spectral analysis on a pixel-by-pixel basis. The extinction is estimated by comparison of theoretical to observed line ratios. Excitation diagrams reveal the excitation conditions and the ortho-to-para ratios.

Analyzing the excitation conditions yields two separate components — a warm one and a hot one — with column densities of $\sim 2 \cdot 10^{18}$ to $\sim 2 \cdot 10^{19}$ cm^{-2} and $\sim 6 \cdot 10^{17}$ to $\sim 3 \cdot 10^{18}$ cm^{-2} respectively. Their excitation temperatures are between 1500 and 2500 K and between 2000 and 3500 K. The uncertainties of the excitation analysis are mostly below 23% for the column densities and below 200 K for the excitation temperatures. The calculated ortho-to-para ratios are mostly close to 2.5 for the warm component and mostly around 2.75 for the hot component. Their corresponding uncertainties are relatively high with 0.25 and 0.5. The estimated visual extinction ranges from 2.5 to 15 mag with errors mostly below 0.5 mag.

Apart from the OPRs, these results are consistent with previous research. Slight differences are likely caused by the different observing strategies that probe different components of the outflow. The tentative trend of higher OPRs for the hot component is contrary to the expected relation and likely an effect of the method used. Given the excellent data quality and my detailed analysis, both the spatial resolution and the uncertainties of inferred properties could be drastically improved compared to prior near infrared studies of the terminal bow shocks of HH211.

Kurzfassung

Das James Webb Space Telescope (JWST) und sein nah-infrarot Spektrograf (NIRSpec) bieten nie dagewesene Möglichkeiten für spektroskopische Beobachtungen von jungen, noch tief in Molekülwolken eingebetteten, Protosternen und ihrer Umgebung. Der enge Zusammenhang zwischen Akkretion auf ein junges stellares Objekt und Materieauswurf in Form von Jets und Winden ist von besonderem Interesse.
Diese Masterarbeit untersucht die äußersten Bugschocks von Herbig-Haro 211 (HH211) – einem System von protostellaren Jets und Schocks, das der sehr junge Protostern HH211 mm antreibt. Untersucht werden die Extinktion, die Anregungszustände und das ortho-zu-para Verhältnis (OPR). Die Korrektur der Extinktion ist essenziell für zuverlässige Ergebnisse. Aus den Anregungszuständen lassen sich die stattfindenden Anregungsprozesse und die Eigenschaften der beteiligten Schocks ableiten. Das OPR enthält Informationen über die thermische Vergangenheit des Systems.
Für die Datenauswertung wurden im Rahmen dieser Arbeit maßgeschneiderte Python Routinen zur pixelweisen Analyse entwickelt. Die Extinktion wird durch den Vergleich von theoretischen mit beobachteten relativen Linienhelligkeiten abgeschätzt. Die Anregungszustände und das OPR werden mittels Anregungsdiagrammen bestimmt.
Daraus ergeben sich zwei eigenständige Komponenten — eine warme und eine heiße. Ihre entsprechenden Säulendichten betragen zwischen $\sim 2 \cdot 10^{18}$ und $\sim 2 \cdot 10^{19}$ cm^{-2} und zwischen $\sim 6 \cdot 10^{17}$ und $\sim 3 \cdot 10^{18}$ cm^{-2}. Die Anregungstemperaturen liegen zwischen 1500 und 2500 K für den warmen Teil und zwischen 2000 und 3500 K für den heißen Teil. Die Unsicherheiten sind überwiegend unter 23 % für die Säulendichten und überwiegend unter 200 K für die Anregungstemperaturen. Die meisten OPRs betragen rund 2,5 für die warme Komponente und rund 2,75 für die heiße Komponente. Ihre zugehörigen Unsicherheiten sind mit 0,25 und 0,5 relativ hoch. Für die visuelle Extinktion werden Werte zwischen 2,5 und 15 mag errechnet; die Unsicherheiten sind niedriger als 0,5 mag. Abgesehen von den OPRs sind die Ergebnisse in guter Übereinstimmung mit jenen vorheriger Untersuchungen. Kleinere Abweichungen sind durch unterschiedliche Beobachtungsstrategien bedingt. Der schwache Trend zu höheren OPRs für die heiße Komponente ist entgegen der Erwartung und liegt wohl an der angewandten Methodik. Durch die ausgezeichnete Datenqualität und die detaillierte Analyse konnten sowohl die räumliche Auflösung als auch die Unsicherheiten der errechneten Eigenschaften gegenüber früheren nah-infrarot Beobachtungen der Bugschocks von HH211 maßgeblich verbessert werden.

Acknowledgments

Words cannot express my gratitude to my supervisor Assoc.Prof. Dr. Alvaro Hacar for his guidance and invaluable feedback.
Thanks should also go to the entire Interstellar Medium Astrophysics (ISMA) Group of the Institute for Astrophysics at the University of Vienna, for their insightful suggestions and questions during our group meetings.
I am also thankful for the support of my friends and family, who helped with proof-reading and offered constant encouragement.

This work is based on observations made with the NASA/ESA/CSA James Webb Space Telescope. The data was obtained from the ESA jwst science archive — available under `https://jwst.esac.esa.int/archive/`. These observations are associated with program 1257. This research has made use of the VizieR catalogue access tool, CDS, Strasbourg, France (DOI : 10.26093/cds/vizier). The original description of the VizieR service was published in 2000, A&AS 143, 23. This research has made use of the Astrophysics Data System, funded by NASA under Cooperative Agreement 80NSSC21M00561.

Contents

List of Figures

1 Introduction

The formation of new stars is a complex multi-step and multi-faceted process. During the early stages of a star's life it is rapidly accreting matter. Not all of this matter reaches the forming star – significant amounts of material are ejected away giving raise to jets and winds. Shocks occur where these ejecta impact the surrounding medium. Parts of the surrounding medium are swept up to form massive outflows. Faster parts of the jet overtaking slower ones create additional shocks along the jet. So, studying the shocks along an outflow presents a unique opportunity to investigate the ejection history of a protostar. Given the tight link between ejection and accretion, this ejection history is a proxy for the accretion history of the protostar (Bally, 2016). The sections 1.1 and 1.2 give an introduction to star formation and the evolution of forming stars; section 1.3 explains the associated jet, wind and outflow structures.

In this thesis I study the terminal bow shocks of the Herbig-Haro 211 (HH211) outflow using data from the Near InfraRed Spectrograph (NIRSpec) of the James Webb Space Telescope (JWST). HH211 is often regarded as a prototypical example of an outflow from a very young protostar. Its proximity and low inclination with respect to the plane of the sky make it a prime target for the study of jets and outflows (Caratti o Garatti et al., 2024). As the studied emission stems from molecular hydrogen, section 1.4 details its properties and its expected excitation. Section 1.5 summarizes the prior research on HH211. My analysis includes extinction, excitation conditions and ortho-to-para ratios. Determining the extinction is a necessity as it needs to be corrected for the study of the excitation conditions. Additionally, studying the effects of extinction allows conclusions about the composition of the dust present in the sightline. From the excitation analysis I infer column densities and excitation temperatures that help investigate the properties of the exciting shocks. Finally, the ortho-to-para ratios encode hints of the thermal history of the material observed.

1.1 General introduction to star formation

Stars form in cool and dense molecular clouds. Typically, stars do not form in isolation, but as a group or cluster within the same cloud or an extended cloud complex. If parts of that cloud reach the critical density, they can collapse under their own gravity. In this early phase the collapse happens essentially in free-fall — the gas is still sparse enough so

that the opacity is low and the produced thermal energy can easily be radiated away. With the continued rise of the central density, the central region becomes increasingly opaque to infrared radiation. Because of this, the cooling becomes less efficient which leads to growing temperature and pressure in the central region. The material can no longer collapse freely and a pressure supported core — the first hydrostatic core — forms. Material continues to accumulate onto that core increasing its density and temperature further. A second collapse occurs once the dissociation temperature of molecular hydrogen is reached at around $\sim$ 2000 K. The additional energy needed for the dissociation slows the rise of temperature and pressure allowing for a second collapse. After that collapse a second hydrostatic core forms. It continues to accrete matter, contract and heat up until nuclear fusion ignites in the young star's core and it reaches the main sequence (Young, 2023).

As a consequence of the conservation of angular momentum, a circumstellar disk forms almost immediately after the initial collapse. This disk is integral in driving accretion onto the young star. Angular momentum plays a central role throughout the entire accretion process and it has to be carried away to allow further transport of matter towards the forming star. Within the disk internal friction or viscosity enable the transport of angular momentum to the outer parts of the disk. This allows material in inner parts to be accreted onto the central core. Angular momentum can also be carried away in the form of outflows (Williams and Cieza, 2011). With the accretion of material the protostar gains a lot of angular momentum — the expected consequence would be an increase in the rotation of the protostar. Observations show that this is not the case — most young sources are slow rotators. Jets and outflows — specifically those emitted directly from the surface of the forming star — are a feasible mechanism of controlling the spin of young stellar objects (YSOs) (Ray and Ferreira, 2021).

1.2 Evolution of young stellar objects (YSOs)

Dense cloud fragments without a hydrostatic core at their center are typically called pre-stellar cores. Objects that have a core but have not reached the main sequence yet are referred to as young stellar objects (YSOs) (Andre et al., 2000). YSOs are typically categorized into 4 classes. The Classes I, II and III are defined based on the object's observed spectral energy distribution (SED) in the 1 to 20 μm range in Lada (1987). Andre et al. (1993) extend this classification scheme with the Class 0, which was defined based on submillimeter observations. The different classes mark an evolutionary sequence.

Class 0 objects are unobservable at wavelengths < 10 μm due to the large amount of surrounding material and the resulting high extinction (Andre et al., 1993). For this reason, they do not appear as detected sources in Lada (1987) and were, thus, not described there. Andre et al. (1993) coin the term Class 0 for submillimeter sources that show evidence for a central heating source — such as the detection of an outflow or an associated radio source — and have a large fraction of cirumstellar mass compared to the mass of the central source. This second criterion can be sampled by the ratio of bolometric luminosity to the luminosity at 1.3 mm. The former grows with the mass of the central source while the later relates mostly to the material around the source.
Consequently, Class 0 YSO are deeply embedded in dense envelopes still containing a majority of the total mass. These envelopes dominate the SED, making it appear as black-body radiation corresponding to a temperature of few 10 K. This class marks the first evolutionary stage of a YSO and for many of them, outflows and disks are observed (Ray and Ferreira, 2021).

Class I objects are further developed, resulting in less dense envelopes and lower extinction. As a consequence, these sources can be observed in the near-infrared. The observed emission is a combination of thermal emission of the envelope around the central source and from the accretion disk (Ray and Ferreira, 2021). The emission observed longwards of 10 μm stems mainly from the cool envelope while the emission at shorter wavelengths is from the warmer disk. Overall the contribution from the envelope still dominates the SED (Robitaille et al., 2006). This results in the defining feature of Class I objects: a rising energy output longwards of 2 μm (Lada, 1987).

As soon as the photosphere of the forming star becomes optically observable, the YSO has reached the T Tauri stage. If the accretion disk is still present, these objects are referred to as Class II objects or classical T Tauri stars. In that stage a combination of the black-body emission from the star and the thermal emission of the accretion disk is observed. The SED is still broader than a single black body — the disk's emission causes an excess in the infrared. Due to the different temperatures present in the disk, the SED typically does not resemble the combined emission of two black bodies, but takes on a more complex form (Ray and Ferreira, 2021). In contrast to Class I objects, the energy output of Class II objects is flat or falling longwards of 2 μm (Lada, 1987).

With the dissipation of the disk, the SED starts to more closely resemble that of a single black body. This is characteristic for Class III YSOs — also called weak-line T Tauri stars. This stage ends with the ignition of nuclear hydrogen burning in the star's core and the star joining the main sequence (Ray and Ferreira, 2021).

1.3 Protostellar jets, winds and outflows

From young stellar objects to accreting neutron stars, active galactic nuclei and quasars, bipolar jets and outflows are observed near many rotating and accreting astrophysical systems. In early studies of star forming regions a number of different, seemingly unrelated objects were discovered and described. These include the pre main-sequence T Tauri stars and Herbig AeBe stars (discovered as objects with emission line dominated spectra and irregular variability), Herbig-Haro (HH) objects, and molecular outflows observed in CO and H_2 (Bally, 2016). Herbig-Haro objects are small nebulous structures, that are observed in optical line emission and generally found in or near dark clouds. In contrast to HII regions (photoionized regions around hot stars), HH objects do not generally coincide with a stellar source and show more and stronger low-excitation lines. This suggests an excitation mechanism different from photoionization for HH objects (Schwartz, 1983). The emission of HH objects is caused by supersonic shock waves, which are powered by the outflows of young stars (Bally, 2016). Outflows are only found near sources that have an accretion disk (Ray and Ferreira, 2021). Additionally, observations show a strong correlation between accretion luminosity and the mass loss rate through outflows. This suggests that outflows are powered by accretion and that the two processes are tightly connected (Hartmann et al., 2016). All of the mentioned phenomena are now known to be directly linked to accretion and ejection processes of young stars (Bally, 2016).

This is a good point to specify the terminology used in this thesis. Jet, outflow and wind all refer to material flowing away from the forming star. The key difference is in their velocities, their degrees of collimation and the source of the material. *Jets* are generally defined as fast — typical velocities higher than 50 km/s — material, that is highly collimated and launched from the inner parts of the star-disk system. In contrast to this, the term *outflow* most commonly describes surrounding material that has been swept up by the jet (van Dishoeck et al., 2025). Typical velocities for outflows are around 1–30 km/s and they are less collimated than jets (Frank et al., 2014). Sometimes the term outflow is used to refer to the entire system of jet, swept-up material and wind. *Winds* are characterized by their low velocities $\sim$ 10 km/s and wide opening angles. Like jets, winds are also ejected from the inner parts of the system — either from the stellar surface or the surface of the disk (van Dishoeck et al., 2025).

1.3.1 Launch mechanisms

For YSOs, the majority of observations show that outflows inject magnitudes more momentum into the surrounding medium than can be explained by the radiation field of the central source. Thus, the launch of outflows needs to include additional mechanisms (Bally, 2016).

There is a general consensus that jets are launched by magnetocentrifugal acceleration along open magnetic field lines. These open lines can occur in the disk's magnetic field, in the area of interaction between the magnetic field of the disk and that of the star or in the stellar magnetic field. So the material is launched from the surface of the accretion disk (called disk winds), from the star-disk interface or from the surface of the star (called stellar winds). The different launch mechanisms have their specific strengths and weaknesses and no single model explains the observed phenomena fully (Bally, 2016).

Stellar winds would provide a very convenient way of regulating the spin of the central source as the two interact directly. The stellar wind models introduced in Ray and Ferreira (2021) have their specific problems making both of them unlikely to be the main launching mechanism. One model requires high rotation rates for the young stars — which does not match observation — and cannot reproduce the observed mass loss rates otherwise. The other model has a problem with massive outflows — the models require increasingly high rates of accretion to power massive outflows. The higher accretion rate brings more angular momentum to the star, which would then have to be counteracted by even more massive outflows. In some cases this model simply does not converge.
One of the most prominent models for outflows from the star-disk interface is the X-wind model. This model assumes a strong magnetic field for the forming star, that significantly disrupts the inner parts of the disk, but opens up and diffuses further out. The closed field lines funnel material onto the star. Simultaneously, material is centrifugally accelerated away along open field lines. The wind is expected to carry away all of the angular momentum of the accreted material, so that the star does not spin up. While elegant, the model is not consistent with observed jet kinematics. The assumed properties for the underlying disk are rather extreme and the transport of angular momentum from the accreted material to the outflow material needs some additional (viscous) process. Some numerical simulations show ejections from the inner edge of the disk. These magnetospheric ejections are episodic and occur along with the recombination of magnetic field lines. They are able to spin down the star and the disk. However, the material from such ejections moves nearly ballistically. Thus, without any other outflows present, magnetispheric ejections alone cannot match the observed YSO jet structures (Ray and Ferreira, 2021).
Disk winds are thought to be one of the main driving forces for accretion, as other mechanisms for the transport of angular momentum — such as the magneto-rotational instability — have been shown to only apply situationally. The corresponding models rely on a large scale, vertical, magnetic field — threading either the entire disk or only the inner parts — to magnetically launch material from the disk's surface and carry away angular momentum. Disk winds are the favored mechanism for providing the bulk of the material in protostellar jets as well as removing the bulk of the angular momentum that needs to be removed to allow accretion. By their very nature, though, disk winds are unable to spin down the star. Additionally, they do not explain observed jet asymmetries or hot inner flows (Ray and Ferreira, 2021).

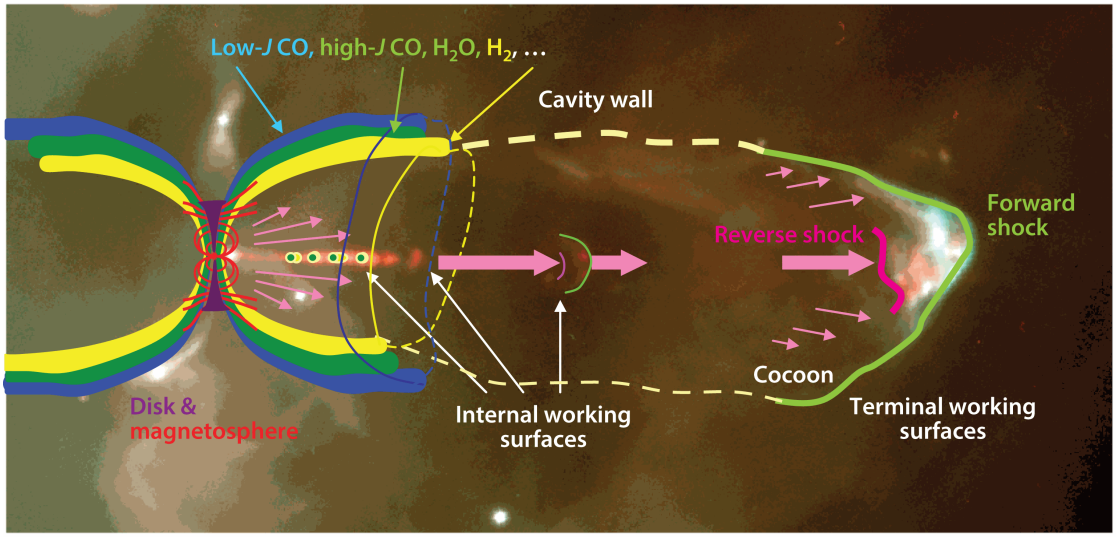
Therefore, a combination of mechanisms seems likely. A probable structure is an on-axis stellar wind surrounded by a multi-component disk wind. Time-variable mass ejections from the star-disk interface could produce internal jet shocks and help explain the knots and inner bow shocks seen in objects like HH211 (Ray and Ferreira, 2021).

1.3.2 Outflow morphology

While the launching zones of jets are generally quite close to the YSO and rather small, which makes them difficult to observe, the jets impact their surrounding on a much larger scale. The prime example for this are outflows. The typical morphology of an outflow depends on the evolutionary stage of the driving source. Outflows from Class 0 YSO are typically smaller than 1 parsec, driven by jets with velocities in the range of 50 to 150 km/s and mostly molecular. They tend to be bright in CO, H_2O and OH and their shocks are seen in H_2 transitions. Optical HH objects are usually heavily obscured and dim if present at all (Bally, 2016). The mass fluxes observed through outflows of Class 0 objects are generally higher than those of more evolved sources (Ray and Ferreira, 2021).

Figure 1.1 — taken from Bally (2016) — visualizes the structure of a typical young outflow. The entire structure is driven by a molecular jet that plows through the parent cloud. A terminal bow shock forms where the jet impacts the surrounding medium. There, the surrounding material is swept up and largely incorporated into the flow — this process is called prompt entrainment. Additional shocks occur on inner working surfaces. These are created as faster material rams into the slower material ahead of it. Causes for these velocity differences within a jet are the variability of ejection velocity as well as the jet being slowed by ambient material in its path. On account of the larger relative velocities, the terminal working surface generally includes the brightest shock. Around the jet a low-density cavity forms. Responsible for this are wide-angle winds or the sidewards expansion of shocked jet material. Kelvin-Helmholtz instabilities and turbulent motion — induced by shear forces — allow the transport of momentum from the outflow to the surrounding medium. Along the cavity walls, shocks and UV radiation excite various molecules. This results in a layered structure of shells of molecular emission with H_2 emission found closest to the jet, followed by emission of H_2O and higher-J CO, and finally lower-J CO emission. HH211 is a very young outflow and follows this structure closely (Bally, 2016).

In contrast to this, outflows from more evolved YSOs are powered by faster, less massive and increasingly atomic and ionized jets (Ray et al., 2023). The outflow cavities tend to widen as time passes and the central source evolves. Emission from molecules along the cavity edges shows a wide range of temperatures and becomes less prominent as more



Bally J. 2016.
Annu. Rev. Astron. Astrophys. 54:491–528

Figure 1.1: A visualization of a young outflow. The main components are the primary jets and winds shown in pink, the working surfaces of the shocks shown in green and the cavity and its walls. Emission traces the terminal bow shock — caused by the interaction of a supersonic jet with the surrounding slow or resting material, the inner shocks — caused by velocity differences between different parts of the jets — and cavity shocks along the walls of the cavity. The colored bands indicate the different molecules that are excited and observable along the cavity walls. In the dashed areas the cavity walls are mostly atomic or ionized. The disk and the magnetosphere of the forming star are also shown; their size is, however, largely exaggerated. The image in the background shows the Herbig-Haro 34 outflow with its jet traced by [S II] emission (red) and the shocks in Hα (cyan). This figure was taken from Bally (2016).

and more material is swept up and the shells decelerate. The full structure often extends multiple parsecs from the source and includes symmetric chains of HH objects, that can extend past their parent clouds (Bally, 2016).

1.3.3 Observational caveats

The detailed properties of a protostellar outflow depend on a number of factors including the luminosity, mass and evolutionary stage of the driving YSO, its accretion history, and the characteristics of the surrounding medium. In addition, different tracers are sensitive to different components of the outflow and, thus, yield different inferred outflow properties (Bally, 2016).

Low-J rotational CO transitions — observable at radio wavelengths — are commonly used because CO is abundant, easily excited even in cold clouds and the lowest transitions are accessible to observation from the ground. They trace cool regions with lower excitation, like the outer parts of the cavity walls. Using radio continuum observations, jets can be studied closely to their sources and — for some outflows — synchrotron radiation caused by relativistic particles can be observed.
The warmer parts of the molecular outflow are best studied in the infrared or (sub)mm-regime. Common tracers are higher-J and ro-vibrational transitions of CO, SiO and H_2O as well as shock excited H_2. For these molecules a large number of lines can be observed within a relatively small wavelength range. This allows for detailed studies of their physical properties such as the excitation conditions.
Line emission ranging from the sub-millimeter to the visual regime reveals the atomic and ionized components. Such tracers include sub-millimeter fine structure lines of [O I] as well as some [C II] and [N II] transitions. In the visual to near-IR, emission lines from the Balmer, Paschen and Brackett series of hydrogen and the forbidden transitions of [S II], [O I], [O II], [O III], [N II] and [Fe II] also trace the atomic and ionized components.
For some sources even X-rays have been detected. They point to the presence of very fast shocks and hot plasmas with temperatures around some mega-Kelvin (Bally, 2016).

1.4 Molecular hydrogen

Being the most abundant molecule in the universe the study of molecular hydrogen H_2 is of particular importance and the focus of this work. As just discussed it is also an excellent tracer for the warmer and higher excited parts of a molecular outflow. Some of the most important properties of H_2 are introduced here.

1.4.1 Spin isomers: ortho and para H_2

H_2 has two spin isomers that are only differentiated by the alignment of the nuclear spins of the two hydrogen atoms. The isomer with parallel spin alignment is a triplet state and referred to as ortho hydrogen or o-H_2. If the spins are anti-parallel, the isomer is called para hydrogen or p-H_2; this state is a singlet state. Due to the ortho configuration being three times degenerate and the para configuration being non-degenerate, it is expected that H_2 forms with an ortho-to-para ratio (OPR) of 3. This OPR of 3 is the equilibrium value for warm environments. The para configuration has, however, a lower binding energy making it the preferred state in cold conditions. This results in an OPR of ~ 0 for very cold systems. As there are no radiative transitions between ortho and para hydrogen, equilibria are reached slowly resulting in many — mostly cool — systems being out of equilibrium. The thermodynamic properties — and especially the heat capacity — of H_2 gas depend on the ortho-to-para ratio and impact the temperature evolution of molecular clouds and the stability of massive circumstellar disks (Vaytet et al., 2014).

This thesis studies the rotational and ro-vibrational line transitions of molecular hydrogen in the near infrared. The vibrational state of the molecule is typically denoted as ν and the rotational state as J. The indices u and l indicate the upper and lower states. Transitions are noted as $H_2(\nu_u - \nu_l)X(J_l)$ — instead of X a letter indicating ΔJ is used. As H_2 is a symmetric molecule, it has no permanent dipole-moment making radiative dipole transitions impossible and quadrupole transitions the predominant ones. They have ΔJ of -2, 0 or +2 — the corresponding letters used in notation are S, Q and O. Only ortho hydrogen can exist in odd rotational states and only para hydrogen in even ones. This means that only one spin isomer contributes to each transition making the differentiation between them easy (Draine, 2011). In the observed wavelength range of ~ 1.6 to 5.2 μm upper state excitation energies starting from ~ 6000 K can be observed. The upper limit for energies of observable states depends on the excitation conditions and instrument sensitivity; some lines originating from highly excited states fall into this wavelength range (Roueff et al., 2019a).

Figure 1.2 shows the spectrum of the terminal bow shock of HH211 as observed by NIRSpec. This figure was done by Ray et al. (2023) and uses the same observational data as will be analyzed in this project. The most prominent lines stem from purely rotational (0-0), and ro-vibrational (1-0) and (2-1) transitions of H_2 as well as the principal (1-0) CO transitions in form of the R- and P-branches.

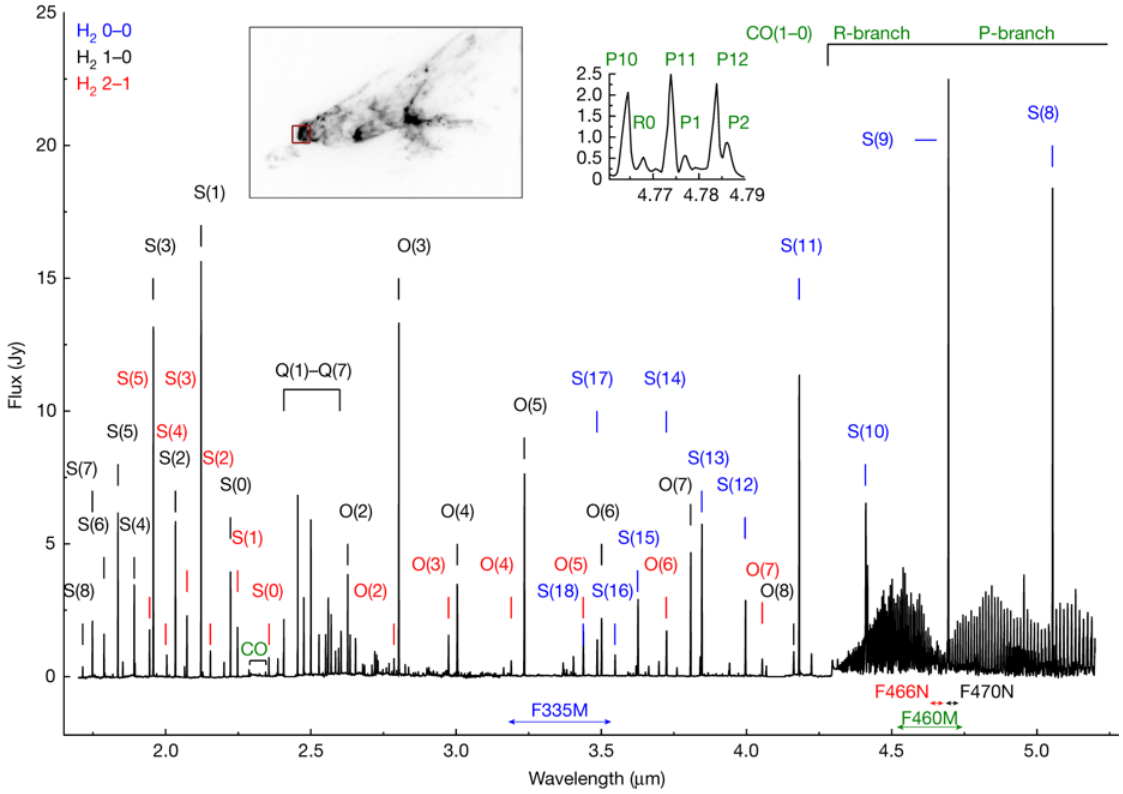


Figure 1.2: A visualization of the most prominent lines in the terminal bow shock of HH211 as seen by NIRSpec. The purely rotational lines are labeled in blue and the (1-0) and (2-1) ro-vibrational lines in black and red respectively. Towards the higher wavelengths the CO (1-0) R- and P-branches are visible. For those transitions secondary peaks caused by the different isotopologues of CO are also present - they are shown in the insert. Additionally, the wavelength coverage of F335M, F460M and F470N filters of NIRCam is indicated — those filters are the ones used for Figure 1.3. This figure was taken from Ray et al. (2023).

1.4.2 Shock excitation of molecular hydrogen

In the observed region — the region of the terminal bow shocks of the protostellar outflow HH211 — the main excitation mechanism for molecular hydrogen is shock excitation. Passing through the shock the material is heated, accelerated and compressed. The medium then cools via line emission, which is studied in this thesis. Depending on the shock properties — such as its velocity, the pre-shock density or the strength of an external magnetic or UV radiation field — different states of H_2 are excited. As these shock parameters cannot be observed directly, they are usually determined by comparison to shock models. Kristensen et al. (2023) provide an extensive grid of such models as well as insight into the effects of each parameter on the resulting emission. The models reported there are of one-dimensional stationary shocks. One key take-away from the shock simulations is that excitation temperatures and column densities inferred from observation of line intensities vary greatly depending on the lines chosen. This is because shocks are inherently out of equilibrium phenomena and the excitation of H_2 states does not follow a single Boltzmann distribution. Generally, the lines associated with higher energy states trace more energetic components and, thus, higher excitation temperatures and lower column densities are recovered from them.

A similar effect occurs for the OPRs inferred from observation. The simulations assume for the gas to have reached the equilibrium values for the OPR before the emergence of a shock. At the initial temperatures of ~ 10 K OPRs are close to 0. As the shock passes through and heats the material, the exchange reactions necessary to convert para to ortho hydrogen become more efficient and increase the OPR towards the high temperature equilibrium value of 3.
Here it is important to point out that the OPR can be calculated for either the column density, which is generally done for the interpretation of observations where the local values are unknown, or using the local (volume) densities. These two values are not necessarily consistent with each other. In the simulations, the OPRs calculated from the column densities are lower than the local values and lag behind in their evolution over time (Kristensen et al., 2023). Additionally, the true column density is unknown and the choice of lines used to infer it has implications on the resulting OPRs — even if lines of a similar energy regime are used to determine the column densities of each isomer. As the material cools after the shock, the higher energy states are no longer being excited. If the exchange reactions continue to convert para to ortho hydrogen, the high energy OPR no longer reflects the current OPR of the bulk of gas — it becomes "frozen". This happens earlier for higher energy states. As a result of this the OPRs recovered from lines of higher energy states are generally lower than those inferred from lower energy states (Wilgenbus et al., 2000).

While I do not carry out quantitative comparison of the observed line emission to the shock simulations from Kristensen et al. (2023), the qualitative comparison still provides valuable insights.

1.5 Herbig-Haro 211 (HH211)

Due to its proximity and favorable inclination, HH211 is a prime target for the detailed study of jets and outflows. The driving source is a very young Class 0 YSO and the outflow often is regarded as prototypical for young sources (Ray et al., 2023). In this thesis I study the new observational data from JWST providing an unprecedented look into the HH211 system. Figure 1.3 shows a near infrared image of HH211 from Ray et al. (2023). That image covers the entire outflow structure, while this work focuses on the three bow shock furthest from the source in the south-east lobe — this set is sometimes referred to as the terminal bow shocks of HH211.

Described first in McCaughrean et al. (1994) Herbig-Haro 211 has become one of the most extensively studied protostellar outflows over the last 30 years. The discovery of HH211 was incidental: as a result of a near infrared survey searching for a proposed deeply embedded young stellar cluster near IC 348 IR in the Perseus dark cloud complex. Following this first detection the authors carried out a number of targeted follow up observations in the infrared, sub-millimeter and millimeter wavelengths.
The near infrared observations include imaging in the J, H and K bands as well as a narrow-band filter centered around 2.122 μm tracing the (1-0) S(1) emission line of molecular hydrogen. The results imply that the observed emission is dominated by line emission with little to no continuum present. Given the observed relative brightness between filters, shocks are a likely excitation mechanism. The morphology of the H_2 emission indicates that it is not tracing the jet itself, but the inner and terminal working surfaces as well as the cavity walls. Based on the observed size of the outflow and an assumed outflow velocity of $\sim$ 100 km/s McCaughrean et al. (1994) estimated a dynamical age of around 1000 years for the outflow.
The observations at longer wavelengths showed a CO outflow coinciding with the H_2 emission. The CO outflow's south-east lobe is blue-shifted and the north-west lobe is red-shifted. The driving source remained hidden in the near infrared regime and is only visible in continuum observations at millimeter wavelengths. It is located right between the two outflow lobes and accordingly named HH211 mm. McCaughrean et al. (1994) characterized it as a "typical young, deeply embedded outflow source".

Figure 1.3: A near infrared image of HH211 from Ray et al. (2023) taken by the Near InfraRed Camera (NIRCam) of JWST. The colors correspond to its F335M (blue), F460M (green) and F470N (red) filters. F335M is a medium-band filter tracing multiple H_2 lines including the (1-0) O(5) line. The other medium-band filter, F460M, also includes multiple H_2 lines — notably the (0-0) S(9) line — and a large number of lines of the CO (1-0) R-branch. Emission observed in the narrow band-filter F470N stems near exclusively from the (0-0) S(9) line of H_2. The wavelength coverage of those filters can also be seen in Figure 1.2.
The data analyzed in this work encompasses the set of three bright bow shocks that can be seen at the tip of the south-eastern lobe.

Later studies included numerous follow up observations using different instruments and covering a range of wavelengths. Below follows a brief and by no means exhaustive overview.

Of particular interest for this project is the near infrared study O'Connell et al. (2005). It included spectroscopic observations covering a wavelength range of 1–2.5 μm and aligned the spectrograph slit through the centers of the terminal bow shocks. They found a singular excitation temperature along the slit to be unlikely and that a part had excitation temperatures higher than 2000 K. From their excitation analysis they found the column densities to be consistent with an ortho-to-para ratio of 3 and adopted an extinction of $A_K \sim 1.8$ mag. Using the relation given in Rieke and Lebofsky (1985) this corresponds to $A_V \sim 16$ mag.

Dionatos et al. (2010) carried out observations using the Spitzer Infrared Spectrograph covering wavelengths of 5.2–37 μm. They studied the excitation conditions of the less energetic part of the H_2 emission, the ortho-to-para ratio of H_2 and employed shock models to recover the observed emission. The fitting of Boltzmann diagrams to observed H_2 line emission shows two separate components with excitation temperatures of about 300 K and about 1000 K. The OPRs determined are around the warm equilibrium value of 3. The authors found the emission to be consistent with C-type shock simulations.

Using the Photoconductor Array Camera and Spectrometer (PACS) of the Herschel Space observatory Dionatos et al. (2018) observed line emission of the species CO, H_2O, OH and [O I] in the wavelength range of 55 – 210 μm. That study included three pointings along the HH211 outflow and investigated the excitation conditions of CO and H_2O. In the pointing covering the blue-shifted lobe there are two components with excitation temperatures of around 370 K and 750 K for the CO emission. The excitation temperature of the H_2O emission is around 90 K. The ortho-to-para ratio of water is 0.65 — this implies that the water observed was formed at very low temperatures and has not been thermally reprocessed.

Lee et al. (2018) studied the highly extincted area around the driving source with the Atacama Large Millimeter/submillimeter Array (ALMA). In the continuum emission around 350 GHz they detect a faint elongated envelope with an embedded dusty disk. The direction of the disk's and the envelope's major axis is nearly perfectly perpendicular to the jet axis. The disk extends ~ 30 au along its major and ~ 16 au along its minor axis. It is optically thick and warm and has a mass of $\sim 35\,M_{Jup}$. From the rotation of the disk Lee et al. (2018) estimated the mass of the central protostar to be $\lesssim 50 M_{Jup}$. Additionally, they detected a compact, rotating outflow in the SO $N_J = 8_9$ - 7_8 transition. This outflow could be carrying away significant amounts of angular momentum from the inner disk.

The most recent publications make use of a guaranteed time observation program (Ray et al., 2017) using the Near InfraRed Camera (NIRCam), the Near InfraRed Spectrograph (NIRSpec) and the Mid InfraRed Instrument (MIRI) of the James Webb Space Telescope (JWST). My thesis investigates data from that same observational program. As such a more detailed introduction to the NIRSpec data can be found in section 2.2.

Ray et al. (2023) identify the most prominent lines of molecular hydrogen and carbon monoxide in NIRSpec's observed wavelength range of 1.7 to 5.2 μm. The NIRSpec observations encompass the terminal bow shocks of the blue-shifted lobe. Other than a weak forbidden line of ionized iron, no notable emission from atomic or ionized species was observed. This implies that most of the outflow is molecular. Additionally, Ray et al. (2023) investigate the ratio of H_2 to CO emission and tangential velocities. They find that the emission in most regions other than the bow shocks is dominated by hydrogen emission. The tangential velocities measured are around 80-100 km/s, it is, however, likely that the shock velocities might be a lot lower. A strong argument for this is the absence of atomic and ionized species as shocks with velocities around 100 km/s would be energetic enough to dissociate the molecules. Internal working surfaces — where shocks are caused by faster parts of the jet overtaking slower ones — could explain these lower shock velocities. The shock velocity would then be determined by the relative velocity between the different parts and could be as low as few 10 km/s.

The focus of Caratti o Garatti et al. (2024) is on the Mid InfraRed Instrument (MIRI) observations of the same program. These observations cover the entire blue-shifted lobe, the position of the driving source and a small part of the red-shifted lobe over a wavelength range of 5 to 28.5 μm. This range includes emission lines of large number of molecular, atomic and ionized species. Along the jet and in the knots and shocks the observed emission stems — among others — from H_2, HD, OH, [S I], [Fe I] and [Fe II], [Ni II], CO and possibly CO_2. These different species trace different parts of the system as discussed in section 1.3.
While both atomic and molecular species trace the jet, their distribution varies significantly. The jet diameter determined by the atomic species is smaller than that of the molecular species. This implies a layered jet structure.
The material observed in the wake of the shocks and near their wings is, however, almost purely molecular and likely ambient gas swept up by the jet. It is mostly observed in low-excitation emission of H_2. These lines also trace a less collimated molecular wind.
The bow shocks are richer in chemistry compared to the rest of the system. This is especially true for the bow shock second-closest to the driving source — in the observed wavelengths it is the brightest of the system. Here, lines from a number of additional species like OH, [Cl I] and [Cl II], [Ne II], [Ar II], [Co II] and [S III] are detectable. The observed OH lines are a characteristic result of photodissociation of water. The UV radiation needed for that would be caused by strong shocks with shock velocities larger than 40 km/s.

Additionally, Caratti o Garatti et al. (2024) investigate the excitation conditions along the blue-shifted lobe of the outflow.
For this, a reliable extinction correction is essential. The authors determine the visual extinction by comparison of NIRCam data using narrow-band filters centered on the (1-0) S(1) and (1-0) O(5) lines of molecular hydrogen. They measure the highest extinction towards HH211 mm — the embedded protostar. Patches of higher extinction are also present along some knots of the jet and at some bow shocks. The values reported for the terminal bow shocks range from 5–15 mag.
Using these values Caratti o Garatti et al. (2024) correct each line in each pixel for the extinction and then calculate a Boltzmann diagram for each pixel. They include rotational (0-0) and (1-1) transitions of H_2 and find that the system is not well described by a single Boltzmann distribution. So they allow for two separate components — this is consistent with the properties of shock excited H_2 introduced in section 1.4.2. The two components — called warm and hot — have significantly different properties that vary strongly along the outflow. The following results apply to the region of the terminal bow shocks. For the warm component, Caratti o Garatti et al. (2024) find excitation temperatures around 900–1000 K and column densities of some 10^{19} cm^{-2}. An exception to this is the bow shock already mentioned, with a column density of $\sim 3 \cdot 10^{20}$ cm^{-2}. The excitation temperatures of the hot component range from about 2000 to about 3500 K. Its column densities are between some 10^{18} and 10^{19} cm^{-2}.

The methods used in this thesis are very similar to those used in Caratti o Garatti et al. (2024) and they are described in detail in chapter 5. This similarity makes that study a prime point for comparison of my results.

2 Instrumentation and Data

This thesis utilizes pipeline processed early science data from the James Webb Space Telescope (JWST) — specifically its Near InfraRed Spectrograph (NIRSpec). The following section provides additional information about the observations, the resulting data and the instruments used to capture it. Furthermore, I introduce the line catalog providing all the theoretical line properties used in my analysis.

2.1 The James Webb Space Telescope (JWST)

As the currently most sensational space telescope there are a lot of facts and details to be mentioned about JWST. A short general introduction is followed by more information about the instruments and observing modes used to capture the data that I analyzed for this thesis.

Developed by an international collaboration of NASA, ESA and CSA, the JWST is sensitive to near infrared and mid infrared radiation at wavelengths of 0.6 to 28 μm. The 6.5 m effective diameter of its primary mirror in combination with its positioning outside earth's atmosphere allow for unprecedented sensitivity and resolution in these wavelengths. This opens new opportunities for high red-shift observations as well as observations of dust-rich environments unobservable in optical wavelengths such as star-forming regions and active galactic nuclei. Additionally, the peaks of the thermal radiation of cool objects and a number of fundamental molecular transition such as those from H_2 and CO fall into this wavelength range. JWST was launched in 2021 and reached its final orbit around the Lagrange point 2 (L2) of the sun-earth system in 2022. Using a combination of active and passive cooling the telescope operates at temperatures between 6 and 55 K. This greatly limits internal thermal emission resulting in the zodiacal light being the limiting factor for wavelengths below 12.5 μm; above this wavelength it is the telescope's thermal emission (Gardner et al., 2023).

A significant part of the telescope was deployed after launch — this includes parts of the primary mirror, the secondary mirror and the sunshields. Following this, the telescope alignment was done on-orbit reaching the diffraction limit for wavelengths down to 1.1 μm and surpassing the design goal of diffraction-limited observations down to

wavelengths of 2 μm. With a pointing stability of $\sim$ 1.5 mas the pointing system also outperforms the pre-launch requirements (Gardner et al., 2023). The following sections describe the four science instruments aboard the JWST. As the data analyzed in this work was captured with NIRSpec, I give more details about this instrument.

2.1.1 Near InfraRed Camera (NIRCam)

NIRCam allows for imaging observations in a wavelength range from 0.6 to 5.0 μm. Its 29 filters include different band widths from extra wide filters to narrow-band filters designed to capture the emission from individual line transitions. Due to its optical design NIRCam typically observes in two parallel fields of view (FOVs) — one for wavelengths shorter than 2.4 μm and one for longer ones — across a total of 10 detectors. Additionally, NIRCam is capable of wide-field slitless spectroscopy from 2.5 to 5.0 μm and coronagraphic observations. A dedicated Bright Object Time Series (BOTS) observing mode is available for both imaging and spectroscopic observations to monitor the variability of relatively bright sources (Gardner et al., 2023).

2.1.2 Near InfraRed Imager and Slitless Spectrograph (NIRISS)

Like NIRCam, Near InfraRed Imager and Slitless Spectrograph (NIRISS) operates in a wavelength range of 0.6 to 5.0 μm. Its 9 filters are made to match those of NIRCam and its imaging capabilities are mostly redundant; it can, however, be used in parallel for an increased area coverage. NIRISS' single object slitless spectroscopy works for wavelengths between 0.6 and 2.8 μm and defocuses bright sources to avoid saturation. Also similar to NIRCam, NIRISS has its own wide-field slitless spectroscopy mode which allows for low spectral resolution observations in the wavelength range of 0.8 to 2.2 μm. Two different gratings in combination with the various filters help avoid overlapping spectra in denser fields. The addition of a seven-aperture mask to the pupil wheel enables the use of Aperture Masking Interferometry — a method for high-contrast imaging for sources with low angular separation (Gardner et al., 2023).

2.1.3 Mid InfraRed Instrument (MIRI)

In contrast to both instruments discussed so far MIRI covers wavelengths of 5 to 28 μm. Being the only instrument aboard the JWST operating in this wavelength range, MIRI is designed for versatility and has standard and corona graphic imaging as well as low-resolution spectroscopy and medium-resolution integral field spectroscopy capabilities. As these observing modes are vastly different from each other, they are optically distinct

with the imagers and the low-resolution spectrometer sharing a field of view, while the Medium Resolution Spectrometer (MRS) has a separate FOV. The subsystems do, however, share an active cooling system that maintains an operating temperature of 7 K (Wright et al., 2015).

The MIRI imager encompasses both imaging modes as well as the low resolution spectroscopy. In the standard imaging mode diffraction limited observations using one of 9 filters can be obtained. The field of view is $74'' \times 113''$ and the detector pixels have a projected size of $0.11''$ — slightly below the Nyquist criterion for the shortest wavelengths. Coronagraphic observations use the side of that same FOV. There are four units operating at different but fixed wavelengths. While the longest wavelength coronagraph has a conventional design with a central occulted stop, the other three rely on the interference introduced by four-quadrant phase masks to cancel out the bright central source.
Using the prism mounted on the imager filter wheel low resolution spectra can be observed. Their spectral resolution varies from $\sim$ 40 to $\sim$ 160 over the nominal operational wavelength range of 4.5 to 10 μm. The $0.51''$ by $4.7''$ slit is located in the structure separating the imaging and coronagraphic fields. Additionally, when using the prism the imager acts a slit-less wide-field spectrometer dispersing all sources present within this field of view (Wright et al., 2023).

The Medium Resolution Spectrometer (MRS) operates in an integral field observing mode. This was chosen over a long-slit for a number of reasons including easier target acquisition, less loss of light and a practical way of observing objects that might have different centers of emission at different wavelengths.
To allow for spectral resolving powers of 1300 to 3700, the MRS operates in 4 separate spectral channels. They have co-spacial fields of view and optics individually optimized for each one. This channel split is done by three serial dichroic filters. Due to the limited detector space only a sub-band of each spectral channel can be observed in an individual observation. Each of these sub-bands covers about one third of the respective channel making three observations necessary to cover the entire 5 to 28.5 μm wavelength range. The use of sub-bands allows for further optimization of the optical components as they are only used for a relatively small wavelength range each.
As mentioned before, each channel has its own integral field unit (IFU). They all utilize image slicers whose slice widths are designed to match the wavelength-scaled point-spread function of each channel. After the FOVs are sliced and rearranged, the light reaches the dispersive elements where each channel uses a different grating for each sub-band. The dispersed beams are re-imagined onto the detectors. Here the two short-wavelength channels share one detector unit and the two long-wavelength channels share the other detector. The resulting observations have different sized FOVs and different projected pixel sizes (Wells et al., 2015).

2.1.4 Near InfraRed Spectrograph (NIRSpec)

NIRSpec offers single object as well as multi-object spectroscopy (MOS) in addition to its IFU. Using one of 7 dispersive elements a number of different spectral resolving powers can be achieved over a wavelength range of 0.6 to 5.3 μm for a low spectral resolution or 0.7 to 5.2 μm for a higher one. They are low resolution prism spectroscopy — which reaches $R \sim 100$ — and mid and high resolution grating spectroscopy — which allow for $R \sim 1000$ and $R \sim 2700$ respectively. Four long-pass interference filters are used for separating the different orders of diffraction. The detector consists of two Teledyne 2048×2048 H2RG arrays that are mounted next to each other; between their light sensitive regions lies a gap of 2.8 μm. This gap appears in dispersion direction and leads to some parts of observed spectra not being recorded. A number of dither patterns are available that can help recover additional wavelength coverage as well as improve the spacial and spectral resolution by allowing for a better sampling rate.
For the single object mode one of five fixed slits of various sizes can be used. These fixed slits occupy their own space on the detector and are always open (Gardner et al., 2023).

The MOS mode utilizes a Micro Shutter Array (MSA) consisting of 4 quadrants of 365×171 individually controllable shutters. This flexibility allows for simultaneous spectroscopic observation of 50 to 200 targets over a total field of view of $3.6' \times 3.4'$. With any desired configuration some of the shutters cannot be controlled correctly. They can fail in either an open or closed state, either contaminating the desired spectra or obscuring part of the desired slit. The multi-object mode is mostly designed for the low and medium dispersive elements. Using the high resolutions grating increases the length of the spectra leading to a large cut-off from the red end of spectra observed through any of more than half of the shutters (Ferruit et al., 2022).

Between the four quadrants of micro-shutters is the aperture of the IFU. It covers a continuous field of view of $3.1'' \times 3.2''$. Over a number of folding and relay mirrors the incoming light is magnified and directed to the image slicer. It consists of 30 mirror facets, that make up the 30 virtual slits of the instrument. From there, each partial beam is reflected towards a dedicated pupil mirror followed by a slit mirror. This is done to arrange the images of the 30 individual slits into a single virtual long slit image, that is folded once more before reaching the spectrometer optics. After filtering and dispersion the beam reaches the detector. The resulting spatial resolution — the spaxel size — is $0.103'' \times 0.105''$. It is determined by the slice size in one direction and by the pixel size in the other one (Böker et al., 2022).

It should be noted that the MOS and IFU observing modes utilize the same detector space making them mutually exclusive. A consequence of this is contamination of IFU observations by failed-open micro-shutters, as well as a low-level leakage through the entire MSA — the upper panels of Figure 2.1 visualize this.

Observing separate "leakage exposures" — with the same observational parameters except closing the IFU aperture — helps mitigate this effect. These exposures can then be used to subtract any additional light from the science exposures (Böker et al., 2022).

Another challenge is the already mentioned physical gap between the two detector units. For the IFU mode the high-resolution gratings will disperse the spectra onto both detectors as well as the gap in-between. Due to the optical construction of the IFU — which introduces a slit-tilt — the wavelength ranges that are lost are different for each slice. Dithering the image helps reduce the impact — the wavelength gaps can, however, not be fully eliminated. These wavelength gaps can be seen in the lower right panel of Figure 2.1 (Böker et al., 2022).

2.2 JWST observations of HH211

At the core of this project are the observations of HH211 associated with the cycle 1 guaranteed time observations program 1257 - "The Young Protostellar Outflow HH211". The program consists of four separate observations around HH211: a mosaic of MIRI MRS observations encompassing HH211's entire south-east lobe, corresponding MIRI MRS background observations, a number of exposures of NIRCam with various different optical elements and a mosaic of NIRSpec IFU pointings observing the terminal bow shocks of the south-east lobe (Ray et al., 2017). As this work makes use of the NIRSpec IFU observations they are described in more detail.

Observed on August 28th of 2022 the NIRSpec data encompasses the three furthest bow shocks of the south-east lobe in a mosaic covering roughly 9″ by 7″. Two grating and filter combinations — G235H/F170LP and G395H/F290LP— were used. Additional on target observations were done with the entry pupil to the IFU unit closed in order to correct for the leakage of failed-open shutters in the MSA. For the dithering pattern a two-point nod was chosen (Ray et al., 2024).

The raw observational data was pre-processed and calibrated using version 1.11.4 of the JWST science calibration pipeline (Ray et al., 2024). I accessed the resulting fully calibrated and combined data sets via the online user interface of the ESA JWST science archive after the end of the exclusive-access period. These two separate data cubes — with wavelength coverage of 1.70 to 3.15 μm and 2.88 to 5.20 μm and wavelength channels widths of 0.000395 μm and 0.000665 μm — are the starting point for my following analysis.

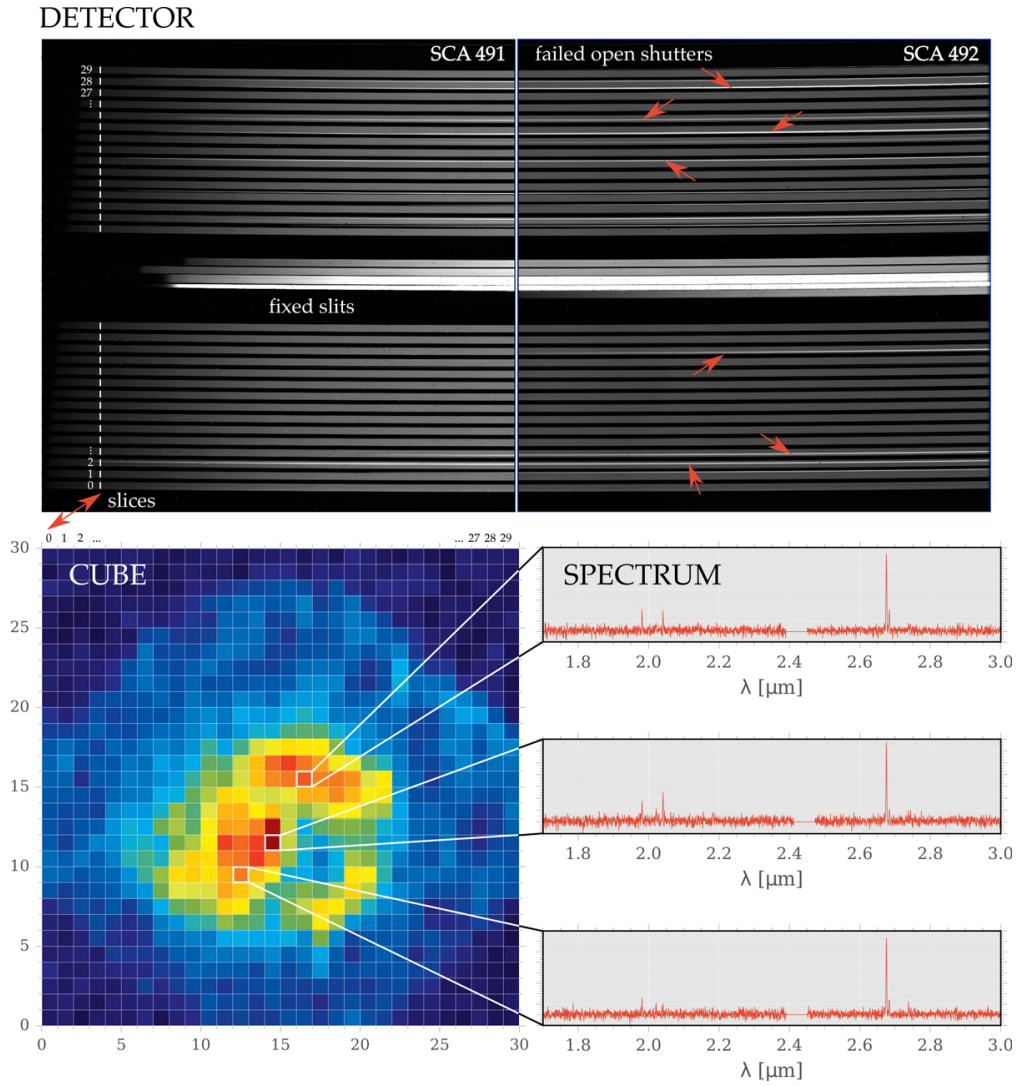


Figure 2.1: A visualization of NIRSpec data taken from Böker et al. (2022). The top two panels show count rate maps as detected by the two detector units — this specific image was observed during ground testing. The virtual slits are labeled by their number and contamination caused by failed-open MSA shutters is indicated by the red arrows. Between the two groups of spectra from the IFU the spectra from the fixed slits can be seen. The bottom panels visualize a reconstructed data cube — this one was generated from a simulated observation. The plot on the left side shows a channel map plotting the fluxes measured in one specific wavelength channel for all pixels. On the right, spectra extracted from 3 pixels are shown. In this example the spectra were simulated for the G235H grating. The wavelength gap — and the variation of its position depending on the slice — can clearly be seen.

The first step after acquiring the data from the archive, was data exploration. Given the nature of IFU data cubes, there are two main ways of studying them: investigating the spectra of each pixel or comparing the flux maps of different wavelengths. This work mostly analyses the spectra on a pixel-by-pixel basis.

In Figure 2.2, I show an approximate recreation of Figure 1.3 — using the NIRSpec data. Comparing the two figures, it is clear just how different the data obtained from the different instruments is. While NIRCam covers the entire outflow in a single field of view, the NIRSpec observation is a mosaic of 10 tiles and still only covers the terminal bow shocks (Ray et al., 2017). Given that both instruments operate in the same wavelength region and on the same telescope, differences in spatial resolution are due the different plate scales. NIRSpec's plate scale is significantly smaller than NIRCam's, which results in a more pixelated appearance. The key difference, though, is that the NIRSpec observations contain the full spectral information of the near infrared emission (Gardner et al., 2023). The image shown in Figure 2.2 represents just three of thousands of spectral channels in those data cubes.

Figure 2.3 shows the spectrum extracted from the pixel centered at RA 03:43:59.931 DEC +32°00'34.38". Disregarding the two small gaps, the spectrum spans the 1.70 to 5.20 μm wavelength region and contains numerous H_2 and CO lines. As this pixel is located in the bright and highly excited terminal bow shock, it shows a particularly line rich spectrum. I will use this spectrum to illustrate the data analysis process in the following chapters.

2.3 Line catalog

The methods described in the following chapters 5 and 6 rely on the accurate knowledge of intrinsic line and transition properties. As I study exclusively the emission from molecular hydrogen and numerous H_2 lines fall into the observed wavelength region — as seen in Figure 1.2 — a catalog of H_2 transitions is needed for line identification. Table 2 from Roueff et al. (2019a) provides all the necessary information. It contains transition energies, wavelengths, probabilities and coefficients of any possible transition of H_2 in the infrared. Additionally, it lists both the upper and lower level rotational and vibrational quantum numbers of those transitions, as well as the corresponding upper state energies and their statistical weights.

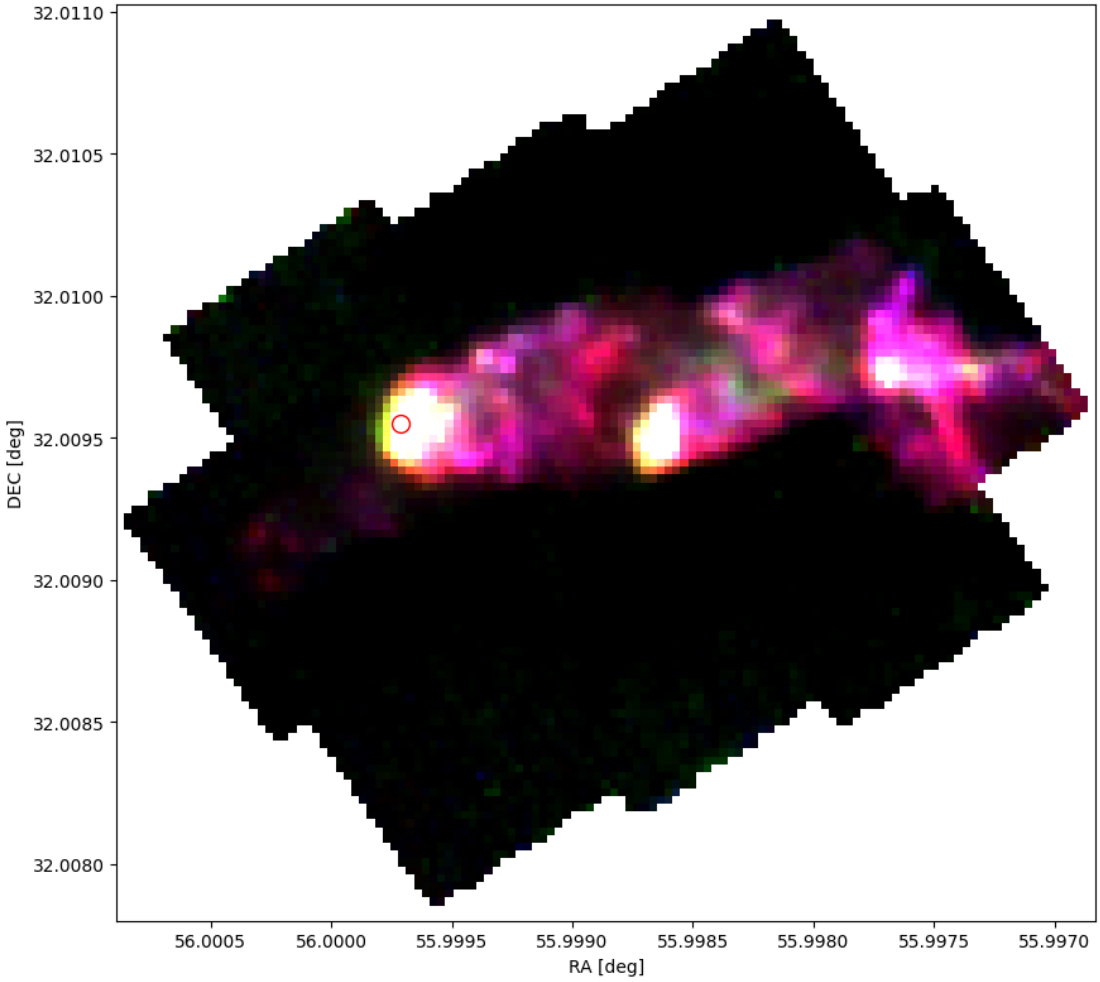


Figure 2.2: A near infrared image of the terminal bow shocks of HH211 as observed by NIRSpec. Similar to Figure 1.3, here blue corresponds to the emission of the H_2 (1-0) O(5) line, green to the CO R-branch and red to the H_2 (0-0) S(9) line. In this plot I used individual spectral channels rather than emulating a filter curve. The specific channels used are centered at 3.235 μm (blue), 4.520 μm (green) and 4.695 μm (red). These narrow channels lead to a different appearance between the two figures. The red circle marks the position of the example pixel centered on RA 03:43:59.931 DEC +32°00'34.38", that I use to illustrate data analysis in the following chapters.

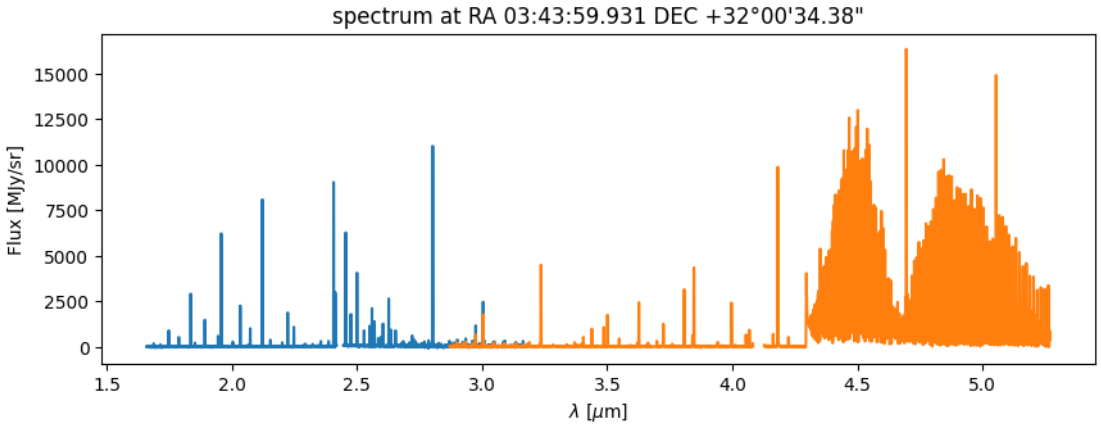


Figure 2.3: The spectrum of the pixel centered at RA 03:43:59.931 DEC +32°00'34.38" — the same one as marked in Figure 2.2. The different colors correspond to the different observations and the two gaps seen around 2.4 and 4.1 μm are the detector gaps discussed in section 2.1.4. Due to this pixel's position in the terminal bow shock, its spectrum is particularly rich in lines. All of the strong and separated lines stem from H_2. The dense groups of lines longwards of 4.25 μm are the R- and P-branches of the CO (1-0) transitions. I use this pixel to illustrate data analysis in the following chapters.

I acquired the catalog via the VizieR catalog access tool. To avoid unnecessary computational time for lines outside of the observed wavelength range, I compare only the relevant part of the VizieR catalog Roueff et al. (2019b) to the observed lines. An additional selection criterion is an upper state energy T_u lower than 30000 K. My testing has shown, that lines with higher energies are only rarely detected in the bow shocks and not at all in the other regions.

Some calculations, such as those of the partition functions required for the excitation analysis, need to consider all possible states. For those calculations I extract the state properties from the full catalog — including states with lines outside of the observed wavelength region and high energy states. Additionally, I add the properties of the two ground states ($\nu = 0$, J = 0 and $\nu = 0$, J = 1) as they are not the upper states for any transitions and, thus, missing from the catalog. I calculate the energies of these states by subtracting the energies of the 0-0 S(0) and 0-0 S(1) transitions from their corresponding upper states.

3 Custom Python Routines: *nirspecpy*

For this project, I developed and implemented a suite of custom python routines, which I named *nirspecpy*. This set consists of three modules: *calc*, which provides powerful functions for data analysis; *ut*, including utility functions to simplify data handling; and *plot*, which facilitates visualizing results.

As these functions are purpose-built for this project and with the studied data set in mind, they align seamlessly with the analysis methodology used here. Because of this, the detailed implementation and choice of necessary hyper-parameters of each function are best introduced as the methodology is explained. Each step of extracting line properties — baseline correction, line fitting and line matching — has a corresponding function, that I explain in detail in sections 4.1, 4.2 and 4.3. Another function allows the fitting of excitation diagrams — a full introduction of this method and the respective function follows in chapter 5.
The functionality of my routines depends on the following existing python packages:

- *NumPy* for fast and easy manipulation of data arrays (Harris et al., 2020)
- *AstroPy* for loading and manipulating the initial fits files as well as construction of proper coordinate systems from the fits file's header (Astropy Collaboration et al., 2013; Astropy Collaboration et al., 2018; Astropy Collaboration et al., 2022)
- *uncertainties* for handling uncertainties (Lebigot, 2010)
- *SciPy* for finding peaks in the spectrum and for physical constants (Virtanen et al., 2020)
- *LMFIT* for curve fitting (Newville et al., 2024)
- *matplotlib* for plotting (Hunter, 2007)
- *pandas* for manipulation of table files (McKinney, 2010; The pandas development team, 2024)

Additionally, I used the package *multiprocess* for parallelization (McKerns et al., 2011; McKerns and Aivazis, 2010) in the further analysis. A custom function forms the interface between my analyzing functions and the *multiprocess* framework — more details are found in chapter 7.

4 Extraction of Line Properties

Before any physical properties — like velocities, densities or temperatures — can be derived, the properties of the individual lines have to be extracted from the spectra. I do this in three separate steps.

In the first step, the spectra need to be baseline corrected and the average noise level needs to be determined. This is important, as the analysis in the later steps will assume that the spectral lines are not vertically offset. The noise levels estimated here will be used for determining signal-to-noise ratio cut-offs.
As the spectra are dominated by line emission and no continuum emission is seen — e.g., in Figure 2.3 — I assume a constant baseline. The value of this baseline is determined by examining the distribution of flux values across each spectrum. Assuming Gaussian noise and an overall low baseline level, one expects to see a nearly Gaussian distribution of flux values centered around some value close to zero. The mean of this distribution would then be the baseline and its width would measure the noise.
Naturally, this distribution is also influenced by the observed spectral lines. However, if the line fluxes are high compared to the baseline and noise level, their fluxes should lie outside of the Gaussian peak caused by the noise. Then it should be possible to separate the lines from the noise by masking high fluxes during the baseline correction. As will be discussed in section 4.1, this works well in most cases, but the CO branches longwards of 4.25 μm are challenging.

The aim of the second step is determining the line parameters — flux, central wavelength and line width — of each observed line in each pixel. This is done by fitting a Gaussian function to each peak in the baseline corrected spectra. From these fit parameters I extract the line properties. For the future steps it is practical to describe the lines in terms of their central wavelength (the mean of the fit), their width (the fit's standard deviation) and the integrated line intensity I (the area under the fit's curve). From this step onward, I only consider spectral lines that have a peak flux above 5 times the noise calculated in the first step. This is done to greatly reduce the probability of misinterpreting noise for spectral lines.

Building on the previous steps, the third step is matching the observed lines to the corresponding lines in the line catalog. After all, the analysis in chapters 5 and 6 relies on a combination of observed and intrinsic line properties. I accomplish this match by comparing the observed central wavelengths to the theoretical ones. As illustrated in section 4.3, this process is rather simple in case of the observations investigated here. The low radial velocities result in minimal Doppler shifts, which do not have to be taken into account separately.

4.1 Baseline correction

I implemented the baseline correction step just described in a routine named *nirspecpy.calc.baseline_fit*. This function accepts a spectrum as input and returns the parameters of the fit to the flux distribution and the baseline-corrected spectrum. If measurement errors are passed to the routine, they are updated to include the uncertainty introduced by the baseline correction using Gaussian error propagation. Future steps use these updated errors. Because of the challenges encountered and discussed in the following, I also implemented an automatic visualization option for inspecting these fits as well as basic error handling ensuring output in the expected format even if the fit does not converge.
In Figure 4.1, I show a zoomed version of the spectrum also shown in Figure 2.3. While the baseline offset appears small compared to the line fluxes, not correcting it will lead to inaccurate line intensities in the later steps.

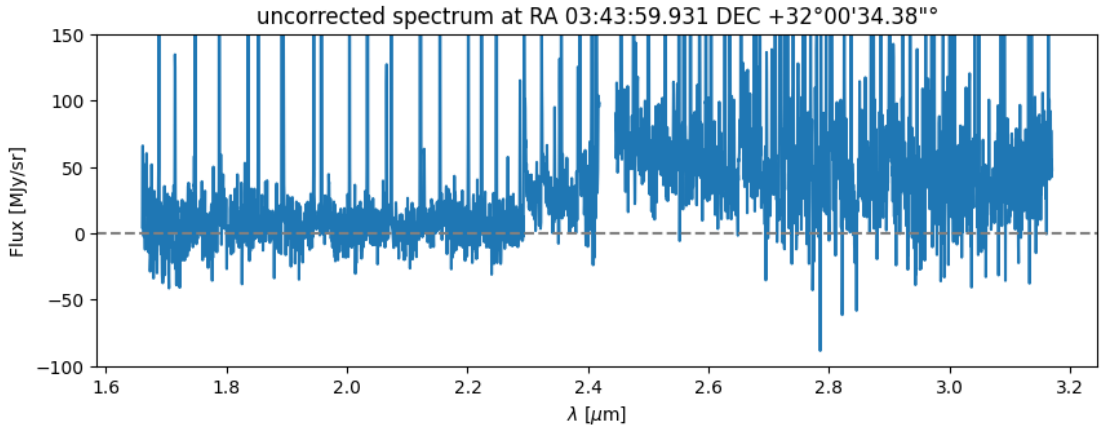


Figure 4.1: Just the shorter wavelength observation of the spectrum shown in Figure 2.3. The range of the y axis is chosen to highlight the baseline of the spectrum. For both parts, the baseline is non-zero. Comparing the two, it is clear, that they share neither the same baseline nor the same noise level. The discussion of this challenge follows in section 4.1.2.

4.1.1 Determining the noise distribution

At its core, this routine investigates the distribution of flux values and fits a Gaussian to a peak close to zero. For this to happen, the routine must first calculate a histogram — with *some reasonable number* of bins, that cover the *relevant flux range*. I opted to use the *numpy.histogram* function for this. After testing this function on multiple spectra, I decided to use 50 bins of uniform width for each histogram and to control the bin width by changing the covered range.
Keeping the number of bins fixed and relatively low helps to limit the number of (nearly) empty bins and the computational effort needed. Doing this it becomes all the more important to cover the relevant flux range precisely. Setting this range too small could lead to the peak, caused by noise, not being fully contained. On the other hand, using a too large range could place the entire peak in the same bin. In both cases the fit to the distribution will not be reliable. That said, the relevant flux range changes between pixels and observations, as the baseline and noise tend to be different.

I addressed this challenge by calculating summary statistics of the spectrum first and using those to determine the needed flux range. In a spectrum containing only noise, the average of fluxes would be a good estimation of the baseline and the standard deviation of fluxes would describe the noise quite well. In that case, one could define the relevant

flux range centered on the mean and with a width of some standard deviations. For a noise-only spectrum, this is enough to expect that the full distribution is covered and there is little excess space to the sides.

However, if the spectrum contains numerous strong spectral lines, they heavily skew those summary statistics towards higher values. A way to mitigate the effect of those lines, is to mask high flux values. The average and standard deviation, calculated using the masked spectrum, can then be used to set the bounds of the flux range covered by the histogram. I set the lower bound to the average minus 2 times the standard deviation and the upper bound to the average plus 2 times the standard deviation. I chose to use twice the standard deviation, as the calculated value is still likely influenced by spectral lines and, thus, likely overestimated. This flux masking value is set using a parameter (*max_noise*) in the function and because it is only used in a first estimation, a wide range of values still lead to good results. Figure 4.2 shows two such histograms with their corresponding fits — once without masking high flux values and once masking all values above 250 MJy/sr. In the former case the distribution is not well defined, as only few bins contribute to the fit. In the latter case, the distribution is nicely spaced across the 50 bins. In the analysis presented in this thesis, I set a flux masking value of 250 MJy/sr after testing several values.

Having gone through all this to determine the histogram of the flux distribution, fitting a Gaussian to said distribution is rather straightforward and can be done using the *GaussianModel* of the *lmfit* package. I subtract the mean of the fitted distribution from the spectrum to correct for a non-zero baseline and adopt the fitted standard deviation as the noise for future signal-to-noise calculations. Referring back to Figure 4.2 again, neither fit is particularly good and the lower plot reveals another problem, that will be discussed in the next section: the distribution of fluxes does not appear to be a (single) Gaussian distribution.

4.1.2 Fitting baselines for partial spectra

As the bottom plot in Figure 4.2 shows, the distribution of the flux values is not well described by a single Gaussian curve. A closer look at the baseline of the spectrum in question — as seen in Figure 4.1 — reveals the reason: baseline and noise are not constant within the spectrum. It becomes clear that the two detectors capturing each spectrum — as described in section 2.1.4 — can have significantly different noise properties. As a result of this, the method described in section 4.1.1 is a lot less effective when applied to the entire spectrum.

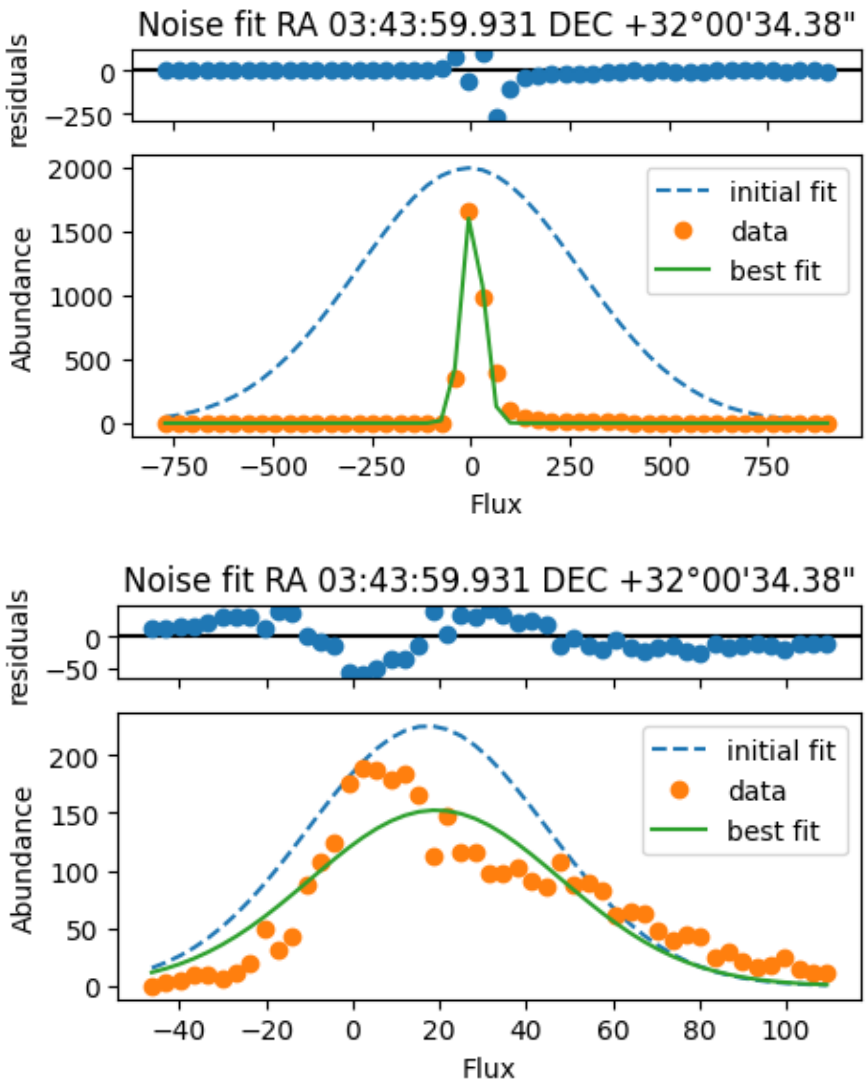


Figure 4.2: Examples of the noise fits for the shorter wavelength observation in the pixel centered at RA 03:43:59.931 DEC +32°00'34.38". In the upper plot no flux values were masked when setting the limits for the histogram and in the lower plot all flux values above 250 MJy/sr were masked. It is clear from the upper plot that not masking high flux values leads to far from ideal results. Interestingly, the fit in the lower plot is also not particularly good. This is due to the different noise properties of the two detectors capturing this spectrum and will be discussed in section 4.1.2.

Rather than fitting two convolved Gaussian distributions to compensate for this, I opt to split the spectra into parts and do their partial noise statics independently. They are cut at the position of missing data caused by the physical gap between the individual detector units. If the spectrum was split into parts, the baseline correction is individually applied to each part and the spectrum reassembled before it is returned. The results of the function also include the parameters for each individual fit as well as the additional *cuts* array. It holds the cuts applied and can be passed to the line fitting routine so that the different noise levels are considered when cutting lines below a certain signal-to-noise threshold. This functionality can be disabled using the *allow_multiple* parameter. From the fits to the partial spectra shown in Figure 4.3, one can see that splitting the spectrum into two parts, indeed, has the desired effect. The flux distribution of either partial spectrum shows a Gaussian peak close to zero, that can be fitted accurately.

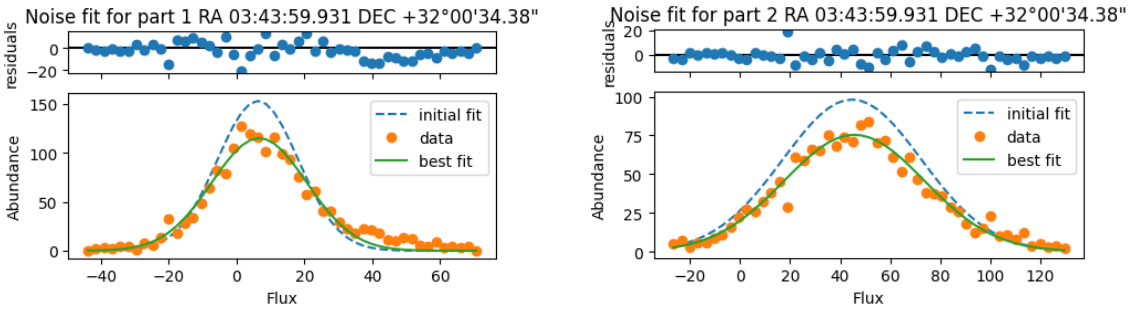


Figure 4.3: Examples of the partial noise fits for the shorter wavelength observation in the pixel centered at RA 03:43:59.931 DEC +32°00'34.38". For both partial spectra the distributions are well described by a single Gaussian. This also shows that the feature seen in the bottom part of Figure 4.2 is indeed a convolution of two Gaussian distributions caused by the different properties of the different detectors.

Some of the spectra show numerous small gaps with only small rows of data between them. Considering each of those partial spectra individually leads to failed noise fits due to insufficient data. Therefore, I implemented an optional requirement for a minimum number of data points in each partial spectrum. If any partial spectrum has fewer data points than this value, it is combined with the next one — via trial and error a value of 20 has proven effective.

Using the baseline parameters obtained by fitting to the partial spectra, any baseline differences can be compensated. Figure 4.4 shows the baseline-corrected version of Figure 4.1. Naturally, the noise is still present and still different between the parts of the spectrum. The baselines, however, are now both zero.

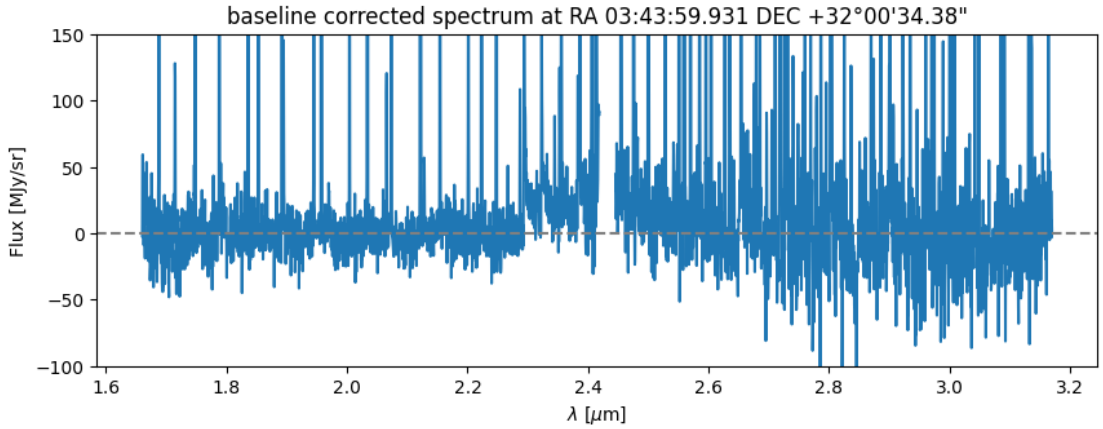


Figure 4.4: The baseline-corrected version of Figure 4.1. This spectrum has been corrected using the results obtained from fitting to the partial spectra. While the noise is still different, the baselines of both parts are now on average zero.

4.1.3 Baselines in spectra containing CO branches

So far the descriptions and examples have focused on the shorter wavelength observation. For that observation, the methods described work well and reliably. The same is true for the shorter wavelength partial spectrum of the observation covering the wavelength range from 2.88 to 4.1 μm. In the other partial spectrum, however, the very dense CO branches observed in some pixels push this method to its limits. As Figure 4.5 shows, there are only few spectral channels contributing to the baseline in the second partial spectrum. This results in baseline fits are shown in Figure 4.6, where a small peak is accompanied by a lot of bins representing the flux of the dense CO lines. Depending on the specific pixel, these fits might converge to a very wide distribution including the channels that measure the wings of spectral lines. Fine-tuning the flux masking parameter helps, limit this influence and ensure a proper baseline fit.

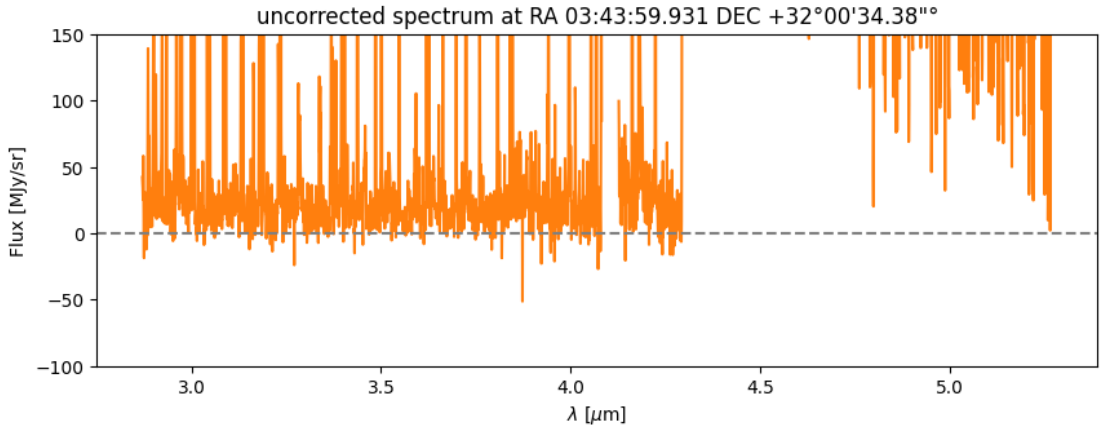


Figure 4.5: Just the longer wavelength observation of the spectrum shown in Figure 2.3. For the parts of the spectrum containing the CO branches there is no baseline present between the individual lines. This is a challenge for baseline correction routine.

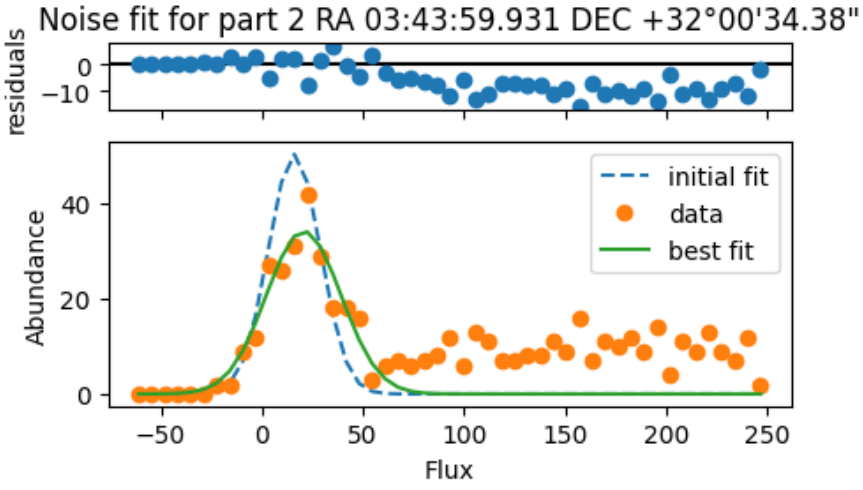


Figure 4.6: The baseline fit done for the partial spectrum containing the CO branches. The small peak is caused by the baseline at wavelengths shorter than ~ 4.25 μm; towards higher fluxes the contribution of the spectral lines can clearly be seen. By masking high flux values as described in section 4.1.1, the baseline correction is challenging, but possible.

4.2 Line fitting

After the baseline correction, the lines in each spectrum can be fitted and their properties extracted. For this purpose I developed the *nirspecpy.calc.line_fit* routine. It takes a corrected spectrum and a required minimum height of spectral lines as parameters and returns the fitted line properties — such as integrated line intensities, central wavelengths and line widths — of all the successful fits. I chose to require a minimum signal-to-noise ratio of 5 using the noise determined in the last step.

In a first step, the routine locates all peaks exceeding the required minimum height within the spectrum. Using this information as an initial guess, it fits each individual line. If the relative errors in amplitude, central position and standard deviation of a given fit are smaller than the maximum allowed errors, it is added to the list of results. These maximum errors are all 0.1 by default, but can be set individually. For the analysis presented in the following, I used values of 0.25 for each of them. After a fit has been attempted for each peak, the results are returned and the numbers of failed fits and fits discarded due to too high errors are reported. All of the parameters mentioned so far, can also accept an array instead of an individual value. This is to support different requirements — such as different noise levels for a required signal-to-noise ratio — for different parts of the spectrum. It only works properly if the cuts found in the baseline fitting routine are also passed to the line fitting routine and the number of values passed to each parameter is either one or equal to the number of partial spectra defined by the cuts.

In Figure 4.7, I show the line fits to the shorter wavelength part of the example spectrum together with the residuals of those fits. The routine successfully fitted all of the bright lines and the residuals are low.
Similar to the challenges encountered in the baseline analysis, the CO branches also present difficulties for my line fitting routine. I illustrate this showing the fits and residuals of the longer wavelength part of the example spectrum in Figure 4.8. The routine fits each spectral line individually, which works well for fairly isolated lines. This is, however, not the case in the CO branches, where lines are very dense and partially overlap. In those cases the fits might not converge to the intended feature and some lines are missed entirely.

This routine also supports recursive line fitting which can help with close lines. Defining a number of steps that is larger than 1 will lead the function to carry out multiple of the steps described above. After each step, the successfully fitted lines are subtracted from the initial spectrum and the function is called again using the residuals as the new input. Once all steps are done, the results are combined and nicely packaged for the output.

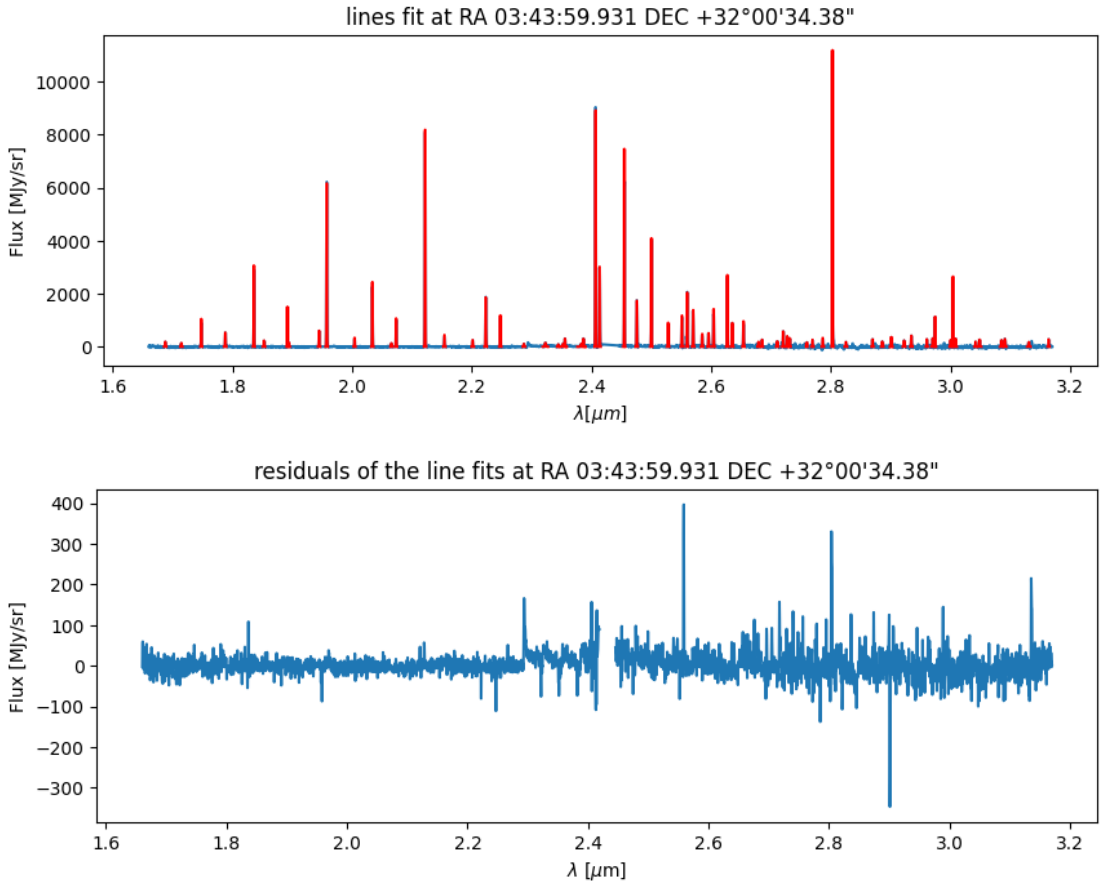


Figure 4.7: The line fits and their residuals done for the shorter wavelength part of the example spectrum. The upper panel shows the fitted line profiles in red plotted over the original data, which is shown in blue. From this plot it is clear, that the fitting routine did not miss any bright lines. By checking the residuals as shown in the lower panel, I can ensure the accuracy of the fits. In this case the residuals are remarkably low, with only a few small peaks that *might* be dim lines.

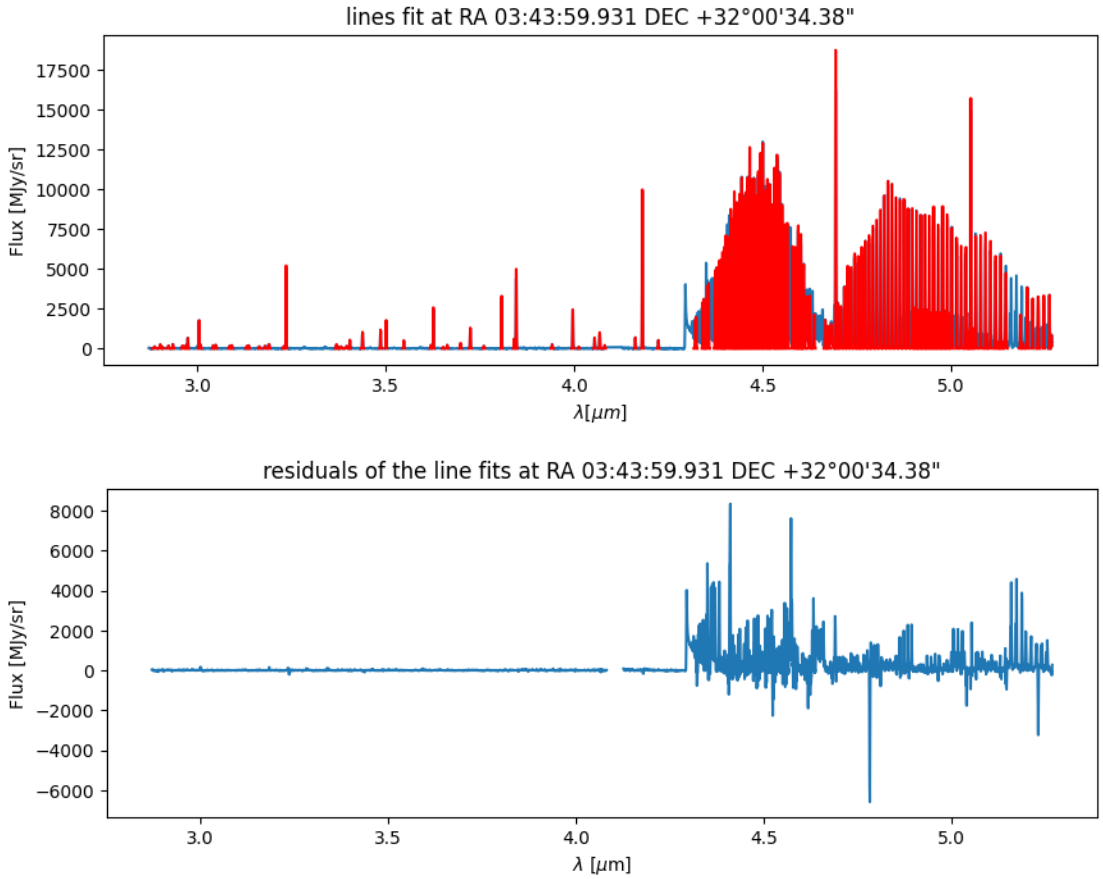


Figure 4.8: The line fits (top) and their residuals (bottom) done for the longer wavelength part of the example spectrum. Again the fitted models are shown in red while the original data is blue. The method used is not suitable for fitting the dense CO branches seen at wavelengths higher that ~ 4.25 μm. As every fit is done in isolation, partially blended lines can not be extracted properly. This can be seen by the missed lines in the upper plot and the higher residuals in the lower plot.

My initial testing has shown that across numerous spectra only a handful of additional lines could be fitted in the second step. Given the high computational effort required for minimal improvement, I carry out the line fitting for the entire data cube in a single step.

Even more adaptations can be made using a number of additional parameters. Uncertainties of each individual flux measurement of the spectrum can be passed to the routine to be considered in the fitting process. The same is true for information on the physical wavelengths associated with the channels. In this case all results are expressed in terms of the wavelengths instead of channels. Setting the mode to "em", "abs" or "all" will result in only emission lines, only absorption lines or all lines to be fitted. If required, the routine will also return the residuals and *lmfit*-models with the other results. Even if no line fit is successful, the routine will return — empty — output in the expected format. Besides visualizing the final results, the routine can also output a visualization of every individual line fit done.

4.3 Line matching

Once all of the lines in the spectrum are fitted, I match them to their literature counterparts using my *nirspecpy.ut.line_matcher* utility function. Given fit results, central wavelengths from literature and a user-defined maximum allowed deviation, it returns the information of all fits, whose central wavelength is sufficiently close to the literature value. For both the fit results and the resting wavelengths, a single object or a list of multiple objects can be given.

Determining a good limit for the maximum allowed deviation between resting wavelength and observed wavelength is mostly a question of the Doppler shift due to radial motion. For high radial velocities the Doppler shift can be substantial and has to be considered separately. In the case of NIRSpec observations of HH211, though, this is not the case. Caratti o Garatti et al. (2024) measure the radial velocities in the bow shock region of HH211 using H_2 as a tracer. They report radial velocities in the system's rest frame that range from -8 km/s to -17 km/s. Additional contributions to the observed radial velocity come from the systemic velocity with respect to the local standard of rest of 9.2 km/s and the conversion from the heliocentric reference frame — as used by default by JWST — to the local standard of rest. In the case of HH211 this conversion is adding 6.7 km/s (Caratti o Garatti et al., 2024). So the expected observed radial velocity in the bow shocks of HH211 is ~ -1 to 8 km/s. This is quite small compared to nominal velocity resolution of the observations analyzed in this thesis. As discussed in section 2.2, these observations have channel widths of 0.000395 and 0.000665 μm and cover wavelength ranges of 1.70 to 3.15 μm and 2.88 to 5.20 μm. Based on these parameters, one spectral channel corresponds to a velocity range of 37–72 km/s, depending on the wavelength.

Given that the expected radial velocities are so much smaller than the channel widths, it is a safe assumption, that lines are shifted by less than one channel. Thus, allowing a maximum deviation of one channel width leads to good results.

Setting the *mode* keyword of the line matching routine controls the output structure and allows for easier data access for the further analysis. In *line mode* an output is returned for each line in the list of resting-frame central wavelengths. Each one of those outputs contains lists of the x and y coordinates as well as an individual list for each fit parameter. This structure is ideal for comparing the same line across the different pixels. It is particularly practical for plotting and visualizing results as these lists can directly be used as the input for plotting routines.
In *pixel mode* the first hierarchy is that of the pixels, meaning an output is returned for each pixel. Within them are the matched fit parameters and pixel coordinates. Additionally, a list of indices relating said parameters directly to the elements of the list of literature values is returned. When carrying out more detailed analysis using multiple lines per pixel this structure is more suitable. I added it for the excitation diagrams and it has proven very helpful for that task.

At this point I identified some lines as problematic and dropped them from the further analysis. Dropped lines include any pair of lines whose central wavelengths are closer than twice the channel width of the observation and lines that fall into the dense CO-branches with wavelengths higher than 4.25 μm. In both cases it is likely that the observed emission has contributions of multiple lines, which leads to inaccurate results in later steps. If lines are further separated there should be at least one channel with lower emission between their peaks and they should be separable.
An exception to this are the hydrogen (0-0) S(8), (0-0) S(9) and (0-0) S(10) lines. They remained in the data, based on them being multitudes brighter than their neighboring CO lines, which makes the effect of a possible contamination less significant.
Another source for potentially problematic lines is the ice extinction feature around 3 μm. It affects lines with wavelengths between 2.75 and 3.75 μm (Decleir et al., 2022). Due to the higher extinction at these wavelengths, the observed line intensities are much lower than their intrinsic intensities. The lines are affected to varying degrees, which distorts their relative intensities and causes problems in the excitation analysis. As will be discussed in chapter 6 this feature is not easily modeled making its correction difficult. For the sake of comparison, I carry out parts of the further analysis twice: once excluding the lines affected by this extinction feature and once including them.

5 Excitation Analysis

By using excitation diagrams — sometimes also called Boltzmann or rotational diagrams — one can determine the total column density as well as the excitation temperature. The key concept is investigating the relative population of the different energy levels of an ensemble, which can be done by relating observed column densities to their corresponding energy levels. If the system is in local thermodynamic equilibrium (LTE), this relation is given by a single Boltzmann distribution, that can be fitted to the data points and reveals the system's total column density as well as its excitation temperature. The methodology described and implemented here is based on the approach detailed in Goldsmith and Langer (1999) with adaptations from Van De Putte et al. (2024). Considering the ortho and para states separately gives insights into the OPRs (Neufeld et al., 1998).

5.1 Theoretical foundation of excitation diagrams

For a system in LTE the distribution of energy levels follows a Boltzmann distribution. In that case, the probability of finding any given particle in the energy level u is proportional to $g_u \cdot \exp(-E_u/(k_B T_{ex}))$. Here E_u denotes the energy of that level, g_u its statistical weight, k_B the Boltzmann constant and T_{ex} the system's excitation temperature. When considering the entire system, the column density N_u of particles in a specific energy level u is given by equation 5.1. In this equation, N_0 is the total column density and Z the partition function. Z is the sum over the probabilities of all possible upper levels and is needed to normalize the total probability to 1. As can be seen in equation 5.2 this function depends on T_{ex} (Van De Putte et al., 2024).

As discussed in section 1.4.2, shocks are inherently out of equilibrium. Assuming a Boltzmann distribution for the energy levels in this analysis is, thus, an approximation and the expectation is, that no single excitation temperature fully describes the system. Nevertheless, Boltzmann distributions can be fitted to the observed column densities and provide insights into the excitation conditions.

$$N_u = \frac{N_0}{Z} \cdot g_u \cdot e^{-\frac{E_u}{k_B T_{ex}}} \tag{5.1}$$

$$Z = \sum_u g_u \cdot e^{-\frac{E_u}{k_B T_{ex}}} \tag{5.2}$$

To plot excitation diagrams and determine the properties of the underlying Boltzmann distribution, the column densities of a number of energy levels have to be determined first. This can be done via the observation of line transitions — with known intrinsic properties. Following Van De Putte et al. (2024) and also assuming optically thin emission, so that optical depth terms vanish, the column densities N_u — of the upper level of each transition — can be calculated from the observed integrated lines intensities I via equation 5.3. Here h is Planck's constant, c is the speed of light, λ is the transition wavelength and A_{ul} is the Einstein A coefficient for the transition from the upper level u to the lower level l. In the following calculations all of the intrinsic transition properties are taken from the catalog Roueff et al. (2019a).

$$N_u = \frac{4\pi}{hc} \cdot \frac{\lambda I}{A_{ul}} \tag{5.3}$$

For practical reasons, a number of further manipulations are done to the equations just introduced. Given that g_u is a known quantity of each transition's upper energy level, it is more convenient to consider the distribution of the individual states by dividing both equation 5.1 and equation 5.3 by g_u. For ease of fitting and analysis, one can further use $\log(N_u/g_u)$ to linearize the fitting task. Additionally, for the assessment of the excitation temperature T_{ex} it is preferable to express the upper state energy in terms of a temperature as well. This is done by dividing it by Boltzmann's constant yielding $T_u = E_u/k_B$. So truly the relation examined is that of T_u to $\log(N_u/g_u)$.

Figure 5.1 shows a visualization of the T_u to $\log(N_u/g_u)$ relation and the corresponding linear fit. This fit function has a slope of $-T_{ex}^{-1} \cdot \log(e)$ and an intercept of $\log(N_0/Z)$, from which T_{ex} and N_0 can be calculated by simple algebra (Van De Putte et al., 2024). Here, the fit only uses the lines of the H_2 (1-0) Q branch of the example spectrum, a choice made to illustrate the technique. If all detected lines are considered by the fit, the situation becomes more complex. This will be discussed in section 5.3.

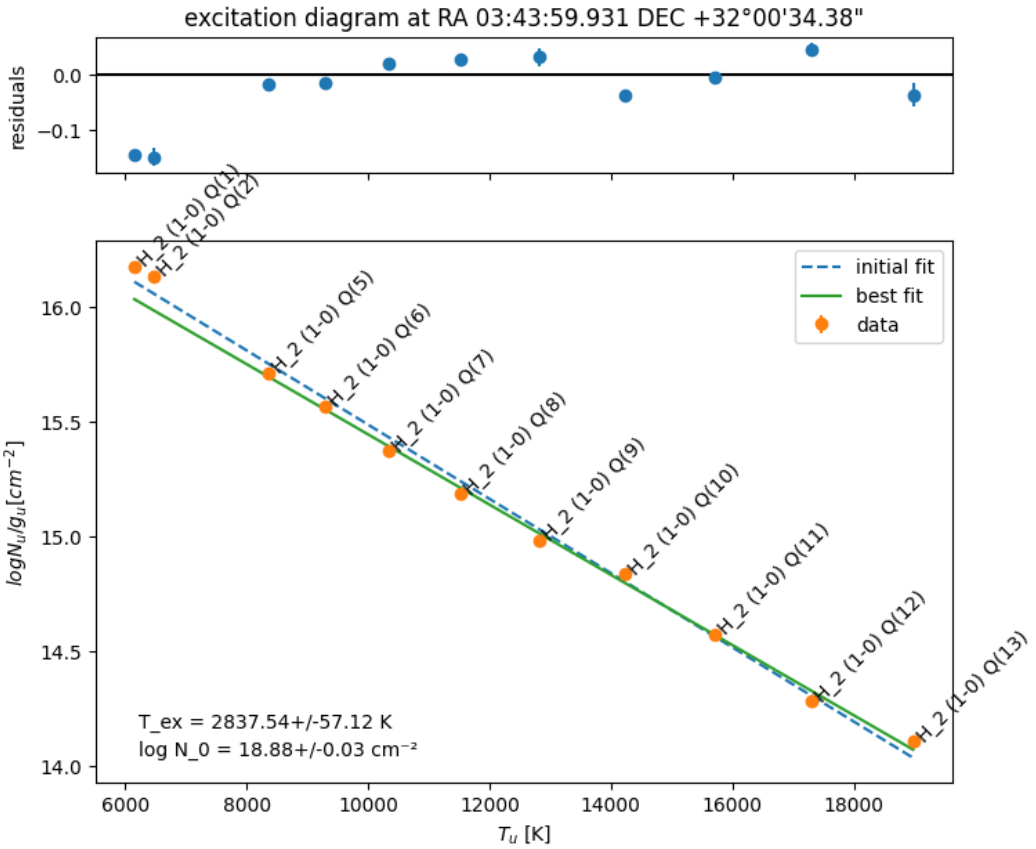


Figure 5.1: An excitation diagram done for the (1-0) Q branch of the H_2 emission in the example pixel (corresponding to the spectrum seen in 2.3). Each data point corresponds to one observed line, their names are annotated. This excitation diagram illustrates the technique; as will be discussed in section 5.3 the analysis becomes more complicated if all detected lines are considered.

5.2 Ortho-to-para ratios

Using the fact that ortho-H_2 only occupies odd rotational states and para-H_2 only even rotational states, the excitation analysis described in section 5.1 can be done separately for ortho-hydrogen and para-hydrogen. In this case, the calculations for each spin isomer use a separate partition function, that only includes states occupied by that isomer. The resulting $N_{0,i}$, thus, describe the total column density of that spin configuration. Dividing the value obtained for the ortho transitions by the value calculated for the para transitions directly gives the OPR.

Ideally, the excitation conditions of the two isomers differ only in their column densities and share the same T_{ex}, as differences in T_{ex} would also influence the derived column densities. A way ensuring the same T_{ex} is keeping it fixed — at the value determined using lines from both isomers — and only varying N_0 during the fitting process.

5.3 Practical application of excitation diagrams

The practical implementation of the method described in section 5.1 is my *nirspecpy.calc.fit_T_ex_N_tot* routine. When given a set of observed line properties as well as theoretical transition properties, it fits a Boltzmann distribution to the data and returns the excitation temperatures, total column densities and the number of lines used for each fit.

The routine includes a number of optional settings such as defining a required minimum number of observed lines. If fewer are found, no fitting is attempted — I used a minimum number of 5. Additionally, a maximum allowed relative flux error can be set, excluding all lines with too high errors from the fit. A functionality for easy visualization of each fit is also included. The function is written to work with individual spaxels as well as the spatial grid. I tested this for a handful of spaxels before up-scaling to the entire data set.

As hinted before, the analysis becomes more complicated if all — or a larger number of — the detected lines are included in the fitting process. I show the excitation diagram for the full spectrum for the example pixel in Figure 5.2. A curvature can be seen in the distribution of the data points. This curvature leads to a systematic underestimation of the column densities of states near the low and high energy limits of the observation and overestimation of the states with intermediate energies. Even though the general trend is the same as in Figure 5.1, the derived excitation properties are significantly different ($T_ex = 2838 \pm 57$K and $\log(N_0) = 18.88 \pm 0.03$ cm^{-2} compared to $T_ex = 3169 \pm 69$K and $\log(N_0) = 18.74 \pm 0.03$ cm^{-2}). The observed discrepancy is a consequence of the non-equilibrium nature of shocks, that was discussed in section 1.4.2.

While the underlying excitation temperatures form a continuous range, I take the varying T_{ex} into account by fitting the excitation diagrams in multiple components. This is implemented into my function via the *comps* keyword. If a value greater than 1 is passed for *comps*, the data is sorted by corresponding T_u and then cut into parts that are fitted separately as described in section 5.1. Following the example of Caratti o Garatti et al. (2024), I adopt two separate components and refer to them as the warm and the hot component.

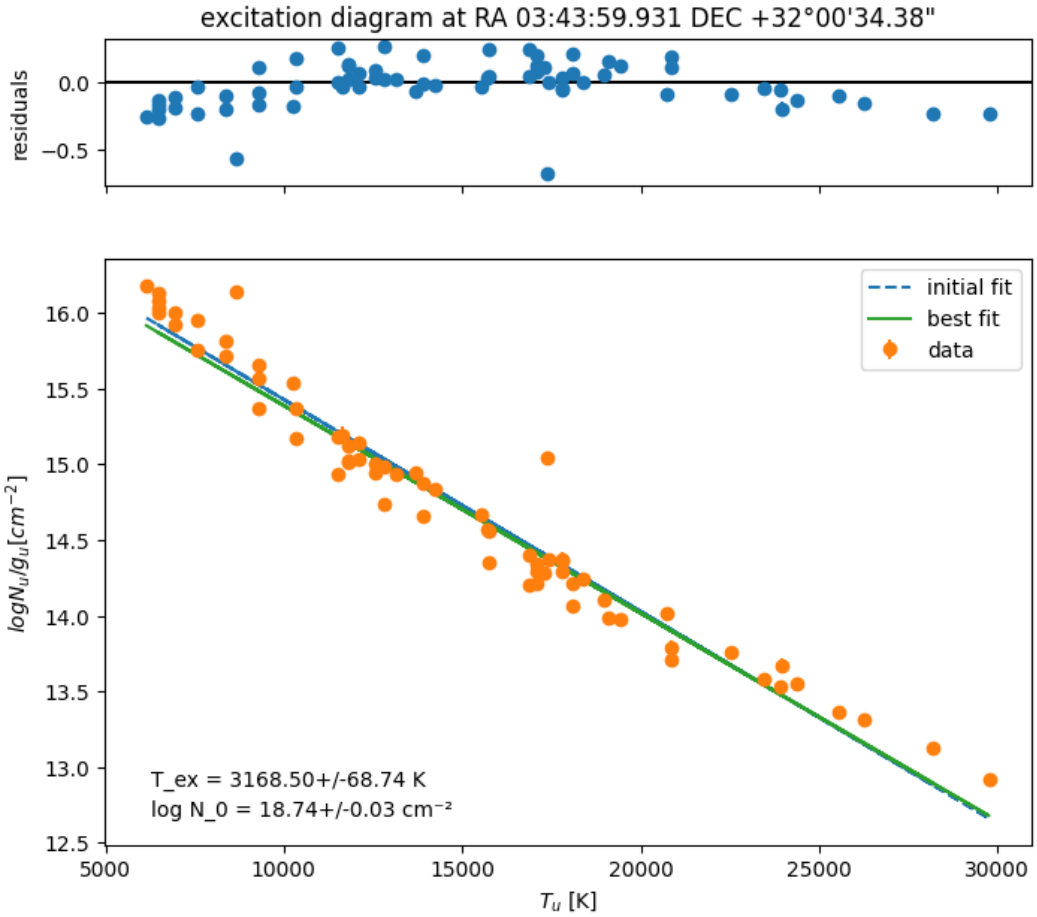


Figure 5.2: An excitation diagram done using all H_2 lines detected in the example pixel (corresponding to the spectrum seen in 2.3). For clarity of the plot the annotations were left out here. This plot illustrates two key aspects. First, the data deviates systematically from the fitted Boltzmann distribution. The distribution of the data seems to follow a curve rather than straight line, which indicates, that there might be more than a single excitation temperature. Second, there are multiple pairs and triplets of data points sharing the same energy but with different column densities. These are lines originating from the same upper state. As their wavelengths differ, extinction affects them differently leading to an additional scatter — a more thorough discussion of this follows in chapter 6.

Figure 5.3 shows the two component fit to all of the detected lines in the example spectrum. In comparison to the single component fit of Figure 5.2, the residuals in the two component fit are overall lower and the deviations no longer follow a clear trend. This means that no parts of the energy distribution are systematically over- or underestimated, which points to an overall better fit. The inferred T_{ex} and N_0 are significantly different for the two components, which in turn are both different from the one component fit. All of this indicates that the excitation structure is more complex than a single Boltzmann distribution.

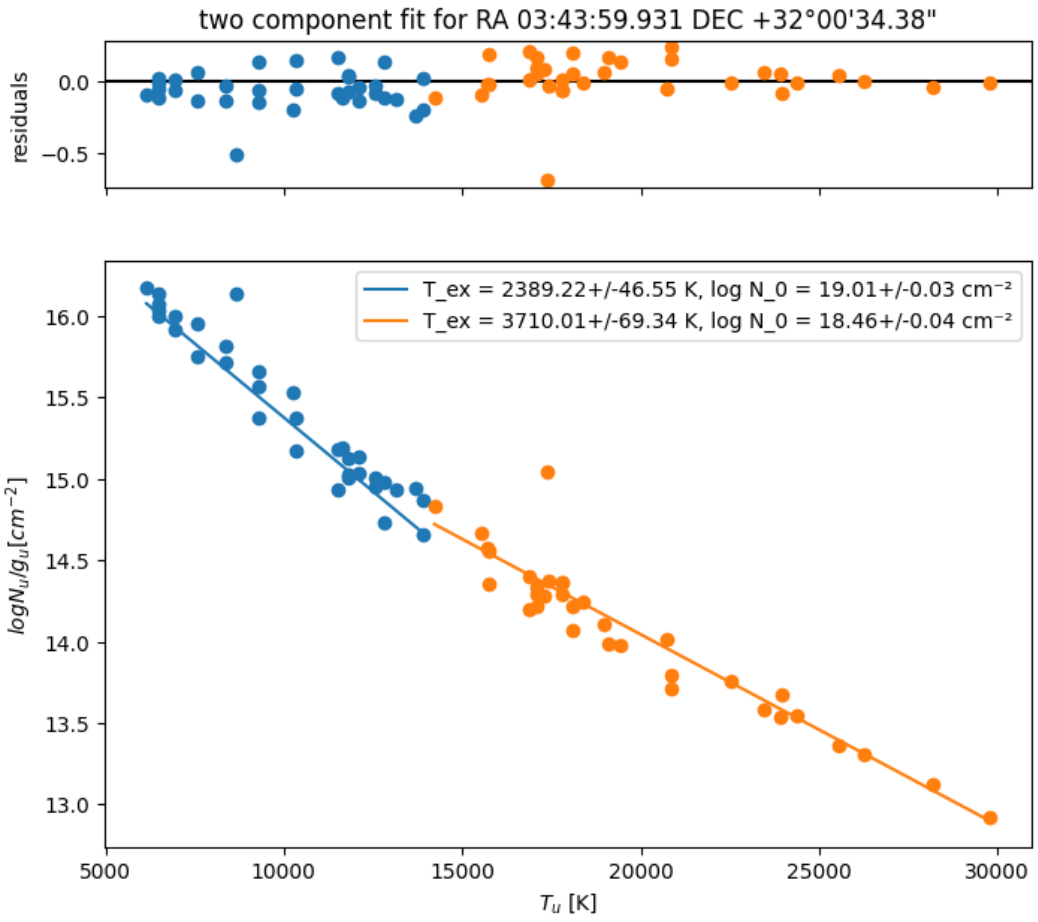


Figure 5.3: Two component fit to the same data as in Figure 5.2. Fitting two separate components helps mitigate the systematic deviations seen in the one component fit. This is most evident by the overall lower residuals and the less systematic deviations of the two component fit. The two components have significantly different T_{ex} and N_0, which further justifies the use of multiple components.

6 Extinction Analysis and Correction

Given the young age of HH211 and the fact that the driving source remains deeply embedded, a significant amount of light attenuation is to be expected. Because of this, the analysis and correction of extinction is of critical importance.

The effects of extinction are visible in the excitation diagram in Figure 5.2 as a vertical scatter of multiple N_u values for the same T_u. These multiple points correspond to multiple lines originating from the same upper level, which should — intrinsically — each recover the same column density. Due to the wavelength dependence of extinction and the different wavelengths of the various lines, they are affected differently. This fact can be used to estimate the extinction.

6.1 Estimation using line ratios

As just mentioned, pairs of different lines originating from the same upper state can be used to determine the extinction. This is done by comparing the intrinsic line ratios to the observed ones and assuming that any deviation is due to extinction. The intrinsic line ratio $(I_1/I_2)_{intr}$ depends only on the transition wavelengths λ_1 and λ_2 of the respective lines and the Einstein A transition coefficients A_1 and A_2 of either transition. The relation is shown in equation 6.1. It is equivalent to using equation 5.3 and setting $N_{u,1} = N_{u,2}$, which is expected as both lines stem from the same state. The theoretical line properties are again taken from Roueff et al. (2019a).

$$\left(\frac{I_1}{I_2}\right)_{intr} = \frac{\lambda_2 A_1}{\lambda_1 A_2} \tag{6.1}$$

The attenuated line ratios can be directly calculated from the observed integrated line intensities and will be denoted as $(I_1/I_2)_{obs}$ in the further equations. Equation 6.2 then gives the difference in extinction $\Delta(A)$ in magnitudes.

$$\Delta(A) = 2.5 \cdot \left[\log\left(\frac{I_1}{I_2}\right)_{intr} - \log\left(\frac{I_1}{I_2}\right)_{obs}\right] \tag{6.2}$$

This differential extinction is related to the total visual extinction A_V to allow the comparison between different pairs of lines or different wavelength regimes. A necessity for that is the knowledge of — or more truthfully an assumption about — the wavelength dependence of the extinction $A(\lambda)$. Using a general reddening law $k(\lambda) = A(\lambda)/A_V$ the visual extinction can be calculated using equation 6.3.

$$A_V = \frac{2.5}{k(\lambda_1) - k(\lambda_2)} \cdot \left[\log\left(\frac{I_1}{I_2}\right)_{intr} - \log\left(\frac{I_1}{I_2}\right)_{obs} \right] \tag{6.3}$$

Across many wavelength regions the reddening law can be modeled decently well by a power law of the form $k(\lambda) = A(\lambda)/A_V = S \cdot \lambda^{-\alpha}$ where α and S are free parameters that depend on the properties of the attenuating dust. Some regions require additional terms to model strong extinction features such as the ice feature around 3 μm (Decleir et al., 2022). I show a comparison of 3 reddening laws presented in Decleir et al. (2022) in Figure 6.1. Two of those reddening laws were observed along dense sightlines and show the ice feature clearly. Based on visual comparison of these two to the third one — the Milky way average power-law fit — I determine the "likely affected" wavelength range to be 2.75 to 3.75 μm. The most prominent H_2 lines in this range are the (1-0) O(3), O(4) and O(5) lines and the (0-0) S(15) line. Multiple lines are blended and are not considered in the analysis regardless.

Once the visual extinction is known, the specific extinction for each wavelength can be calculated and the observed fluxes and intensities corrected.

6.2 Practical implementation

While the calculation of the extinction is fairly straightforward using the formulas described in section 6.1, they rely on a number of assumptions and choices. The two big ones are the reddening law and the pair of lines used.

In theory, any pair of lines from the same upper state can be used for extinction correction. Ideally, though, both lines are prominent and somewhat isolated, so that their intensities can be accurately determined. Using the 68 lines detected in the example pixel, one can form 24 different pairs of lines, that share their upper state. Of those 24 pairs 12 stem from bright (1-0) lines. For those pairs I calculate the visual extinction using equation 6.3 and the Milky Way average curve from Decleir et al. (2022) — a power-law of the shape $k(\lambda) = S \cdot \lambda^{-\alpha}$ with parameters $S = 0.386$ and $\alpha = 1.71$. Then I correct the line intensities for extinction, convert them to column densities and plot excitation diagrams to check the effects of the different A_V values.

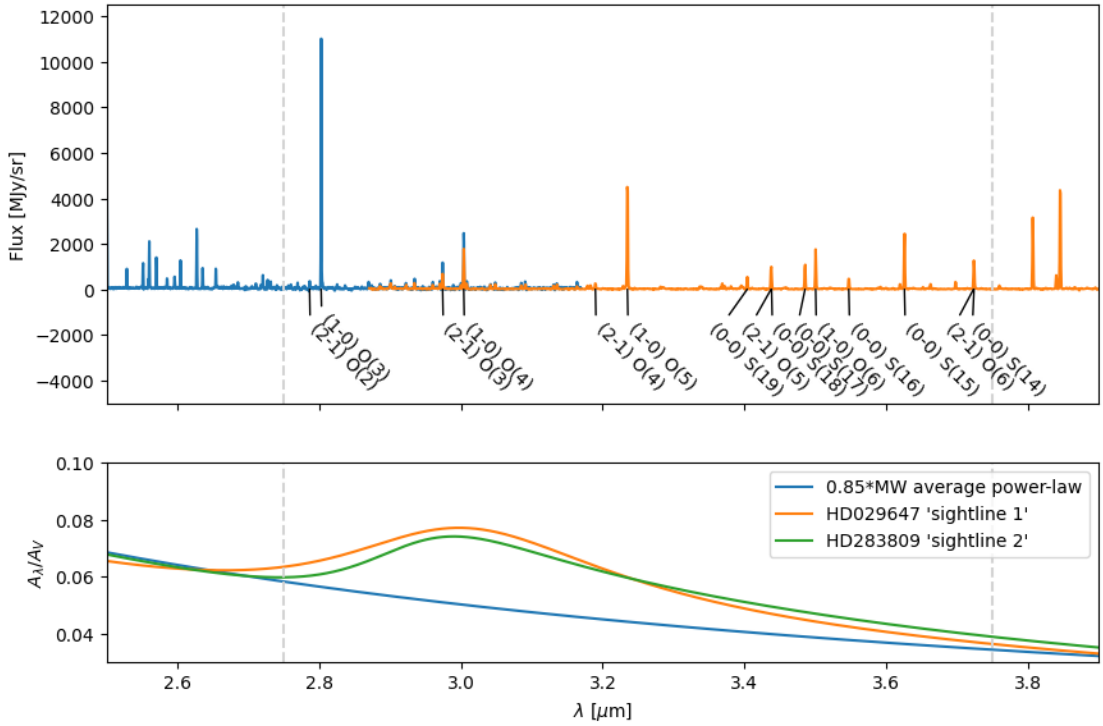


Figure 6.1: Visualization of the 3 μm ice feature. The upper plot shows the spectrum in the example pixel with prominent H_2 lines, that are likely affected by the ice feature, labeled with their name. In the lower panel three reddening laws presented in Decleir et al. (2022) are compared. The orange and green lines are the reddening laws found in the sightlines to HD029647 and HD283809 — in the following they will be called sightline 1 and sightline 2. The blue line is the Milky Way average power-law fit, which was re-scaled by a factor of 0.85 to match the other two reddening laws. The gray dashed lines mark the wavelength range of 2.75 to 3.75 μm, that I consider "likely affected".

For any pair including the (1-0) O(4) line the A_V values calculated are negative. This is a clear pointer that a power law is not suitable for describing the extinction at that line's wavelength and that the 3 μm ice feature plays a relevant role here.
All pairs consisting of one (1-0) S line and the corresponding (1-0) Q line, lead to similar A_V values, which are around 9–10 mag.
Finally, pairs including one of the other (1-0) O lines and a corresponding (1-0) S or (1-0) Q line result in lower but positive A_V.

I show a comparison of the uncorrected excitation diagram against a corrected one in Figure 6.2. In that plot an extinction correction for $A_V = 9.02 \pm 0.18$ is applied, which is the value determined for the (1-0) S(5) and the (1-0) Q(7) pair. Given the similar values determined for (1-0) S and (1-0) Q line pairs, this correction also works well for the other such pairs. This is seen by the multiple lines coinciding in the corrected excitation diagram at the bottom of Figure 6.2. In general, the vertical scatter is lower in this excitation diagram, pointing to the fact that a lot of it was caused by extinction and could successfully be corrected for.

The entire calculation of the extinction is also dependent on the reddening law used. In this work I adopt mainly the Milky Way average curve from Decleir et al. (2022) for extinction correction. For the purpose of comparison, I also calculate visual extinctions and extinction-corrected excitation diagrams using the curves reported in Decleir et al. (2022) for the two dense sightlines of their sample. These curves include a modified Drude profile to better match the ice feature around 3 μm and are visualized in Figure 6.1.
In Figure 6.3, I compare the effects of the two different reddening laws — a power-law and one including a modified Drude profile. When focusing on the highlighted lines with wavelengths between 2.75 and 3.75 μm, it is clear, that including the ice feature in the reddening law improves the correction for affected lines. Upon closer inspection, though, there is still an offset visible in the low energy regime. This is unsurprising, considering that the reddening law has not been adapted besides re-scaling to match the specific extinction present.

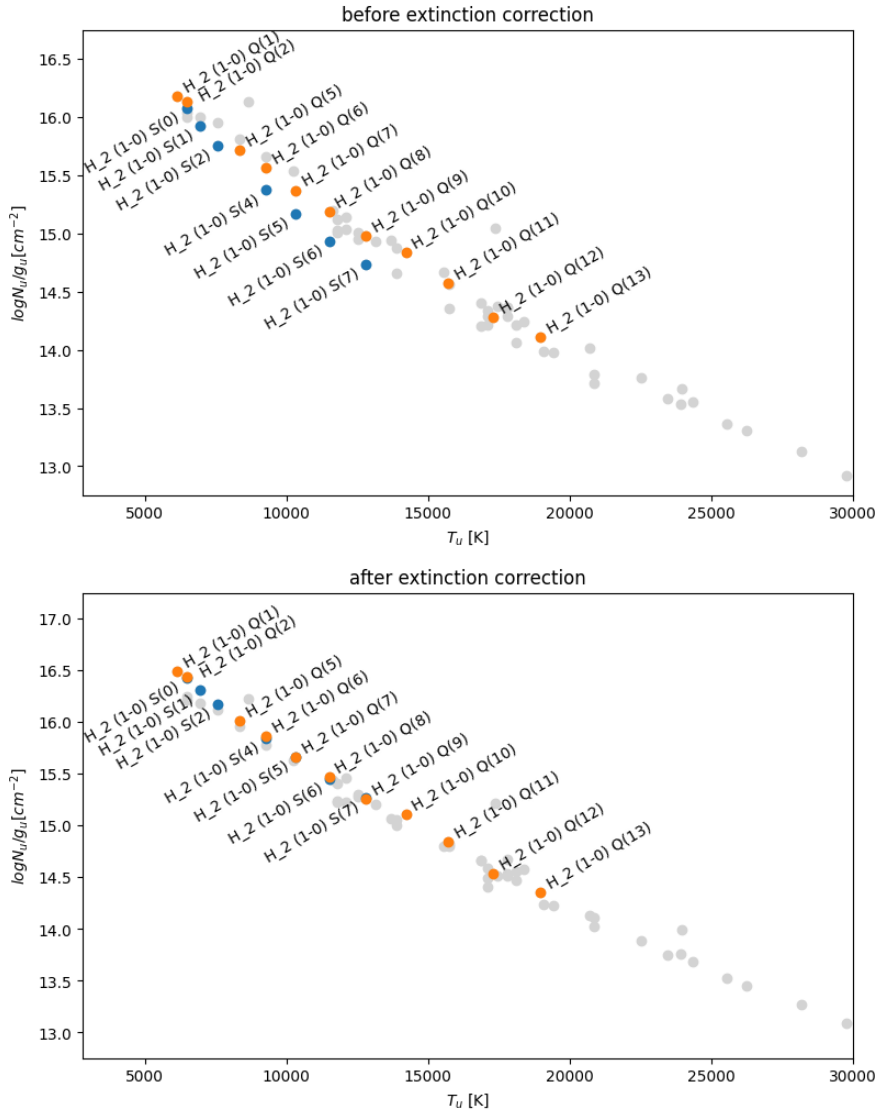


Figure 6.2: The uncorrected excitation diagram (top) compared against the extinction-corrected one (bottom), with the (1-0) S and (1-0) Q lines marked. The extinction correction was done using the Decleir et al. (2022) Milky Way average reddening law with a visual extinction $A_V = 9.02 \pm 0.18$ mag as calculated from the (1-0) S(5) and the (1-0) Q(7) pair. By definition, the column densities of these two lines are identical after the extinction correction. For the other (1-0) S and (1-0) Q lines the correction is also rather accurate as highlighted by the nearly invisible blue (1-0) S data points in the lower plot.

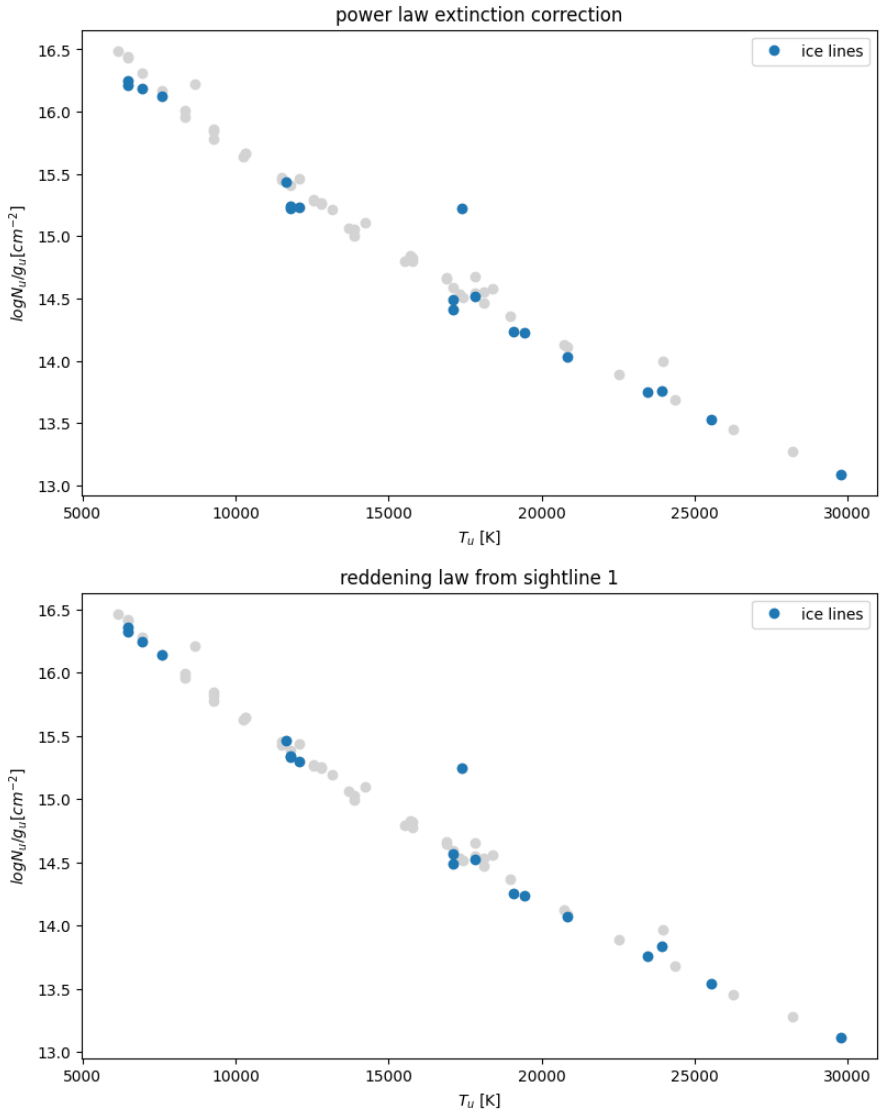


Figure 6.3: Excitation diagrams corrected using the Milky Way average power-law (top) and a power-law plus modified Drude profile (bottom). In both cases the reference lines are — again — the (1-0) S(5) and the (1-0) Q(7) lines. The corresponding A_V are 9.02 ± 0.18 and 10.34 ± 0.24 mag. I showcase the lines in the wavelength range of 2.75 to 3.75 μm, that are likely affected by the 3 μm ice feature. The data points corresponding to these lines appear below the general trend in the power-law corrected diagram (top). In the lower plot they appear much closer to the bulk of the data points, so including that modified Drude profile helps to model the ice feature.

7 Data Cube Analysis

In the preceding three chapters, I detailed the methodology and specific implementation used and illustrated the entire process through the example of one single spaxel. The full NIRSpec data set is a cube of over 10000 spaxels, of which over 1500 show significant emission, that is to be analyzed. This makes a more automated approach and multiprocessing necessary — I used the *multiprocess.pool* method as it works with externally defined target functions. To further facilitate the process I implemented two additional utility functions — *nirspecpy.ut.fits_utility* and *nirspecpy.ut.make_wrappper*. The full data set analysis proceeds as follows.

The first step — extraction of line properties — is done for each observation separately. First, I load the entire data set using the *fits_utility* function. It takes a path pointing to the pre-processed observation as input and returns the data and the associated errors, the coordinate system and the wavelength axis constructed from the header information, and — optionally — the header itself. For practical reasons, I convert all fluxes from MJy/sr to $\mathrm{W/m^2/\mu m/sr}$ right after loading the data. The flux masking value of 250 MJy/sr introduced in section 4.1.1 for the baseline correction step cannot be directly converted to $\mathrm{W/m^2/\mu m/sr}$. This is because the conversion depends on the wavelength, which leads to different values for different parts of the spectrum. I use the upper limit for each observation to ensure that no noise is cut out of the baseline estimation. The resulting flux masking values are $\sim 2.7 \cdot 10^{-4}$ $\mathrm{W/m^2/\mu m/sr}$ for the shorter wavelength observation and $\sim 9.1 \cdot 10^{-5}$ $\mathrm{W/m^2/\mu m/sr}$ for the longer wavelength observation.

Then I use the *make_wrappper* utility to define a wrapper function around the analyzing functions for baseline correction and line fitting. This utility's inputs include the data cube, coordinates and wavelength information, required signal-to-noise ratio as well as optional keyword arguments that are passed to the analyzing functions. I use the returned wrapper function as the pool's target function and a simple array of all the spaxel indices as the input values. This results in a long list of fit results and coordinates, that I save for the later analysis.
Matching the lines to their literature counterparts is straightforward, as the matching routine supports input consisting of fit information of multiple pixels. By setting the *mode* keyword this routine returns its output with regard to each input pixel so that the data is easily accessible for the future steps. After matching the lines, I combine the data from the two observations into a single data set, which is used in the further analysis.

Applying extinction corrections proves more tricky, as only few lines are detected across a large number of pixels. A large part of the (1-0) Q branch is missing in patches as the transitions fall into the wavelength gap of the short wavelength observations and a number of higher excitation lines are only detected in the very bright regions. Ultimately, I used the (1-0) S(5) and (1-0) Q(7) pair and the Milky Way average reddening law for extinction corrections, as those appeared to be the most plausible after some preliminary testing. These tests included extinction maps — criteria are spatial correlation and plausibility of overall values — as well as studying the effect of the applied correction on the excitation diagrams — as discussed in section 6.2. An additional point in favor of this pair is that both lines' wavelengths are at a respectable distance from the ice feature around 3 μm.
Additionally, I carried out extinction analysis for the (1-0) S(1) and the (1-0) O(5) pair for comparison with Caratti o Garatti et al. (2024). Furthermore, I calculate and visualize extinction maps for the other two reddening laws introduced in section 6.2.

Using the extinction-corrected line intensities, I carry out excitation analysis for each pixel. This is again a rather simple task as the excitation diagram fitting routine *nirspecpy.calc.fit_T_ex_N_tot* works with multi-pixel input. Each excitation diagram is fitted using two components. Based on the discussion in section 6.2, I carry out this step twice: once for all the lines that are left after excluding blended lines and once excluding the lines falling into the 3 μm ice feature as well.
Additionally, for the study of ortho-to-para ratios I also do the same excitation analysis for the ortho lines and the para lines separately. Again, each excitation diagram fit uses two components. These separate fits might have different excitation temperatures which would affect the derived OPRs. Because of this, I carry out this analysis twice: once varying both the slope and the intercept during fitting and once with a fixed slope, which corresponds to a fixed T_{ex}. In the case of the fixed excitation temperatures, I use the value derived for the full ensemble of lines.
Finally, I visualize all the results in the form of maps showing spatial distribution, as well as some histograms to facilitate comparison between the two components.

8 Results and Discussion

In this chapter I report and discuss the key results of my master thesis. Here I mainly focus on the excitation temperature, column density and ortho-to-para ratio maps as well as the extinction maps. The excellent data quality together with the high resolution and fine-tuned analysis routines resulted in a very detailed view on the excitation and extinction conditions around HH211's terminal bow shocks. Overall, the results discussed here are consistent with the prior results and expectations introduced in section 1. There are, however, some deviations that I explore in more detail below.

8.1 Excitation conditions

From the flux maps in Figure 2.2 it is clear, that there are three relatively small regions of strong H_2 and CO emission surrounded by regions of weaker H_2 emission. The emission from the H_2 (1-0) O(5) and the H_2 (0-0) S(9) lines have very similar spatial distributions. Given their similar upper level energies of $\sim$ 10340 K and $\sim$ 10260 K, it is not surprising to find both states excited in the same regions (Roueff et al., 2019a). The spectral channel at 4.520 μm, which was chosen for the CO emission, roughly coincides with the $^{12}C^{16}O$ (1-0) R(19) transition. This transition has an upper level energy of $\sim$ 4020 K (Goorvitch, 1994). In the plot shown, the CO emission appears compact and only in the region of the bow shocks. Based on it corresponding to a lower energy state one would expect more extended emission. Upon closer inspection, there is more extended CO emission. It is best visible in the wake of the middle bow shock. This extended emission appears dim in comparison to the H_2 emission, which can be explained by the lower overall abundance of CO and the higher number of other CO levels with similar energies.

Focusing now on the excitation of H_2 exclusively, I have shown in Figures 5.2 and 5.3 that fitting the excitation diagrams with two separate components leads to better fits overall. The two component fits have lower residuals with less systematic deviations compared to the single component fits. This is a result of the non-equilibrium nature of shocks and was discussed in section 1.4.2. From the shock simulations introduced in that section, one would expect to derive higher excitation temperatures and lower column densities when using lines corresponding to higher energy states. The excitation diagram for the example spectrum of the pixel centered at RA 03:43:59.931 DEC +32°00'34.38" directly follows this expectation. The warm component (defined by the lines of lower upper level energies) has $T_{ex} \sim 2389 \pm 47$ K and $\log(N_0) \sim 19.01 \pm 0.03$ cm^{-2}.

In contrast to this, I derived $T_{ex} \sim 3710 \pm 70$ K and $\log(N_0) \sim 18.46 \pm 0.04$ cm^{-2} for the hot component. While the line intensities used for these diagrams were not yet corrected for extinction, the derived values are within the typical range that I find for the highly excited region around the bow shocks.

As mentioned in chapter 7, I carried out the excitation analysis of the entire data set twice: once including the lines affected by the ice feature and once excluding them. The results of these two variations are quite similar, with the more restrictive choice of lines showing slightly smaller uncertainties and lower spatial noise. In the following I will focus on the latter results.

Using the extinction-corrected line intensities and the more restrictive selection of lines I find the following typical excitation conditions.
With logarithmic values mostly within $\log(N_0) \approx 18.25$–19.25 cm^{-2} the column densities of the warm component are overall higher than those of the hot component, whose logarithmic column densities are in the range $\log(N_0) \approx 17.75$–18.50 cm^{-2}. For a majority of fits the logarithmic errors are below 0.1 cm^{-2} with the ones for the warm component being overall higher.
The excitation temperatures of the warm component range from 1500 up to 2500 K. For the hot component they are approximately between 2000 and 3500 K. There is, however, a number of pixels with excitation temperatures in the 3500 to 4000 K range. The errors of the excitation temperatures of both components are mostly below 200 K. A number of fits for the warm component show higher errors.

As Figure 8.1 illustrates, there are clear trends for the spatial distribution of total column densities. For both components, the highest column densities are found in a compact region around the bow shocks. The material in the wake of the shocks shows lower column densities and similar values over an extended region. The spatial variation of the excitation temperatures follows a similar trend and can be seen in Figure 8.2. One key difference is the relatively large area of high excitation temperatures around the bow shocks seen in the hot component. Here no heightened column densities are seen.

As shocks both heat and compress the material, it is unsurprising that the column density and the excitation temperature follow the same general spatial trends. A closer inspection, however, reveals that these trends are not exactly the same and that the shocks are not fully symmetric either. For the furthest bow shock (leftmost in all of the maps), both components show their maximum column density near the north edge of the shock structure. The maximum in excitation temperature, though, is located near the apex of the shock and appears shifted towards the south. The space between the closest shock and middle shock contains a region, in which the excitation temperatures of the hot component reach values higher than those found near the middle shock front – seen in the

bottom panel of Figure 8.2. Finally, the area around the closest (rightmost) bow shock exhibits a number of unexpected trends. In its wake lies a region with very low column densities and some of the highest excitation temperatures of the entire field. Ahead of the same shock front is a region with exactly the opposite trend: very low excitation temperatures and some of the highest column densities. While the errors in those last two regions are overall higher than the average ($\Delta T_{ex} < 400$ K and $\Delta \log(N_0) < 0.3$ cm^{-2}), these deviations are still outside of the uncertainty.

All of the different trends observed point to a more complex picture. Spatial differences in the surrounding medium or some inherent fluctuation within the shocks could be a cause for this. Alternative explanations could also be possible. For example, O'Connell et al. (2005) carry out detailed simulations of the shocks in the northwest outflow lobe of HH211. They explain observed asymmetries in the bow shocks with a mis-alignment between the outflow axis and the direction of the (large-scale) magnetic field.

Given the similarity of the methodology of the two studies, it is rather fitting to compare the results obtained in this thesis with those of Caratti o Garatti et al. (2024) — particularly to the column density and excitation temperature maps in their Figure 9 and the reported ranges of values. As also introduced in section 1.5, these ranges were $T_{ex} \approx$ 900–1000 K and N_0 of some 10^{19} cm^{-2} for the warm component and $T_{ex} \approx$ 2000–3500 K and $N_0 \approx$ some 10^{18}–10^{19} cm^{-2}.

For the hot component the properties derived here and in Caratti o Garatti et al. (2024) are overall similar. However, while the bulk of excitation temperatures is similar (in the 2000–3500 K range), there is a significant number of pixels with higher T_{ex} in this thesis. The column densities are overall lower (ranging from $\sim 6 \cdot 10^{17}$ to $\sim 3 \cdot 10^{18}$ cm^{-2}). With the low uncertainties of my derivation ($\Delta T_{ex} \lesssim 200$ K and $\Delta \log(N_0) \lesssim 0.1$ cm^{-2}) this is a significant difference.
Comparing the properties of the warm components, there is a much stronger discrepancy between my values and those of Caratti o Garatti et al. (2024). The warm component defined in my work reaches T_{ex} of 1500 to 2500 K and N_0 of $\sim 2 \cdot 10^{18}$ to $\sim 2 \cdot 10^{19}$ cm^{-2}. This is also a significant difference and follows the same trend of higher excitation temperatures and lower column densities as the hot component.

These trends can, however, be well explained using the characteristics of shock excited H_2 introduced in section 1.4.2. The derived excitation temperatures and column densities depend on the upper level energies of the lines used as tracers. As seen in excitation diagrams like Figure 5.2, the data points do not fall onto two distinct straight lines. Rather, the entire distribution shows a curvature.

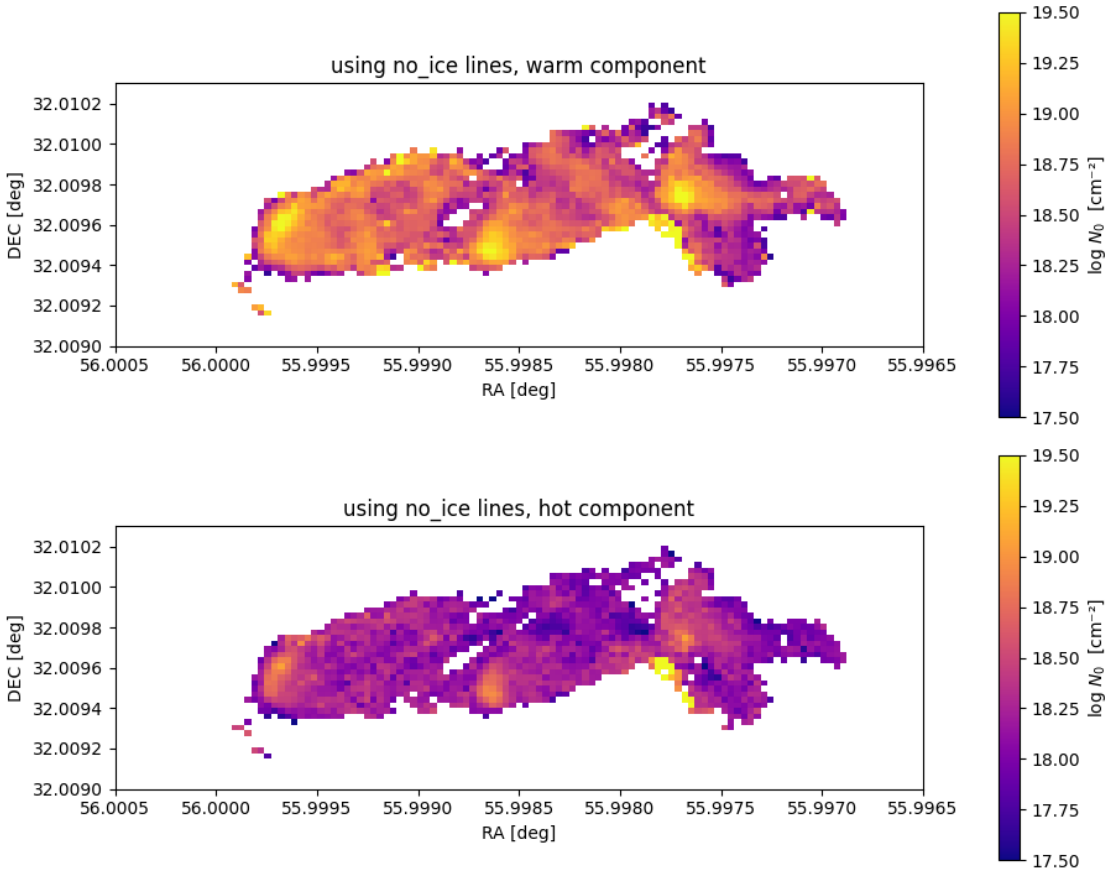


Figure 8.1: The total column densities of the warm (top plot) and the hot (bottom plot) component. The column densities for the warm component are higher, but the distribution is quite similar with both components having their maxima in the area of the bow shocks.

So, even though they share the same names, the warm and the hot component in this work are not exact matches for the warm and the hot component in Caratti o Garatti et al. (2024). Here I used near infrared observations, which include lines originating from energy levels from 6000 to 28000 K. The mid infrared emission studied in Caratti o Garatti et al. (2024), on the other hand, covers emission from states with energies ranging from 1000 to 16000 K. This means that neither study captures the full picture of the H_2 excitation. My analysis does not include the lines from lower energy states, which are associated with lower excitation temperatures and higher column densities. Therefore, I derive higher T_{ex} and lower N_0 for each ensemble of lines.

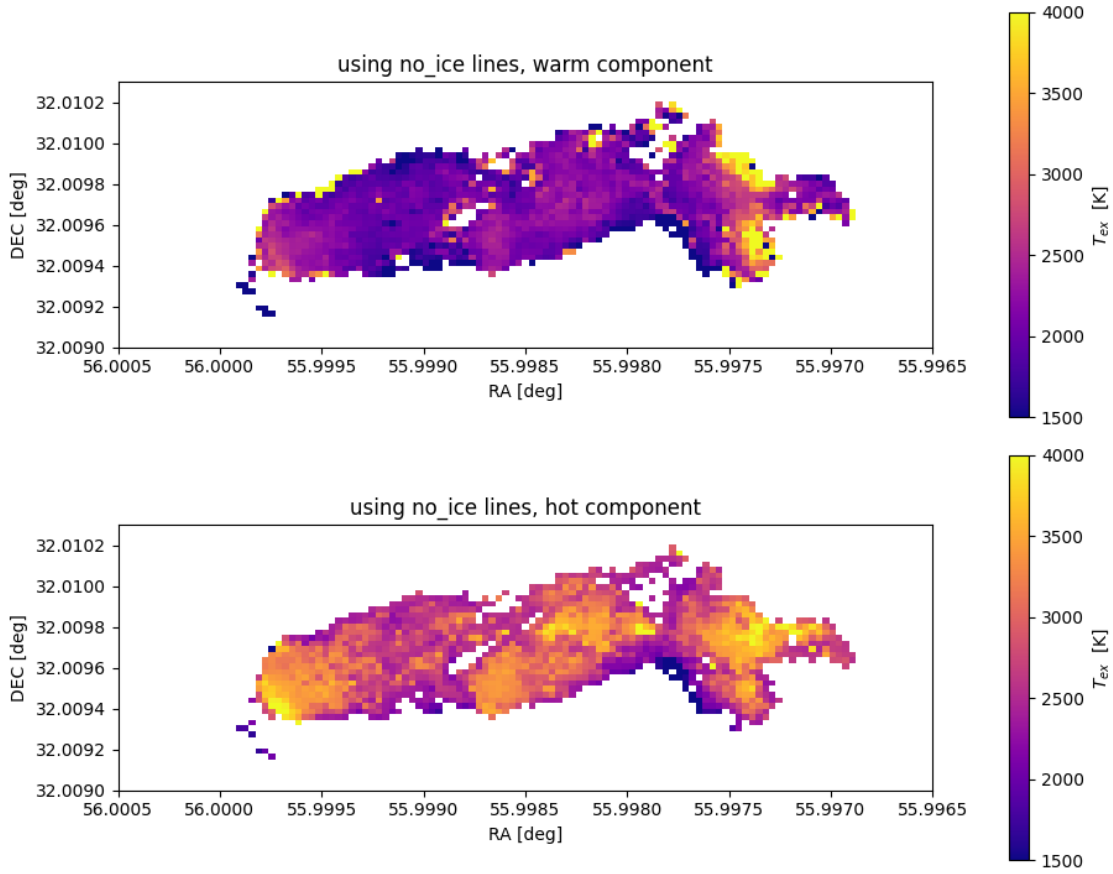


Figure 8.2: The excitation temperatures of the warm (top plot) and the hot (bottom plot) component. As the names suggest, the excitation temperatures are indeed higher for the hot component. For both components higher temperatures are found in the bow shocks. In the warm component of the bow shock closest to the source, an area of very high excitation temperatures can be seen; however, the errors for these fits are rather high.

8.2 Ortho-to-para ratios

While the results reported in section 8.1 used lines of ortho and para hydrogen to derive the excitation conditions, I also carried out another set of calculations separating them. A problem with this method is the lower number of lines that is available for each individual fit. Especially for the more restrictive line filter — cutting all lines that might be affected by the ice feature — only a limited number of pixels could be fitted. If the OPR is higher than 0.5, this affects the analysis for the para lines especially as they are dimmer than similar ortho lines and, thus, fewer para lines are detected. After some deliberation — and inspecting all of the results — I value reliability more than a higher number of calculated OPRs and report the properties derived for the more restrictive set of lines.

In a first attempt, I carried out the excitation analysis as before, but just for different subsets of lines. For this approach to work well, the derived excitation temperatures have to be similar for the ortho and the para lines. If this is not the case, the different slopes of the fits will affect the derived column densities too. So before inspecting the column densities and OPRs, I compare the excitation temperatures. I show the comparison of ortho and para excitation conditions for the warm and hot components in Figure 8.3. While the T_{ex} derived for the respective warm components (top two panels) are very similar to each other, this is not the case for the hot component (bottom two panels). The excitation temperatures of the hot component derived from the ortho lines (3rd panel from the top) closely resemble the values found for the full set of lines in the bottom panel of Figure 8.2. This is consistent with the expectation that more ortho lines are detected — especially near the high energy limit. The excitation temperatures derived for the hot component of the para lines (bottom panel of Figure 8.3), on the other hand, do not share this similarity. The T_{ex} found there are significantly lower and barely reach values above 3000 K.

These different trends can be explained quite easily, when considering the distribution of ortho and para lines in the studied spectra. Overall more ortho lines are detected and they have a larger range of corresponding upper level energies. In contrast to this, the para lines detected are related to lower upper level energies. I visualize this for the lines of the example pixel in Figure 8.4. My excitation analysis routine cuts the — sorted with respect to their T_u — list of line fits in the middle to create the two components. This means that, as the ortho and the para lines sample different ranges of upper level energies, their cut-offs between the warm and the hot component are located at different T_u. This difference in the cut-offs leads to different derived parameters, just like discussed in section 8.1 while comparing my results to those of Caratti o Garatti et al. (2024).

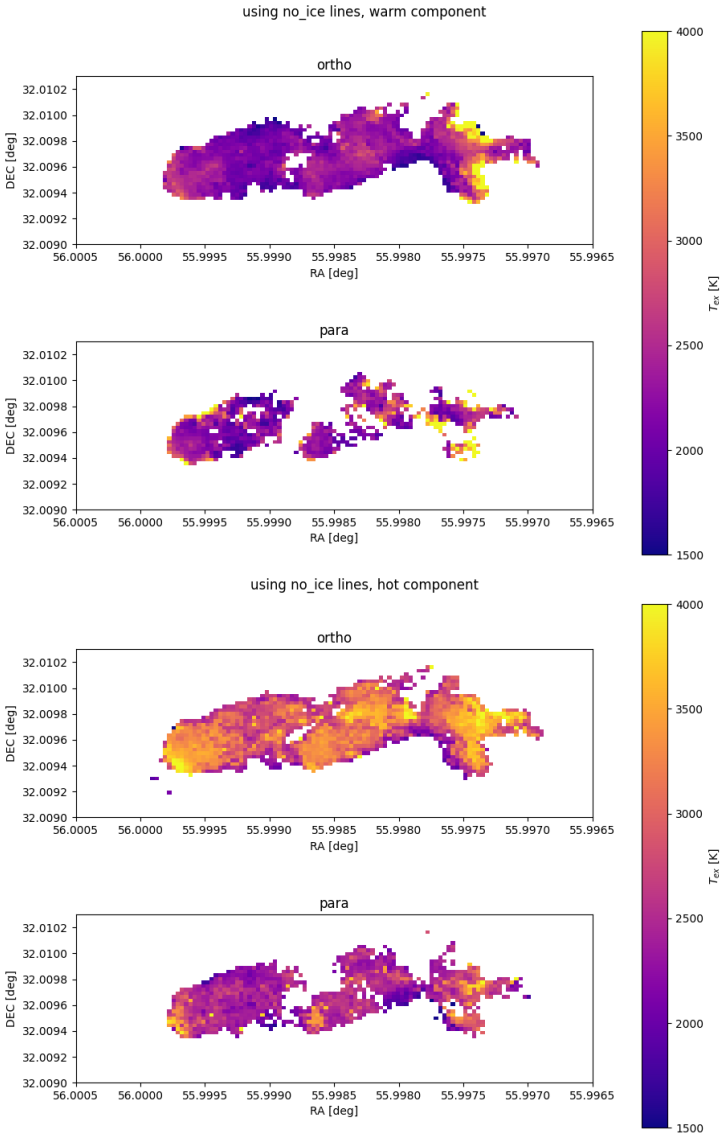


Figure 8.3: Comparison of the excitation temperatures of the hot and warm components derived from ortho or para lines. From top to bottom the panels show the warm component of the ortho lines, the warm component of the para lines, the hot component of the ortho lines and the hot component of the para lines. Comparing the two warm components the values are very similar to each other and to the values derived for the combined excitation diagrams. The two hot components show quite different values. From the third panel and its close resemblance to the bottom panel in Figure 8.2, it is clear that the hot component of the combined excitation diagrams is dominated by the contribution of the ortho lines. The overall lower number of para lines becomes evident by the lower number of successfully fitted pixels.

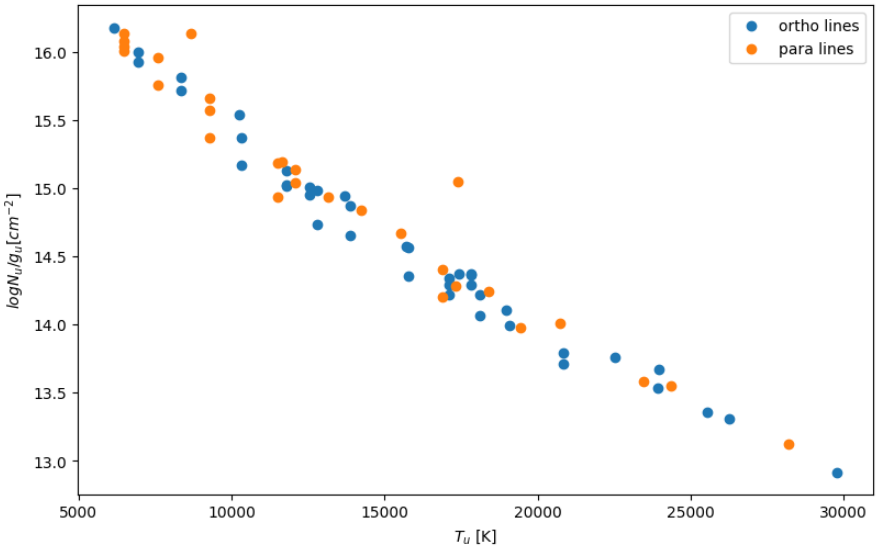


Figure 8.4: A comparison of the (T_u, N_u/g_u) data points associated with the ortho and para lines. The data is from the spectrum of the example pixel and the same as shown in the top panel of Figure 6.2. In this specific pixel I detected 28 para and 40 ortho lines. In addition to the different number of detected lines, the lines of the two spin isomers also probe different upper level energies. Near the high energy limit — e.g. at $T_u > 20000$ K — the large majority of detected lines are ortho lines. This leads to the ortho lines dominating the hot component of the combined excitation diagram. Another consequence are the lower derived T_{ex} for the hot component of the para lines.

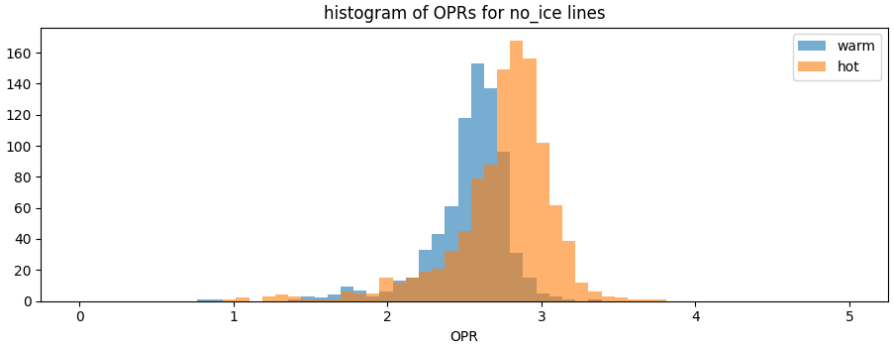


Figure 8.5: A histogram showing the OPRs derived for the warm and the hot component. While there appear to be two separate peaks, this difference is mostly within the errors of the warm component and fully within the errors of the hot component.

While I can explain the different T_{ex} derived for the ortho and para lines, they are very detrimental to the study of the OPRs. I handle this problem by fixing T_{ex} to the same value — I used the values derived from the combined excitation diagrams — for the analysis of both the para and ortho lines. With this, I can guarantee the same T_{ex} so that the only difference is in the N_0. The OPRs derived this way range from 2 to 3 for the warm component and from 2.25 to 3.25 for the hot component. The corresponding errors are mostly below 0.25 for the warm component and mostly below 0.5 for the hot component. When visualizing these values in a histogram — like in Figure 8.5, two separate peaks are visible. Most OPRs of the warm component are close to 2.5 while most OPRs of the hot component are around 2.75. Within the uncertainties most values of the hot component are close to the expected equilibrium value of 3, which is not the case for the warm component. Given the relatively high errors, though, this difference between the components is not quite significant. The tentative difference is contrary to what would be expected from the shock simulations discussed in section 1.4.2. These simulations predict lower OPRs if lines from higher energy levels are used as tracers. A plausible reason for this unexpected trend might again be in the way, in which the lines are separated into the two components. As discussed before, the limiting upper level energy is different between ortho and para lines. So, keeping the excitation temperature fixed might also lead to some non-optimal results for the column densities.

In Figure 8.6, I show the spatial distribution of the calculated OPRs. Given the very similar results for the warm (top) and hot (bottom) components, these maps appear near identical. The distribution of values is overall mostly uniform. There seems to be a small region of lower OPRs near the northern edge of the furthest bow shock. This is, again, a cautious claim due to the somewhat high uncertainties.

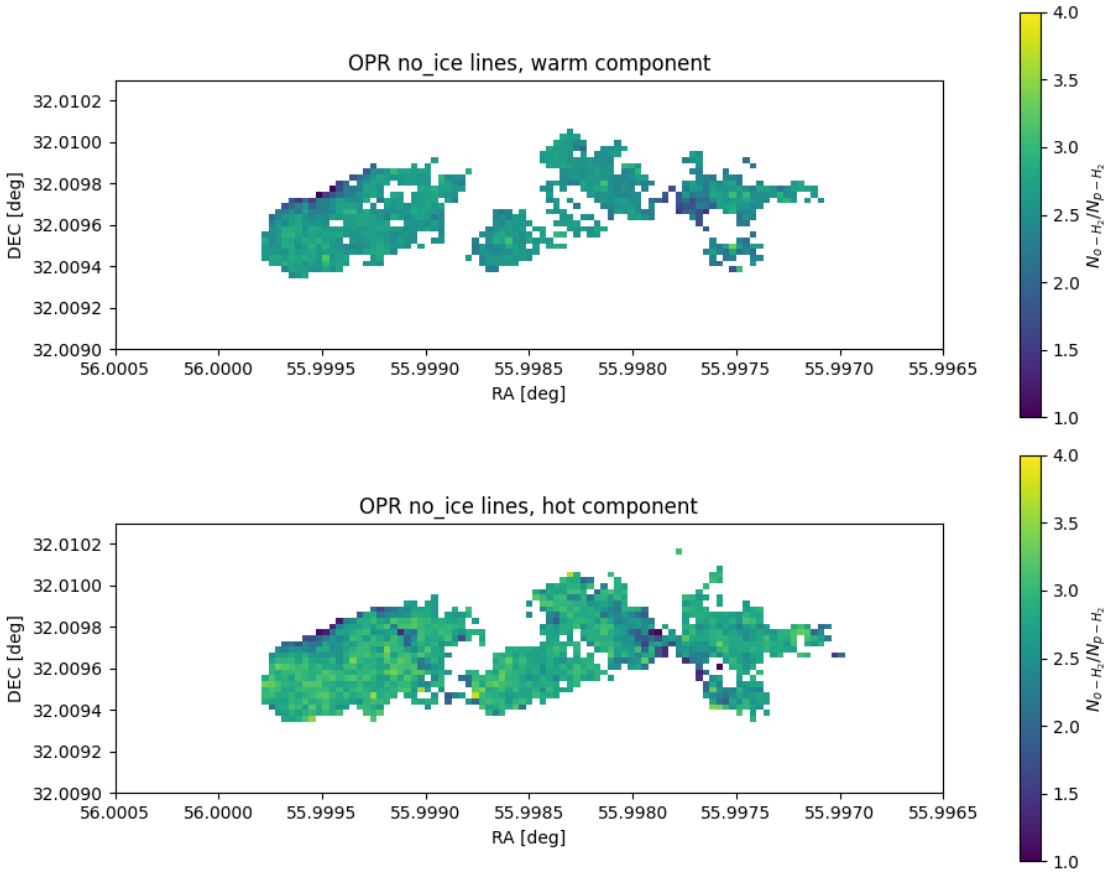


Figure 8.6: The spatial distribution of the OPRs calculated for the warm (top) and hot (bottom) component. The values found are — within the uncertainties — mostly the same for both components and there are no clear spatial trends visible.

8.3 Extinction correction and maps

During the testing phase in chapter 6, I carried out extinction analysis for a number of reddening laws and pairs of lines. As I illustrated in Figure 6.2 the extinction noticeably affects the distribution of data points in the excitation diagrams. This was expected as HH211 is associated with a young embedded protostar. Using a power law correction worked well for the (1-0) S and (1-0) Q branches. However, as seen in the top panel of Figure 6.3 the lines affected by the 3 μm ice feature recover systematically lower upper level column densities. That can be mitigated by using a reddening law that includes the ice feature as seen in the bottom plot of Figure 6.3. Although the data points corresponding to the same upper states are closer with such a reddening law, there is still a discrepancy of A_V values determined depending on the line pair used as a reference. This discrepancy would indicate that the used reddening law is not a good fit for the intrinsic one.

As mentioned in chapter 7, I carried out extinction analysis for the full data set in 6 variations. I used two different references — the (1-0) S(5) and (1-0) Q(7) pair and the (1-0) S(1) and (1-0) O(5) pair — and 3 different reddening laws — the Milky Way power law and the ones from the two dense sightlines from Decleir et al. (2022).
For the power-law reddening curve I determine A_V values between 2.5 and 15 mag for the (1-0) S(5) and (1-0) Q(7) pair and values between 0 and 7.5 mag for the (1-0) S(1) and (1-0) O(5) pair. For the first pair most A_V are around 10 mag and for the second the most common value is around 4.5 mag. For both pairs the associated uncertainties of A_V are mostly below 0.5 mag — the biggest uncertainty is introduced by the choice of the reddening law and not by the measuring uncertainties of the values used to calculate A_V. I show the corresponding extinction maps in Figure 8.7. The map used for the extinction corrections is the one in the bottom panel.

The additional extinction maps calculated using the reddening laws from the dense sightlines and taking into account the 3 μm ice feature are shown in Figure 8.8. These maps illustrate, how the calculated A_V values depend on the choice of the reddening law. Of particular note is the much larger effect of this choice on the extinctions estimated for the (1-0) S(1) and (1-0) O(5) pair. On the other hand, the values calculated for the (1-0) S(5) and (1-0) Q(7) pair are quite consistent across the different reddening laws. Across all maps shown, the extinction is higher in the furthest and the middle bow shocks, with an area of particularly high A_V values near the apex of the terminal bow shock.

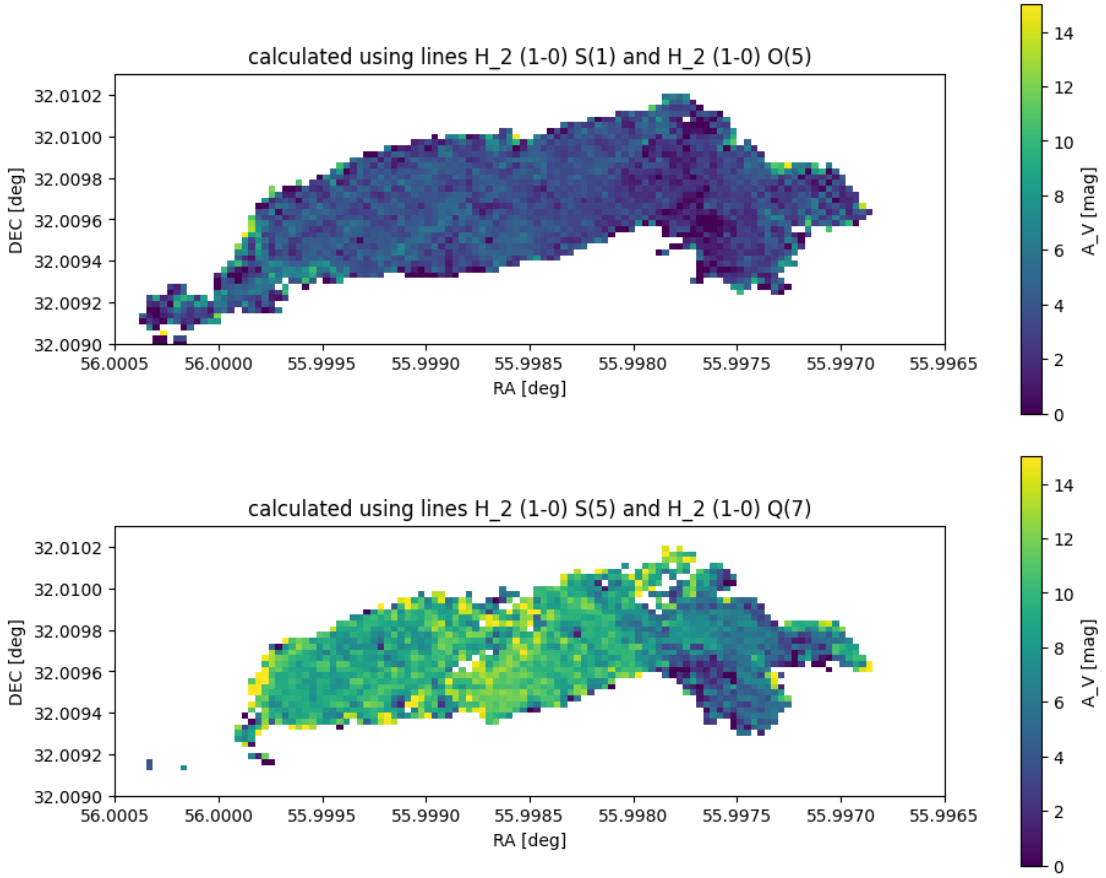


Figure 8.7: Extinction maps calculated using a power-law approximation for the reddening law. The reference lines for the upper plot are the lines (1-0) S(1) and (1-0) O(5) and for the lower plot they are the (1-0) S(5) and (1-0) Q(7) lines.

The values derived from the (1-0) S(5) and (1-0) Q(7) pair seem mostly invariant with changes of the reddening law — this was to be expected as the curves mainly differ in their modeling of the 3 μm ice feature, which none of the lines of this pair fall into. The wings of that feature affect the (1-0) O(5) line resulting in significant differences depending on the chosen reddening law. However, the values obtained by the different pairs of lines do not agree on the visual extinction for any of the tested reddening laws. This is an indicator, that the underlying reddening law is not simply a re-scaling of any of the tested curves and that the ice feature is not well described by any of them.
This is unsurprising as neither of the dense sightlines are particularly close to HH211 and both reddening laws were observed for main-sequence B stars that would already have removed a lot of ice from their immediate surroundings (Decleir et al., 2022).

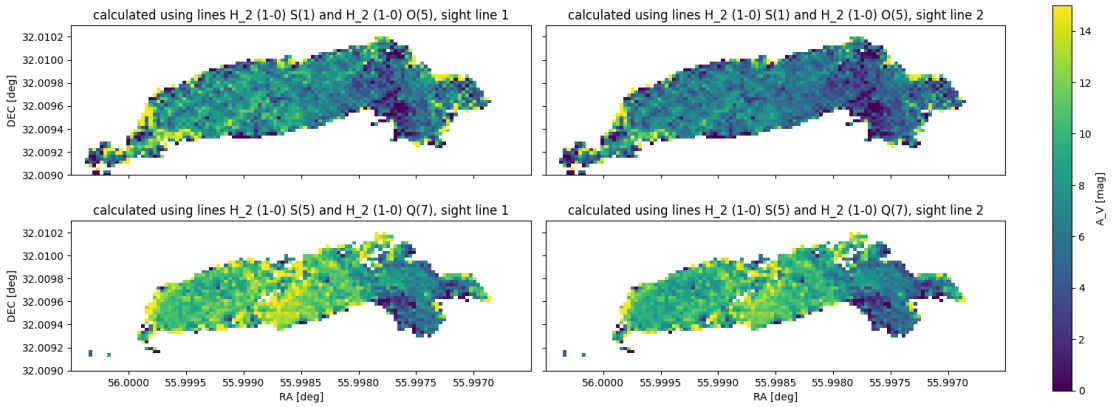


Figure 8.8: Additional extinction maps calculated using the modified Drude profiles for the reddening law; the columns correspond to the different sightlines. Again, the reference lines for the upper plots were the lines (1-0) S(1) and (1-0) O(5) and those for the lower plots (1-0) S(5) and (1-0) Q(7). Note the significant differences of extinctions calculated for the (1-0) S(1) and (1-0) O(5) pair especially compared to the power-law extinction calculations in the upper panel of Figure 8.7.

The extinction correction could be improved by adapting the parameters of the included Drude profiles to directly match the observed line ratios. Using a larger number of line pairs originating for common upper states concurrently might be enough to determine good estimates for all of the necessary parameters. From the testing it seems like the ice feature is underestimated by all of the tested reddening laws, which is again not surprising considering where they were observed.

In comparison to the extinctions reported in O'Connell et al. (2005) (∼16 mag in the bow shock region) and Caratti o Garatti et al. (2024) (5–15 mag), the values found in this work are overall a bit lower. Given the different lines and extinction curves used, they are, however, in the same range.

9 Future Prospects

During my work on this project it became apparent just how much information could be extracted from this data set and how deeply it could be analyzed. The following is a short overview of analysis that could be carried out using the same observations, but was not done as it would far exceed the intended scope of this thesis.

- **Shock modeling or quantitative comparison to existing shock models:** The emission observed and studied in this thesis stems from shock excitation. This is evident by the shape of the emitting regions as well as the excitation structure seen in the excitation diagrams. As mentioned in section 1.4.2, the observed line emission contains information about the intrinsic shock properties like the shock velocity, the pre-shock density or the strength of an external magnetic or UV radiation field. These properties can be constrained using shock simulations. In this work I did a qualitative comparison of my excitation diagrams against those found in Kristensen et al. (2023). I did, however, not determine which model of the grid introduced in that paper provides the best approximation of the observed data. Taking things even further, one could carry out independent and tailor-made shock simulations to ensure an optimal modeling of the observed emission.

- **Detailed velocity structure of the bow shocks:** From the line properties extracted in section 4.2, the study of the velocity structure is only a Doppler shift calculation away. I did this during the early phases of the project to ensure that the line fitting and matching worked properly. The recovered velocity maps vary slightly depending on the line that is used as a tracer and show quite significant differences between the shocks and the region between them. Given NIRSpec's high resolution and the excellent signal-to-noise ratios of this observation, these calculations could be done with a high accuracy. The detailed processes leading to the observed velocity structure were not further investigated.

- **OPRs with the MIRI data:** While Caratti o Garatti et al. (2024) conducted most of the same analysis as done here, OPRs were not calculated for the MIRI data. The routines developed for my project have also been successfully tested on the MIRI data. The required reprocessing of the data available from the archive for proper background subtraction and the necessary re-sampling to match the different pixel scales, prevented me from extending my analysis to this additional data set.

- **Detailed Extinction analysis:** As discussed in section 8.3, the reddening laws used here, only work well for a part of the spectrum with the 3 μm ice feature causing significant deviations. Making full use of the large number of lines observed in the bow shocks, this feature could be better characterized. In some cases not only pairs of lines corresponding to the same upper state are observed, but also triplets and in some rare cases even quadruplets. My testing has shown, that the differential extinction caused between the (1-0) S and (1-0) Q branch can be corrected quite well using a power law reddening curve. This reference together with the lines falling into the ice feature would allow for the characterization of the deviation from that power law.

10 Conclusion

In this thesis I present my analysis of the JWST NIRSpec observations of the terminal bow shocks of HH211. My investigation includes the excitation conditions of molecular hydrogen in that region as well as the OPRs and the extinction conditions. I developed and implemented a dedicated tool set — called *nirspecpy*. These fine-tuned analyzing routines together with the excellent resolution and sensitivity of the NIRSpec data provide an unprecedented view into the conditions of HH211's terminal bow shocks. My main conclusions can be summarized as follows:

- The observational data set is incredibly information-dense with numerous bright H_2 and CO lines and a low noise level.
- Given the size of the data set and the two separate detectors capturing each spectrum, the extraction of line properties provides some challenges. With a number of tactically chosen hyper-parameters and adaptations — like fitting the baselines in multiple components and masking high fluxes during baseline correction, allowing for different line flux thresholds in different parts of the spectra and parallelizing the entire process — I could accomplish this for the entire data set in a mostly automatic fashion.
- The analysis of the excitation conditions on a single-pixel scale reveals the non-LTE nature of the observed emission, as data points in the excitation diagrams do not follow a linear trend but rather a curve. Allowing two separate components greatly improves the accuracy of the excitation fits.
- The accuracy of the excitation analysis is greatly hindered by the effects of extinction, which cause a large vertical dispersion of the data points in the excitation diagrams. Of the tested reddening laws none seems to provide a particularly good match for the entire wavelength range. The 3 μm ice feature has significant influence and is likely underestimated in the reddening laws used.
- The two components have excitation temperatures ranging from 1500 to 2500 K and from 2000 to 3500 K. The column densities vary from $\sim 2 \cdot 10^{18}$ to $\sim 2 \cdot 10^{19}$ cm^{-2} and $\sim 6 \cdot 10^{17}$ to $\sim 3 \cdot 10^{18}$ cm^{-2} across the observed region. With uncertainties predominantly below 23 % for the column density and below 200 K for the excitation temperature, these properties are mostly well constrained. Generally, high excitation

temperatures coincide with high column densities. There are, however, some regions with different trends. The bow shocks — most prominently the terminal one — appear asymmetric in the spatial distribution of column densities and excitation temperatures.

- The determined OPRs are similar across the entire observed region. There is a slight difference between the two components, with the most common OPR for the warm component being around 2.5 and most values being around 2.75 for the hot component. This trend is rather uncertain as the errors are relatively high at 0.25 and 0.5 respectively. If genuine, the trend is contrary to the expectation from shock simulations. Another reason for the observed difference could be in the specifics of the method used in this thesis.
- In the extinction analysis I established the importance of the choice of reference lines and the influence of the reddening law. I find the highest values for the visual extinction near the apex of the furthest and the middle bow shock; the region of the closest bow shock shows rather low extinctions. Overall, the extinctions found are between 2.5 and 15 mag with errors mostly below 0.5 mag.
- There are numerous additional avenues for further analysis of this data set. Examples include modeling of the shock conditions, detailed analysis of the velocity structure of the shocks and detailed extinction analysis using multiple lines concurrently.

Overall, the analysis carried out here allows for a detailed look into the excitation conditions of the terminal bow shocks of HH211, the OPRs of that region and the extinction caused by the surrounding material. My results are consistent with prior studies of HH211 and improve on the spatial resolution as well as the uncertainties of derived properties.

References

Andre, P., D. Ward-Thompson and M. Barsony (May 2000). "From Prestellar Cores to Protostars: the Initial Conditions of Star Formation". In: *Protostars and Planets IV*. Ed. by V. Mannings, A. P. Boss and S. S. Russell, p. 59. DOI: 10.48550/arXiv.astro-ph/9903284. arXiv: astro-ph/9903284 [astro-ph].

Andre, P., D. Ward-Thompson and M. Barsony (Mar. 1993). "Submillimeter Continuum Observations of rho Ophiuchi A: The Candidate Protostar VLA 1623 and Prestellar Clumps". In: *ApJ* 406, p. 122. DOI: 10.1086/172425.

Astropy Collaboration, A. M. Price-Whelan, B. M. Sipőcz et al. (Sept. 2018). "The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package". In: *AJ* 156.3, 123, p. 123. DOI: 10.3847/1538-3881/aabc4f. arXiv: 1801.02634 [astro-ph.IM].

Astropy Collaboration, A. M. Price-Whelan, P. L. Lim et al. (Aug. 2022). "The Astropy Project: Sustaining and Growing a Community-oriented Open-source Project and the Latest Major Release (v5.0) of the Core Package". In: *ApJ* 935.2, 167, p. 167. DOI: 10.3847/1538-4357/ac7c74. arXiv: 2206.14220 [astro-ph.IM].

Astropy Collaboration, T. P. Robitaille, E. J. Tollerud et al. (Oct. 2013). "Astropy: A community Python package for astronomy". In: *A&A* 558, A33, A33. DOI: 10.1051/0004-6361/201322068. arXiv: 1307.6212 [astro-ph.IM].

Bally, J. (Sept. 2016). "Protostellar Outflows". In: *ARA&A* 54, pp. 491–528. DOI: 10.1146/annurev-astro-081915-023341.

Böker, T., S. Arribas, N. Lützgendorf et al. (May 2022). "The Near-Infrared Spectrograph (NIRSpec) on the James Webb Space Telescope. III. Integral-field spectroscopy". In: *A&A* 661, A82, A82. DOI: 10.1051/0004-6361/202142589. arXiv: 2202.03308 [astro-ph.IM].

Caratti o Garatti, A., T. P. Ray, P. J. Kavanagh et al. (Nov. 2024). "JWST Observations of Young protoStars (JOYS): HH211: Textbook case of a protostellar jet and outflow". In: *A&A* 691, A134, A134. DOI: 10.1051/0004-6361/202451350. arXiv: 2409.16061 [astro-ph.SR].

Decleir, M., K. D. Gordon, J. E. Andrews et al. (May 2022). "SpeX Near-infrared Spectroscopic Extinction Curves in the Milky Way". In: *ApJ* 930.1, 15, p. 15. DOI: 10.3847/1538-4357/ac5dbe. arXiv: 2204.13716 [astro-ph.GA].

Dionatos, O., B. Nisini, S. Cabrit, L. Kristensen and G. Pineau Des Forêts (Oct. 2010). "Spitzer spectral line mapping of the HH211 outflow". In: *A&A* 521, A7, A7. DOI: 10.1051/0004-6361/200913650. arXiv: 1006.0821 [astro-ph.GA].

Dionatos, O., T. Ray and M. Güdel (Aug. 2018). "Herschel spectral-line mapping of the HH211 protostellar system". In: *A&A* 616, A84, A84. DOI: 10.1051/0004-6361/201833057. arXiv: 1806.00311 [astro-ph.SR].

Draine, B. T. (2011). *Physics of the Interstellar and Intergalactic Medium*.

Ferruit, P., P. Jakobsen, G. Giardino et al. (May 2022). "The Near-Infrared Spectrograph (NIRSpec) on the James Webb Space Telescope. II. Multi-object spectroscopy (MOS)". In: *A&A* 661, A81, A81. DOI: 10.1051/0004-6361/202142673. arXiv: 2202.03306 [astro-ph.IM].

Frank, A., T. P. Ray, S. Cabrit et al. (Jan. 2014). "Jets and Outflows from Star to Cloud: Observations Confront Theory". In: *Protostars and Planets VI*. Ed. by H. Beuther, R. S. Klessen, C. P. Dullemond and T. Henning, pp. 451–474. DOI: 10.2458/azu_uapress_9780816531240-ch020. arXiv: 1402.3553 [astro-ph.SR].

Gardner, J. P., J. C. Mather, R. Abbott et al. (2023). "The James Webb Space Telescope Mission". In: *Publications of the Astronomical Society of the Pacific* 135.1048, p. 068001. DOI: 10.1088/1538-3873/acd1b5. URL: https://dx.doi.org/10.1088/1538-3873/acd1b5.

Goldsmith, P. F. and W. D. Langer (May 1999). "Population Diagram Analysis of Molecular Line Emission". In: *ApJ* 517.1, pp. 209–225. DOI: 10.1086/307195.

Goorvitch, D. (Dec. 1994). "Infrared CO Line List for the X 1 Sigma + State". In: *ApJS* 95, p. 535. DOI: 10.1086/192110.

Harris, C. R., K. J. Millman, S. J. van der Walt et al. (Sept. 2020). "Array programming with NumPy". In: *Nature* 585.7825, pp. 357–362. DOI: 10.1038/s41586-020-2649-2. URL: https://doi.org/10.1038/s41586-020-2649-2.

Hartmann, L., G. Herczeg and N. Calvet (Sept. 2016). "Accretion onto Pre-Main-Sequence Stars". In: *ARA&A* 54, pp. 135–180. DOI: 10.1146/annurev-astro-081915-023347.

Hunter, J. D. (2007). "Matplotlib: A 2D graphics environment". In: *Computing in Science & Engineering* 9.3, pp. 90–95. DOI: 10.1109/MCSE.2007.55.

Kristensen, L. E., B. Godard, P. Guillard, A. Gusdorf and G. Pineau des Forêts (July 2023). "Shock excitation of H_2 in the James Webb Space Telescope era". In: *A&A* 675, A86, A86. DOI: 10.1051/0004-6361/202346254. arXiv: 2307.04178 [astro-ph.GA].

Lada, C. J. (Jan. 1987). "Star formation: from OB associations to protostars." In: *Star Forming Regions*. Ed. by M. Peimbert and J. Jugaku. Vol. 115. IAU Symposium, p. 1.

Lebigot, E. O. (2010). *Uncertainties: a Python package for calculations with uncertainties*. https://uncertainties.readthedocs.io/en/latest/index.html. Accessed: 2024-08-08.

Lee, C.-F., Z.-Y. Li, N. Hirano et al. (Aug. 2018). "ALMA Observations of the Very Young Class 0 Protostellar System HH211-mms: A 30 au Dusty Disk with a Disk Wind Traced by SO?" In: *ApJ* 863.1, 94, p. 94. DOI: 10.3847/1538-4357/aad2da. arXiv: 1807.05336 [astro-ph.GA].

McCaughrean, M. J., J. T. Rayner and H. Zinnecker (Dec. 1994). "Discovery of a Molecular Hydrogen Jet near IC 348". In: *ApJ* 436, p. L189. DOI: 10.1086/187664.

McKerns, M., L. Strand, T. Sullivan, A. Fang and M. Aivazis (2011). "Building a framework for predictive science". In: *Proceedings of the 10th Python in Science Conference*.

McKerns, M. and M. Aivazis (2010). *pathos: a framework for heterogeneous computing*. https://uqfoundation.github.io/project/pathos.

McKinney, W. (2010). "Data Structures for Statistical Computing in Python". In: *Proceedings of the 9th Python in Science Conference*. Ed. by Stéfan van der Walt and Jarrod Millman, pp. 56 –61. DOI: 10.25080/Majora-92bf1922-00a.

Neufeld, D. A., G. J. Melnick and M. Harwit (Aug. 1998). "ISO observations of molecular hydrogen in HH 54: measurement of a non-equilibrium ortho- to para-H2 ratio". In: *arXiv e-prints*, astro-ph/9808101, astro–ph/9808101. DOI: 10.48550/arXiv.astro-ph/9808101. arXiv: astro-ph/9808101 [astro-ph].

Newville, M., R. Otten, A. Nelson et al. (July 2024). *lmfit/lmfit-py: 1.3.2*. Version 1.3.2. DOI: 10.5281/zenodo.12785036. URL: https://doi.org/10.5281/zenodo.12785036.

O'Connell, B., M. D. Smith, D. Froebrich, C. J. Davis and J. Eislöffel (Feb. 2005). "The near-infrared excitation of the HH 211 protostellar outflow". In: *A&A* 431, pp. 223–234. DOI: 10.1051/0004-6361:20041821. arXiv: astro-ph/0502107 [astro-ph].

Ray, T. P. and J. Ferreira (Dec. 2021). "Jets from young stars". In: *New A Rev.* 93, 101615, p. 101615. DOI: 10.1016/j.newar.2021.101615. arXiv: 2009.00547 [astro-ph.SR].

Ray, T. P., M. J. McCaughrean, A. Caratti o Garatti et al. (Oct. 2023). "Outflows from the youngest stars are mostly molecular". In: *Nature* 622.7981, pp. 48–52. DOI: 10.1038/s41586-023-06551-1.

Ray, T. P., H. Beuther, A. Caratti o Garatti et al. (July 2017). *The Young Protostellar Outflow HH211*. JWST Proposal. Cycle 1, ID. #1257.

Ray, T. P., A. Caratti o Garatti, M. J. McCaughrean et al. (2024). *The Young Protostellar Outflow HH211*. Mikulski Archive for Space Telescopes at the Space Telescope Science Institute https://ui.adsabs.harvard.edu/link_gateway/2023Natur.622...48R/MAST.

Rieke, G. H. and M. J. Lebofsky (Jan. 1985). "The interstellar extinction law from 1 to 13 microns." In: *ApJ* 288, pp. 618–621. DOI: 10.1086/162827.

Robitaille, T. P., B. A. Whitney, R. Indebetouw, K. Wood and P. Denzmore (Dec. 2006). "Interpreting Spectral Energy Distributions from Young Stellar Objects. I. A Grid of 200,000 YSO Model SEDs". In: *ApJS* 167.2, pp. 256–285. DOI: 10.1086/508424. arXiv: astro-ph/0608234 [astro-ph].

Roueff, E., H. Abgrall, P. Czachorowski et al. (Oct. 2019a). "The full infrared spectrum of molecular hydrogen". In: *A&A* 630, A58, A58. DOI: 10.1051/0004-6361/201936249. arXiv: 1909.11585 [physics.atom-ph].

– (Aug. 2019b). *VizieR Online Data Catalog: Full infrared spectrum of molecular hydrogen (Roueff+, 2019)*. VizieR Online Data Catalog: J/A+A/630/A58. Originally published in: 2019A&A...630A..58R. DOI: 10.26093/cds/vizier.36300058.

Schwartz, R. D. (Jan. 1983). "Herbig-haro objects." In: *ARA&A* 21, pp. 209–237. DOI: 10.1146/annurev.aa.21.090183.001233.

The pandas development team (Apr. 2024). *pandas-dev/pandas: Pandas*. Version v2.2.2. DOI: 10.5281/zenodo.10957263. URL: https://doi.org/10.5281/zenodo.10957263.

Van De Putte, D., R. Meshaka, B. Trahin et al. (July 2024). "PDRs4All. VIII. Mid-infrared emission line inventory of the Orion Bar". In: *A&A* 687, A86, A86. DOI: 10.1051/0004-6361/202449295. arXiv: 2404.03111 [astro-ph.GA].

Vaytet, N., K. Tomida and G. Chabrier (Mar. 2014). "On the role of the H_2 ortho:para ratio in gravitational collapse during star formation". In: *A&A* 563, A85, A85. DOI: 10.1051/0004-6361/201322855. arXiv: 1401.4299 [astro-ph.SR].

Virtanen, P., R. Gommers, T. E. Oliphant et al. (2020). "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17, pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

Wells, M., J. W. Pel, A. Glasse et al. (July 2015). "The Mid-Infrared Instrument for the James Webb Space Telescope, VI: The Medium Resolution Spectrometer". In: *PASP* 127.953, p. 646. DOI: 10.1086/682281. arXiv: 1508.03070 [astro-ph.IM].

Wilgenbus, D., S. Cabrit, G. Pineau des Forêts and D. R. Flower (Apr. 2000). "The ortho:para-H_2 ratio in C- and J-type shocks". In: *A&A* 356, pp. 1010–1022.

Williams, J. P. and L. A. Cieza (Sept. 2011). "Protoplanetary Disks and Their Evolution". In: *ARA&A* 49.1, pp. 67–117. DOI: 10.1146/annurev-astro-081710-102548. arXiv: 1103.0556 [astro-ph.GA].

Wright, G. S., D. Wright, G. B. Goodson et al. (July 2015). "The Mid-Infrared Instrument for the James Webb Space Telescope, II: Design and Build". In: *PASP* 127.953, p. 595. DOI: 10.1086/682253. arXiv: 1508.02333 [astro-ph.IM].

Wright, G. S., G. H. Rieke, A. Glasse et al. (Apr. 2023). "The Mid-infrared Instrument for JWST and Its In-flight Performance". In: *PASP* 135.1046, 048003, p. 048003. DOI: 10.1088/1538-3873/acbe66.

Young, A. K. (Dec. 2023). "Insights into the first and second hydrostatic core stages from numerical simulations". In: *Frontiers in Astronomy and Space Sciences* 10, 1288730, p. 1288730. DOI: 10.3389/fspas.2023.1288730. arXiv: 2312.03039 [astro-ph.SR].

van Dishoeck, E. F., Ł. Tychoniec, W. R. M. Rocha et al. (May 2025). "JWST Observations of Young protoStars (JOYS): overview of program and early results". In: *arXiv e-prints*, arXiv:2505.08002, arXiv:2505.08002. DOI: 10.48550/arXiv.2505.08002. arXiv: 2505.08002 [astro-ph.GA].

Acronyms

CSA	Canadian Space Agency
ESA	European Space Agency
FOV	field of view
HH	Herbig-Haro
HH211	Herbig-Haro 211
IFU	integral field unit
JWST	James Webb Space Telescope
LTE	local thermodynamic equilibrium
MIRI	Mid InfraRed Instrument
MOS	multi-object spectroscopy
MRS	Medium Resolution Spectrometer
MSA	Micro Shutter Array
NASA	National Aeronautics and Space Administration
NIRCam	Near InfraRed Camera
NIRISS	Near InfraRed Imager and Slitless Spectrograph
NIRSpec	Near InfraRed Spectrograph
OPR	ortho-to-para ratio
SED	spectral energy distribution
YSO	young stellar object