



# MASTERARBEIT | MASTER'S THESIS

Titel | Title

Gas Kinematics in the Nessie Cloud with ALMA

verfasst von | submitted by

Mathias Lechthaler BSc BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of  
Master of Science (MSc)

Wien | Vienna, 2026

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 Astronomy

Betreut von | Supervisor:

Assoz. Prof. Dr. Alvaro Hacar Gonzalez MA



*Nicht von sich auf andere schließen.*





# Acknowledgements

Zuallererst möchte ich mich bei meinen Eltern, Roland und Roswitha Lechthaler, bedanken, die mir überhaupt erst ermöglicht haben, mein Studium durchzuführen. Sie haben mich dabei in jeder Hinsicht unterstützt, niemals unter Druck gesetzt und mich meinen Weg gehen lassen. Ohne sie würde ich heute nicht hier sein und diese Arbeit würde nicht existieren.

Next, I have to thank Alvaro Hacar for being an excellent supervisor, for guiding me through the intricacies of radio astronomy, for the insightful discussions we had, and for supporting me in finishing this thesis.

I also want to thank the entire EMERGE team, especially Francesca Bonanomi, Andrea Succi, Lukas Orion, and Darwin Lorber. They helped me starting with interferometric data reduction, data combination, and spectral line fitting by providing essential parts of my scripts. Without them, I would have taken much longer to work through all these topics.

Lastly, it should be declared that parts of the code used for analysis were written with the help of currently widely used AI tools, such as ChatGPT and Copilot, or Consensus for literature research. In the code, several Python libraries were used, among them `numpy` [Harris et al., 2020], `scipy` [Virtanen et al., 2020], `matplotlib` [Hunter, 2007], and `astropy` Astropy Collaboration et al. [2022]. For the phrasing in the thesis, ChatGPT or Claude were used. The thesis was written in  $\text{\LaTeX}$ .



# Abstract

This master’s thesis investigated the interstellar molecular cloud ”Nessie”. Due to its large extent of  $\sim 2.5^\circ$  or 160 pc at a distance of 3.1 kpc, Nessie can be classified as a giant molecular filament. The main goal of this thesis is to characterize the kinematic structure of Nessie, as the fragmentation process in molecular clouds – which sets the initial conditions of star formation – is still poorly understood. High-resolution spectral line observations allow us to probe both small-scale clump motions and large-scale velocity patterns along the filament. Nessie was observed in 2019 with two components of the ALMA interferometric telescope, the Total Power and the 7m array, which were combined to recover emission across multiple spatial scales. We process this interferometric data with the CASA software using the Model Assisted CLEAN + Feather approach, resulting in a high-resolution ( $28''$ ) map in the  $\text{N}_2\text{H}^+(1-0)$  and  $\text{HNC}(1-0)$  molecular transition lines. The distribution of velocity centroids of the spectral lines reveals a sinusoidal oscillation in position-velocity space with a wavelength of 102 pc, whose characteristic scale cannot be explained by classical gravo-turbulent fragmentation models. We propose that the origin of this undulation is a process on a Galactic scale, specifically Toomre’s instability, whose characteristic length scale matches the observed wavelength.



# Kurzfassung

Im Rahmen dieser Masterarbeit wurde die interstellare Molekülwolke "Nessie" untersucht. Aufgrund der ungewöhnlich großen Ausdehnung von  $\sim 2.5^\circ$  oder 160 pc bei einer Distanz von 3.1 kpc kann Nessie auch als Giant Molecular Filament klassifiziert werden. Das Hauptziel dieser Arbeit ist es, die kinematische Struktur von Nessie zu charakterisieren, da der Fragmentierungsprozess in Molekülwolken – der die Anfangsbedingungen für die Sternentstehung bestimmt – noch immer schlecht verstanden ist. Hochauflösende Beobachtungen von Spektrallinien ermöglichen es uns, sowohl kleinräumige Bewegungen in den Clumps als auch großräumige Geschwindigkeitsmuster entlang des Filaments zu untersuchen. Nessie wurde 2019 mit zwei Komponenten des ALMA Interferometrieteleskops beobachtet, dem Total Power und dem 7m Array, welche zusammengeführt werden, um Emissionen über mehrere Größenskalen zu erfassen. Wir führen diese interferometrischen Daten mit der Software CASA zusammen, indem wir die Model Assisted CLEAN + Feather Methode verwenden, was in einer hochauflösenden ( $28''$ ) Karte in den molekularen Übergangslinien  $\text{N}_2\text{H}^+(1-0)$  und  $\text{HNC}(1-0)$  resultiert. Die Verteilung der Geschwindigkeitsschwerpunkte der Spektrallinien zeigt eine sinusförmige Oszillation im Ort-Geschwindigkeits Raum mit einer Wellenlänge von 102 pc auf, dessen charakteristische Größe nicht mit klassischen gravo-turbulenten Fragmentationsmodellen erklärt werden kann. Wir vermuten, dass der Ursprung dieser Wellenbewegung ein Prozess auf galaktischer Skala ist, im speziellen die Toomre Instabilität, deren charakteristische Größenskala mit der beobachteten Wellenlänge übereinstimmt.



# Contents

|                                                                     |            |
|---------------------------------------------------------------------|------------|
| <b>Acknowledgements</b>                                             | <b>i</b>   |
| <b>Abstract</b>                                                     | <b>iii</b> |
| <b>Kurzfassung</b>                                                  | <b>v</b>   |
| <b>List of Tables</b>                                               | <b>ix</b>  |
| <b>List of Figures</b>                                              | <b>xi</b>  |
| <b>1. Introduction</b>                                              | <b>1</b>   |
| <b>2. Fragmentation of Giant Molecular Filaments</b>                | <b>5</b>   |
| 2.1. Gravitational and Dynamical Fragmentation . . . . .            | 5          |
| 2.2. Molecular Tracers of Kinematics in the ISM . . . . .           | 8          |
| <b>3. Interferometry, Data Reduction, and Data Combination</b>      | <b>11</b>  |
| 3.1. Atacama Large Millimeter/submillimeter Array . . . . .         | 11         |
| 3.2. Interferometry in a Nutshell . . . . .                         | 12         |
| 3.3. Data Combination . . . . .                                     | 15         |
| 3.3.1. The Common Astronomy Software Applications Package . . . . . | 15         |
| 3.3.2. Model Assisted CLEAN + Feather . . . . .                     | 16         |
| 3.3.3. Image Quality Assessment . . . . .                           | 17         |
| <b>4. Observational Data, Set-Up in CASA, and Resulting Maps</b>    | <b>19</b>  |
| 4.1. Observations . . . . .                                         | 19         |
| 4.2. MACF . . . . .                                                 | 19         |
| 4.2.1. Issues in (u,v)-Plane . . . . .                              | 19         |
| 4.2.2. MACF Set-Up . . . . .                                        | 23         |
| 4.2.3. Quality Assessment after MACF . . . . .                      | 25         |
| 4.2.4. Nyquist sampling . . . . .                                   | 28         |
| 4.3. Re-Projection to Galactic Coordinates and Mosaicing . . . . .  | 28         |
| 4.4. Resulting Maps in Two Molecular Tracers . . . . .              | 29         |
| 4.5. Spectral Line Fitting . . . . .                                | 30         |
| 4.5.1. $\text{N}_2\text{H}^+$ Line Fitting . . . . .                | 31         |
| 4.5.2. HNC Line Fitting . . . . .                                   | 32         |

## Contents

|                                                                                  |           |
|----------------------------------------------------------------------------------|-----------|
| <b>5. Kinematic Analysis of the Velocity Centroids</b>                           | <b>35</b> |
| 5.1. Sinusoidal Undulation in Position-Velocity Space . . . . .                  | 35        |
| 5.2. New Distance Estimate to Nessie . . . . .                                   | 38        |
| 5.2.1. Distance Estimation Based on Kinematic Distance . . . . .                 | 40        |
| 5.2.2. Distance Estimation Based on the Break Point in <i>Gaia</i> Colour Excess | 40        |
| 5.2.3. Distance Estimation Based on the Break Point in 2MASS Extinction          | 42        |
| 5.3. (No) De-Projection of Velocities . . . . .                                  | 44        |
| 5.4. Power Spectrum of PV-Space . . . . .                                        | 45        |
| <b>6. Discussion</b>                                                             | <b>49</b> |
| 6.1. Outlook . . . . .                                                           | 52        |
| <b>7. Conclusion</b>                                                             | <b>55</b> |
| <b>Bibliography</b>                                                              | <b>57</b> |
| <b>A. Appendix</b>                                                               | <b>67</b> |
| A.1. Detailed Spitzer, $\text{N}_2\text{H}^+$ , and HNC Maps . . . . .           | 67        |
| A.2. Comparison of Cloud Placement with Goodman et al. 2014 . . . . .            | 70        |



# List of Tables

|                                                                        |    |
|------------------------------------------------------------------------|----|
| 4.1. Observational Details for each Field in Nessie . . . . .          | 21 |
| 4.2. Comparison of Data Combination Methods . . . . .                  | 23 |
| 4.3. A-Par and Flux Recovery for all Individual Fields . . . . .       | 26 |
| 5.1. Fit Parameters to Both Molecules and the Combined Value . . . . . | 37 |



# List of Figures

|                                                                                               |    |
|-----------------------------------------------------------------------------------------------|----|
| 3.1. Interferometry Observation Sketch . . . . .                                              | 12 |
| 4.1. Spitzer False-Colour Image with $\text{N}_2\text{H}^+$ and HNC Integrated Intensity Maps | 20 |
| 4.2. $(u, v)$ -Coverage and Filling Fraction of F6 . . . . .                                  | 22 |
| 4.3. Workflow of Data Combination . . . . .                                                   | 24 |
| 4.4. Distribution of A-Par for F6 in $\text{N}_2\text{H}^+$ . . . . .                         | 27 |
| 4.5. Example Spectra of $\text{N}_2\text{H}^+(1-0)$ and $\text{HNC}(1-0)$ . . . . .           | 33 |
| 5.1. Velocity-Maps in $\text{N}_2\text{H}^+$ and HNC . . . . .                                | 36 |
| 5.2. Position-Velocity and Position-Intensity Plot . . . . .                                  | 39 |
| 5.3. Kinematic Distance Estimate to Nessie . . . . .                                          | 41 |
| 5.4. Jump in Colour Excess and Extinction with Increasing Distance . . . . .                  | 43 |
| 5.5. Geometry for De-Projecting Velocities along the Spine . . . . .                          | 44 |
| 5.6. Power Spectrum . . . . .                                                                 | 46 |
| 5.7. Mock Power Spectrum . . . . .                                                            | 48 |
| A.1. First Section Detail . . . . .                                                           | 67 |
| A.2. Second Section Detail . . . . .                                                          | 68 |
| A.3. Third Section Detail . . . . .                                                           | 69 |
| A.4. Cloud Placement in Galactic Plane . . . . .                                              | 70 |



# 1. Introduction

Most Galactic star formation takes place in molecular clouds (MCs) [Stahler and Palla, 2004]. These structures are not uniform but internally structured, and their sub-structure arises through fragmentation; a process in which gravitational instability causes the cloud to break apart into smaller, denser units [Vázquez-Semadeni et al., 2019]. The fragmentation of molecular clouds into sub-structures and the mass distribution of these sub-structures directly set the initial conditions for star formation. Understanding this process is therefore central to understanding how stars, and in particular high-mass stars, form [Peretto et al., 2013].

Molecular clouds commonly exhibit an internal filamentary structure, as revealed by observations in many different tracers [Arzoumanian et al., 2011, Myers, 2009]. Filaments consist of gas and dust, have higher densities than their surroundings, and appear as dark lanes against the molecular cloud background [André et al., 2010]. They can be observed through molecular line emission, extinction maps at optical and infrared wavelengths, as well as far-infrared and sub-millimeter dust emission [André et al., 2014]. In particular, the *Herschel* survey revealed the widespread presence of filaments and highlighted their central role in the star formation process [André et al., 2010]: dense star-forming cores are found along filament spines, a configuration often described as “beads on a string” (e.g. [Feng et al., 2024]). Filaments are therefore key sites of star formation, yet the exact mechanisms governing their fragmentation and the formation of dense cores remain uncertain [André et al., 2014].

To understand what drives filament fragmentation, we must first consider how a filament becomes gravitationally unstable. In the simplest configuration, a self-gravitating gas cloud is in equilibrium when the internal gas pressure, which acts to disperse the cloud, is balanced by self-gravity, which acts to compress it. This balance defines a maximum stable size: any cloud exceeding this critical diameter, the Jeans length  $\lambda_J$  [Jeans, 1902], can no longer be supported by pressure and will collapse. In this picture, within a larger molecular cloud, the medium naturally breaks apart into characteristic density enhancements separated by  $\lambda_J$ , set by the gas kinetic temperature  $T_K$  and density  $\rho$ . These density enhancements are gravitationally bound and serve as the seeds for the subsequent collapse and formation of stars [Padoan et al., 2001, André et al., 2014]. This description is highly idealized, however, and cannot be applied to filaments without simplifying reality to a large extent. Filaments have a prominent cylindrical geometry, whereas Jeans fragmentation assumes spherical symmetry.

For filaments, a more adequate description is provided by the model of an infinite, self-gravitating cylinder, first derived by Ostriker [1964] and Stodólkiewicz [1963], which predicts a characteristic separation of density enhancements  $\lambda_O$  arising from the geometry of the filament itself. The same physical principle – the balance between gravity and

## 1. Introduction

pressure – is now applied to the cylindrical geometry. An extension of this model to include magnetohydrodynamical (MHD) processes [Chandrasekhar and Fermi, 1953] predicts an additional characteristic separation  $\lambda_S$ , arising from instabilities along the filament axis – commonly referred to as “sausage instability”. These models and their predictions are discussed in more detail in Chapter 2.

Giant Molecular Filaments (GMFs) are among the largest continuous molecular structures in the Galaxy. With lengths ranging from 10 pc to 100 pc, they cover large areas of the sky and can reach total masses of up to  $10^6 M_\odot$ . Their column densities  $N_0$  range from  $6 \times 10^{21} \text{ cm}^{-2}$  to  $8.4 \times 10^{21} \text{ cm}^{-2}$ , with line masses (mass of a filament per length unit) around  $1000 M_\odot \text{ pc}^{-1}$  and aspect ratios between 4 and 100 [Hacar et al., 2023]. GMFs are not uniform structures: their interiors are themselves fragmented, and the densest, coldest regions along their spines can appear as dark absorption features against the bright mid-infrared background of the Galactic plane. These regions are identified as Infrared Dark Clouds (IRDCs) [Rathborne et al., 2016], and are often embedded within the larger GMF [Kainulainen et al., 2011]

IRDCs can be identified in *Spitzer* survey images through prominent mid-infrared absorption features, and due to this selection effect they are predominantly found at distances of several kiloparsecs [Rathborne et al., 2016]. Compared to GMFs, IRDCs are more compact, with lengths of 1 pc to 10 pc, but are characterized by higher column densities ( $N_0 \gtrsim 10 \times 10^{21} \text{ cm}^{-2}$ ), high line masses ( $m \gtrsim 100 M_\odot \text{ pc}^{-1}$ ), and high total masses ( $M \gtrsim 30 M_\odot$ ) [Hacar et al., 2023]. Their large masses, high column densities, and low temperatures make them ideal environments to study the initial conditions of massive ( $>8 M_\odot$ ) star formation [Henshaw et al., 2014]. The most massive IRDCs are proposed to represent the earliest stage of this process [Rathborne et al., 2006], before the cloud has been significantly influenced by feedback from newly formed massive young stellar objects [Henshaw et al., 2016]. Some IRDCs show complex filamentary networks and hub-filament structures within them [Myers, 2009, Busquet et al., 2016]. Within IRDCs, fragmentation produces clumps: gravitationally bound concentrations of gas and dust with typical masses of  $50 M_\odot$  to  $500 M_\odot$ , sizes of 0.3 pc to 3 pc, and densities of  $10^3 \text{ cm}^{-3}$  to  $10^4 \text{ cm}^{-3}$ , that are themselves often filamentary in morphology [Williams et al., 2000]. These clumps in turn contain dense cores – the sites where individual stars ultimately form [Williams et al., 2000]. Cores are characterized by high densities ( $10^4 \text{ cm}^{-3}$  to  $10^5 \text{ cm}^{-3}$ ), typical sizes of  $\sim 0.1$  pc, and masses of  $0.5 M_\odot$  to  $5 M_\odot$  [Bergin and Tafalla, 2007].

The fragmentation process leaves direct kinematic imprints on the velocity structure of filaments. As gas flows toward local gravitational potential minima, coherent velocity gradients develop along the filament spine, and these are often observed in the form of oscillatory patterns [Hacar and Tafalla, 2011]. A structure is considered velocity coherent if its internal velocity variations are small compared to a characteristic velocity scale; the exact criterion depends on the physical context. Such velocity fluctuations in the ISM are ubiquitous [Henshaw et al., 2020], and have recently been reported across a wide range of filamentary structures: in GMFs (CFG024.00+0.48: He et al. [2026]), IRDCs (G28.37+0.07: Beuther et al. [2026], G035.39-00.33: Henshaw et al. [2014]), and HI filaments (“Maggie”: Syed et al. [2022]). One commonly used parameter to characterise

the global gas motion along a filament is the plane-of-the-sky longitudinal velocity gradient  $\nabla v_{LSR}$ . It is defined as the ratio of the change in line-of-sight velocity centroid  $v_{LSR}$  to the filament length  $L$ :  $\nabla v_{LSR} \equiv \Delta v_{LSR} / \Delta L$ . For GMFs, central velocity gradients  $\nabla v_{LSR}$  range from  $0.04 \text{ km s}^{-1} \text{ pc}^{-1}$  to  $0.08 \text{ km s}^{-1} \text{ pc}^{-1}$ , whereas resolved IRDCs show typical values from  $0.20 \text{ km s}^{-1} \text{ pc}^{-1}$  to  $0.90 \text{ km s}^{-1} \text{ pc}^{-1}$ , reflecting the denser, more compact nature of these structures.

In this case study, we investigate the kinematic behaviour of the GMF "Nessie". The object was identified by Jackson et al. [2010] as a velocity coherent structure and classified as an IRDC because of the absorption features in *Spitzer* images. Nessie thus shows properties associated with both filament categories. In the original study by Jackson et al. [2010], Nessie was reported to extend over  $1^\circ$  on the sky, from  $l = 337.85^\circ$  to  $l = 339.1^\circ$  in Galactic ( $l, b$ ) coordinates. At a distance of 3.1 kpc, this corresponds to a projected length of approximately 80 pc, placing Nessie in the category of GMFs. Nessie is one of the best-studied GMFs and, with its exceptional length, an ideal target to study fragmentation simultaneously at clump scales and across tens of parsecs – a dynamic range rarely accessible in a single object. Nessie was observed in August 2008 with the Australia Telescope National Facility 22 m Mopra Telescope in HNC with a spatial resolution of  $15''$ . The HNC(1-0) line emission with a spectral resolution of  $0.11 \text{ km s}^{-1}$  across the entire filament shows the same radial velocity of  $(-38.0 \pm 3.4) \text{ km s}^{-1}$ , demonstrating that the individual, infrared extinction features belong to a single, velocity-coherent cloud, with a  $\pm 3.4 \text{ km s}^{-1}$  scatter around the velocity of the system  $v_{sys} = -38 \text{ km s}^{-1}$  [Jackson et al., 2010].

Later, this structure was named "Nessie Classic" by Goodman et al. [2014]. They found larger extensions with similar line-of-sight velocities in CO [Dame et al., 2001] and in  $\text{NH}_3$  sources in the  $\text{H}_2\text{O}$  Southern Galactic Plane (HOPS) Survey [Purcell et al., 2012]. They defined a moderately extended structure, "Nessie Extended", spanning approximately  $3^\circ$  ( $\sim 160 \text{ pc}$ ), and a more speculative structure, "Nessie Optimistic", which may reach up to  $8^\circ$  ( $\sim 430 \text{ pc}$ ). Due to its large spatial extent and its alignment with the Galactic plane, Nessie was proposed to be a "bone" of the Milky Way [Goodman et al., 2014] – a GMF embedded in the spiral arm potential. Nessie is the prototype of this class of structures [Zucker et al., 2015].

To identify the main driver of fragmentation in Nessie, we analysed the kinematics of HNC and  $\text{N}_2\text{H}^+$ . These two molecular tracers are complementary:  $\text{N}_2\text{H}^+$  traces the denser, star-forming gas component at  $N(\text{H}_2) > 9 \times 10^{21} \text{ cm}^{-2}$  [Bergin and Tafalla, 2007], whereas HNC is abundant over a broader range of conditions ( $N(\text{H}_2) > 3 \times 10^{21} \text{ cm}^{-2}$  [Pety et al., 2017]), tracing more extended gas. Sausage instability has been suggested to play a major role in the fragmentation of Nessie at clump-scale [Jackson et al., 2010, Mattern et al., 2018], but cannot adequately describe the kinematics on the scale of the entire filament. High-resolution observations of the full filament in both tracers allow us to compare their spatial and kinematic structure, and to probe fragmentation across a wide range of physical scales simultaneously.

Resolving the sub-pc filamentary structure of star-forming regions located at kpc distances requires high-resolution interferometric observations. At such distances, recovering

## *1. Introduction*

emission across the full range of spatial scales requires combining interferometric data with single-dish observations [Bonanomi et al., 2024]. The observations presented in this thesis were conducted at the Atacama Large (sub)Millimeter Array (ALMA) (PI: A. Hacar, project ID 2018.1.00697.S). Due to the large spatial extent of Nessie, 9 individual fields were observed and later combined into a mosaic, each observed with the 7m array and the Total Power (TP) array. A brief introduction to interferometry and data combination is given in Chapter 3.



## 2. Fragmentation of Giant Molecular Filaments

### 2.1. Gravitational and Dynamical Fragmentation

Stars form in molecular clouds, and understanding their formation, evolution, and fragmentation is essential to understand the star formation process. In a homogeneous, isothermal gas cloud, the outward-directed pressure force of the gas works against the inward-directed gravitational force. If the forces are in equilibrium, the cloud is stable. If, however, the cloud exceeds a critical size, gravity overcomes the internal pressure and the medium becomes unstable and will collapse. This inward motion of collapsing gas is directly observable via Doppler-shifted spectral lines, in our case through the molecular transitions of  $\text{HNC}(1-0)$  and  $\text{N}_2\text{H}^+(1-0)$ .

Star formation proceeds through the gradual gravitational collapse of MCs into smaller and denser sub-structures, from cloud-scale filaments, to clumps, to cores [Vázquez-Semadeni et al., 2019]. MCs are inherently turbulent [Mac Low and Klessen, 2004], and turbulence plays a dual role in this process. On the scale of the entire cloud or filament, the varying turbulent velocity field generates density enhancements that are dense enough for gravity to overcome gas pressure and initiate local collapse [Padoan et al., 2001]. On smaller scales, within individual clumps, turbulence acts as an additional support against gravitational collapse, much like pressure [Heitsch et al., 2008]. This interplay between gravity and turbulence is often referred to as gravo-turbulent fragmentation [Klessen et al., 2004], and is central to the dynamic model of MC evolution [Schneider et al., 2010].

Within the gravo-turbulent framework, two distinct fragmentation mechanisms can operate. Dynamical fragmentation dominates on large scales, where turbulence shapes the overall cloud structure and generates the density enhancements that seed collapse [Heitsch et al., 2008]. Gravitational fragmentation, in contrast, dominates on smaller scales, particularly within dense clumps and cores where self-gravity becomes the governing force [Thomasson et al., 2024]. In IRDCs, both mechanisms may operate simultaneously, and it remains unclear which process drives the fragmentation of Nessie.

As gravitational fragmentation proceeds, dense regions of gas form at the intersection of convergent flows, and if this compression is strong enough, these regions become self-gravitating and decouple from the surrounding turbulent flow, collapsing independently under their own gravity, and become gravitationally bound [Padoan et al., 2001]. Gravity then takes over locally, driving the collapse of these bound structures into progressively smaller and denser sub-structures – clumps and cores – in which star formation ultimately takes place [Bergin and Tafalla, 2007]. Dense cores in nearby dark clouds are the birth sites of solar-mass stars, and they represent the simplest environments in which the

## 2. Fragmentation of Giant Molecular Filaments

fundamental processes of star formation can be studied [Myers, 1983].

In the limiting case where fragmentation is driven purely by self-gravity, the process can be described by the classical Jeans instability [Jeans, 1902]. In this model, a self-gravitating gas cloud becomes unstable when gravity overcomes the internal gas pressure, defining a critical scale, the Jeans length  $\lambda_J$  [Jeans, 1902]. Any region exceeding this scale will collapse independently, producing a sequence of overdensities with a characteristic separation of order  $\lambda_J$ . A cloud can span many Jeans lengths, allowing multiple such regions to cross the instability threshold simultaneously [Vázquez-Semadeni et al., 2019], each collapsing independently and giving rise to the observed hierarchy of dense structures. This model is, however, highly idealized: it assumes a uniform, isothermal, spherical cloud in perfect equilibrium; conditions that are not observed in real molecular clouds, which are clumpy, turbulent, and far from spherical [Larson, 1981].

To apply the Jeans model to the fragmentation process, one additional assumption has to be made:

1. The medium is non-turbulent; therefore, the velocity dispersion  $\sigma_V$  is equal to the isothermal sound speed  $c_s$  of the gas.

The isothermal sound speed is given by  $c_s = \sqrt{k_B T_K / \mu_{ISM} m_P}$ , with  $k_B$  being Boltzmann's constant and  $m_P$  the proton mass. If the cloud is composed predominantly of  $H_2$  with a minor contribution of He the mean molecular weight  $\mu_{ISM}$  is equal to 2.33 [Hacar et al., 2023]. The mean number density  $n$  is assumed to be  $n = 10^3 \text{ cm}^{-3}$  and the gas kinetic temperature  $T_K$  is assumed to be 20 K, this ensures consistency with Mattern et al. [2018] in our calculations.

The separation  $\lambda_J$  is then linked to the mean mass density  $\rho$  via

$$\lambda_J = c_s \sqrt{\frac{\pi}{G\rho}}, \quad (2.1)$$

where  $G$  is the gravitational constant. Adopting the values above yields a Jeans length of  $\lambda_J = 0.95 \text{ pc}$ .

The Jeans instability is derived under the assumption of spherical symmetry, which is not appropriate for filaments. A more adequate description is provided by the model of an isothermal, self-gravitating cylinder in radial pressure equilibrium, derived by Ostriker [1964] and Stodólkiewicz [1963]. This description introduces the scale height  $H$ , which is the characteristic radial width of the filament over which the density drops from its central value  $\rho_0$ :

$$\rho(r) = \frac{\rho_0}{(1 + (r/H)^2)^2} \quad (2.2)$$

For an isothermal cylinder, the scale height is given by:

$$H = \frac{c_s}{\sqrt{4\pi G\rho}}. \quad (2.3)$$

Adopting the sound speed  $c_s$  and central density  $\rho_0$  from above yields a scale height  $H$  for an Ostriker filament of 0.15 pc. For an isothermal, self-gravitating cylinder, the

## 2.1. Gravitational and Dynamical Fragmentation

characteristic fragmentation scale [Larson, 1985] is:

$$\lambda_O \approx 7.82 \cdot H, \quad (2.4)$$

which corresponds to the expected characteristic spacing between clumps along the filament spine [Inutsuka and Miyama, 1992]. Following the description above, this yields a mean separation of 1.18 pc.

Chandrasekhar and Fermi [1953] introduced an alternative model, which describes the fragmentation of an infinitely long, non-turbulent, self-gravitating cylinder of an incompressible fluid, including MHD processes. This model is commonly referred to as "sausage instability". Because of its cylindrical geometry, it is as well more suited to describe the physical mechanisms that produce cores and clumps in filamentary IRDCs than Jeans' instability, and it has been proposed as the dominant mechanism for clump formation in Nessie [Jackson et al., 2010]. The predicted mean separation  $\lambda_S$  depends now only on the scale-height  $H$  of the filament. In the regime of the filament radius  $R \ll H$  the characteristic spacing between clumps

$$\lambda_S = 22 \cdot H. \quad (2.5)$$

Adopting the values from above yields  $\lambda_S = 3.32$  pc, which is in good agreement with the observational value of 3.5 pc found by Mattern et al. [2018]. This model is likewise idealized. For instance, it assumes an infinitely long cylinder, whereas real filaments have finite lengths, but it serves as a good approximation to reality.

These instabilities describe gravitational collapse and are therefore primarily acting on clump-scales in the context of GMFs. To address dynamical processes on Galactic scales, Toomre [1964] introduced a stability criterion for a cold fluid disk (see also Binney and Tremaine [2008] for a detailed derivation). The dispersion relation below can be used to determine whether the disk is locally stable against axisymmetric perturbations:

$$\omega^2 = \kappa^2 - 2\pi G \Sigma |k|, \quad (2.6)$$

where  $\Sigma$  denotes the stellar surface density, and  $k$  the wavenumber of the perturbation. The ISM is a complex multi-phase medium but for simplification we treat it here as a fluid, where the surface density is approximated by  $\Sigma$ . If  $\omega^2 < 0$ , the perturbation grows exponentially and the disk is unstable. Conversely, if  $\omega^2 > 0$ , the perturbation oscillates and the disk remains stable.

For the special case of  $\omega = 0$ ,  $k$  is equal to the critical wavenumber  $k_{crit}$ . The corresponding wavelength  $\lambda_{crit} \equiv 2\pi/k_{crit}$  is also known as Toomre's length  $\lambda_T$ . All perturbations with wavelength  $\lambda < \lambda_T$  are unstable, with

$$\lambda_T = \frac{4\pi^2 G \Sigma}{\kappa^2}. \quad (2.7)$$

$\Sigma$  for the typical ISM in the solar neighbourhood it is around  $13 \text{ M}_\odot \text{ pc}^{-2}$  [Binney and Tremaine, 2008].

## 2. Fragmentation of Giant Molecular Filaments

$\kappa$  is the epicycle or radial frequency, and for a flat rotation curve, it can be approximated as  $\kappa = \sqrt{2}\Omega$ , where the circular angular frequency  $\Omega$  is given by

$$\Omega = \frac{V_c}{R_{gal}}, \quad (2.8)$$

with  $V_c$  being the circular velocity and  $R_{gal}$  the Galactocentric radius.

Since  $\Omega$  depends on  $R_{gal}$ , both  $\kappa$  and  $\lambda_T$  are a function of Galactocentric distance. The Toomre length sets an upper limit to the fragmentation length scale, and the growth of perturbations with  $\lambda < \lambda_T$  are suppressed due to shear [Li et al., 2017], as the shear force can provide an additional stabilizing force in the case of a differentially rotating disk [Toomre, 1964].

## 2.2. Molecular Tracers of Kinematics in the ISM

Molecular emission from the ISM is a key diagnostic tool for investigating the physical processes in GMFs and filaments in general. Molecular spectra carry information about, for instance, gas kinematics, temperature, and densities. By spectral line fitting, i.e., fitting a Gaussian profile to the spectrum, we can obtain the velocity centroid  $v_{LSR}$  of the gas (the position of the line), the opacity  $\tau$ , the intensity  $I$  (the integral of the line), or total velocity dispersion  $\sigma_{tot}$  (related to the full width at half maximum (FWHM) as  $FWHM = 2\sqrt{2\ln 2}\sigma_{tot}$ ).

A wide range of molecular species have been detected over the past decades, with  $H_2$  and CO being the most abundant. In the densest parts of GMFs, carbon-bearing species typically freeze out onto dust grains [Bergin and Tafalla, 2007]. As a result, these parts then become invisible in CO. Other species, such as HNC or  $N_2H^+$ , remain abundant and therefore trace these higher density regions, though the underlying reasons are more complex and involve a combination of factors such as excitation conditions, chemical abundances, and optical depth effects.

$N_2H^+(1-0)$  (diazanylium) was first observed by [Turner, 1974] and has a rotational transition at a rest frequency of 93.173 764 GHz, corresponding to the  $J = 1 \rightarrow 0$  transition.  $N_2H^+$  can be described as linear molecule, with an inner and outer  $^{14}N$  nucleus, and H bound to the inner N arranged in a straight line ( $N \equiv N - H$ ). The rotational energy levels are characterised by the angular momentum quantum number  $J$ , which is further split by hyperfine interactions due to the nuclear spins of the two nitrogen atoms. In particular, coupling with the outer  $^{14}N$  nucleus produces three  $F_1$  levels within the upper  $J = 1$  state, coupling with the inner nitrogen nucleus further splits these  $F_1$  levels into seven hyperfine levels characterised by the total angular momentum  $F$ . In total, there are 15 allowed hyperfine transitions for the  $J = 1 \rightarrow 0$  transition [Pagani et al., 2009], but only seven lines are observed because there is no hyperfine splitting of the  $J=0$  level, leading to degeneracies in the transitions [Caselli et al., 1995].

$N_2H^+$  is most often observed in dense molecular clouds, where it has proven to be one of the last molecules to freeze out onto dust grains. It can therefore trace the dense ( $n(H_2) \gtrsim 10^5 \text{ cm}^{-3}$ ) central regions of clouds [Bergin and Tafalla, 2007]. The enhanced

abundance of  $\text{N}_2\text{H}^+$  is also linked to CO depletion: CO is a major destroyer of  $\text{N}_2\text{H}^+$  in the gas phase, and its freeze-out onto dust grains reduces the destruction rate of  $\text{N}_2\text{H}^+$  [Aikawa et al., 2001].

HNC(1-0) (hydrogen isocyanide) was first detected by [Snyder and Buhl, 1972], and its  $J = 1 \rightarrow 0$  rotational transition occurs at a rest frequency of 90.663 564 GHz. The upper  $J = 1$  is split into three hyperfine  $F$  levels, while the ground state  $J = 0$  has only one hyperfine level. This results in three allowed hyperfine transitions. There is no further hyperfine splitting because HNC only contains one nitrogen nucleus, and  $^{12}\text{C}$  does not have a nuclear spin that contributes to hyperfine splitting [Faure et al., 2007]. In practice, these three lines blend together due to limited spectral resolution in astronomical observations and intrinsic line broadening. HNC traces the more diffuse part of the molecular gas. It is visible at lower volume densities ( $n(\text{H}_2) \gtrsim 10^3 \text{ cm}^{-3}$  to  $10^4 \text{ cm}^{-3}$  Hacar et al. [2024]) than  $\text{N}_2\text{H}^+$ .



### 3. Interferometry, Data Reduction, and Data Combination

At the distance of Nessie ( $\sim 3.1$  kpc), resolving sub-parsec structures requires angular resolutions of the order of arcseconds. This can be achieved in an interferometer by combining signals from multiple antennas spread over long baselines. However, interferometric arrays are insensitive to emission on angular scales larger than those sampled by their shortest baselines, causing extended, diffuse emission to be filtered out. Single-dish observations must therefore be incorporated to recover this missing flux [Bonanomi et al., 2024]. To obtain a complete picture of the cloud structure across all spatial scales, we combined the observations from two components of the Atacama Large Millimeter/submillimeter Array, namely the 7m-array and the Total Power array, using a technique called *feathering*.

The following provides a brief summary of Chapters 2, 3, and 7 of the ALMA handbook [Cortes et al., 2024]. The purpose is to introduce the basic principles of interferometry, a topic that can be difficult at first encounter. This overview intends to clarify the most important terms and concepts. It should serve as a guide for the reader, especially for myself, should I come back to that topic in the future, when I will inevitably have forgotten everything.

#### 3.1. Atacama Large Millimeter/submillimeter Array

Our observational data were obtained with the Atacama Large Millimeter/submillimeter Array. ALMA is an interferometric array operated by the European Southern Observatory ESO. It is located on the Chajnantor plain of the Chilean Andes, which is known for exceptionally dry and clear sky conditions. Together with the high altitude of 5030 m, the location offers the conditions needed for millimeter and submillimeter observations.

ALMA consists of 66 high-precision antennas. 50 of these antennas are antennas with 12 m diameter that form the 12-m Array, which is used for sensitive, high-resolution imaging. These are complemented by the Atacama Compact Array (ACA), also known as the Morita Array. The ACA comprises twelve closely spaced 7 m antennas (the 7-m Array), and four 12 m antennas that operate as the Total Power Array for single-dish (SD) observations. The TP arrays are used to recover the emission at the most extended spatial scales. In this work, we use observational data from the 7m and TP array only.

### 3.2. Interferometry in a Nutshell

An interferometer is a telescope that combines the signals from two or more physically separated antennas using interferometry. By correlating these signals, the array achieves an effective aperture corresponding to the maximum separation  $L_{max}$  between the antennas rather than the diameter of an individual dish. The technique is therefore often referred to as *aperture synthesis*. This results in a significantly higher angular resolution  $\theta_{res}$ , depending on the observed wavelength:

$$\theta_{res} = \frac{k\lambda}{L_{max}}, \quad (3.1)$$

where  $k$  is a weighting factor, typically  $\sim 1$ .

The measured signals must then be processed to reconstruct an image of the sky, a step commonly referred to as data reduction. In CASA, see Sec. 3.3.1, this imaging process can be performed with the task `tclean` [CASA Team et al., 2022].

The physical separation  $b$  between two antennas 1 and 2 can be expressed in units of the observing wavelength  $\lambda$ ; the quantity  $L = b/\lambda$  is known as the baseline. In Fig. 3.1, the two antennas observe a common position  $s_0$ , located at an angle  $\theta$  from the zenith. From the perspective of the source, the projected separation of the two antennas is  $u = b \cos \theta$ . The incoming wavefront reaches telescope 2 a little later than telescope 1 because it travels a small additional path of length  $b \cdot s_0 = b \sin \theta$ . The role of the correlator is to compensate for this geometrical delay and to bring the two signals in phase before they are combined.

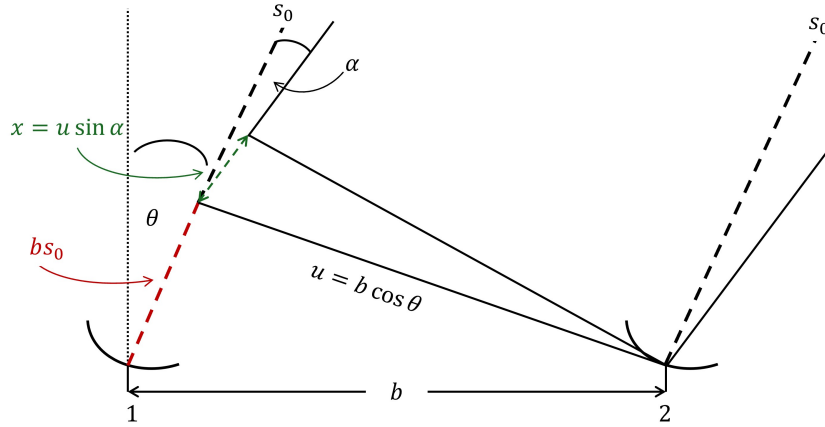


Figure 3.1.: Antennas 1 and 2 observe a common position  $s_0$  at the sky. This introduces a geometrical separation  $u$  and an extra path length  $x$  for the incoming wave front to travel. Adapted from Cortes et al. [2024].

For a source located at an angle  $\alpha$  away from the phase center, the signal reaching antenna 1 travels a slightly longer path than the signal reaching antenna 2, even after



applying the geometrical delay that compensates for an on-axis signal. This extra path length is  $x = u \sin \alpha = ul$ , where  $l = \sin \alpha$ . All distances are expressed in units of the observing wavelength  $\lambda$ , therefore  $x$  is simply the number of wavelengths. Expanding this description to two dimensions introduces a second angular offset  $\beta$ , as the equivalent of  $\alpha$ , and  $m$ , as the equivalent of  $l$ , where the associated path difference in direction is  $y = v \sin \beta = vm$ .

The signal received by the interferometer can be described as a sinusoidal response, and it is proportional to the sky brightness. By sampling many such signals at different baselines, the interferometer effectively measures the sky brightness distribution. Each individual measurement is called a *visibility*. A visibility contains information about the brightness of the emission on the angular scale sampled (related to the amplitude of the sinusoid) and the relative position of that source on the sky (related to the phase of the sinusoid). The completed set of recorded visibilities is called a measurement set.

The correlator measures a quantity known as the complex visibility  $\mathcal{V}$ , which is formally the Fourier transform of the sky brightness distribution, i.e, the image  $I(l, m)$ :

$$\mathcal{V}(u, v) = \int \int I(l, m) e^{2\pi i(ul+vm)} dl dm, \quad (3.2)$$

where  $u$  and  $v$  are identified as spatial frequency components of the sinusoid in E-W and N-S directions, respectively. This relation is known as the van Cittert-Zernike theorem, and it is the basis of aperture synthesis. It implies that the sky brightness distribution  $I$  can be recovered, in principle, by taking the inverse Fourier transform of the complex visibility distribution  $\mathcal{V}(u, v)$ .

The visibility, or  $(u, v)$ -plane, is sampled by observing the sky. By measuring the distribution of complex visibilities at different baselines, we can recover the sky brightness distribution. In practice, however, only a finite number of measurements can be obtained, which leads to discrete, irregular, and incomplete sampling of the  $(u, v)$ -plane. This incomplete coverage can be described by a sampling function  $B(u, v)$ . In the image domain, this corresponds to a convolution of the true sky brightness distribution with the point spread function (PSF)  $b(l, m)$ . The PSF is given by the inverse Fourier transform of the  $(u, v)$ -sampling function  $b(l, m) = FT^{-1}\{B(u, v)\}$ .  $b(l, m)$  is sometimes also called the dirty beam or synthesized beam. The dirty image  $I_V^D$  is therefore given – for now – as the convolution of the visibilities  $\mathcal{V}(u, v)$  with the dirty beam:

$$I_V^D(l, m) = b(l, m) \times I(l, m) = FT^{-1}\{B(u, v)\mathcal{V}(u, v)\}, \quad (3.3)$$

and it can be retrieved as well using the inverse Fourier transform.

If the  $(u, v)$ -plane were sampled perfectly and completely, the image created by inverse Fourier transforming the  $(u, v)$ -plane would be an exact representation of the true sky brightness distribution. In practice, however, it is impossible to sample the  $(u, v)$ -plane completely and obtain all visibilities. One fundamental limitation is the "short-spacing" problem.

This sets a lower limit on the shortest achievable baseline: the origin of the  $(u, v)$ -plane, corresponding to zero spacing, is never sampled, and baselines shorter than the antenna

### 3. Interferometry, Data Reduction, and Data Combination

diameter, the so-called short spacings, are inaccessible. In the  $(u, v)$ -plane, this appears as a central gap, compare Fig. 4.2, known as the "short-spacing" problem. This limitation has a considerable impact on observations of extended objects, such as molecular clouds like Nessie, because large-scale emission is filtered out by the interferometer, as given by the maximum recoverable scale  $\theta_{MRS}$ :

$$\theta_{MRS} \approx \frac{0.6\lambda}{L_{min}}, \quad (3.4)$$

where  $L_{min}$  is the minimum baseline in the array configuration. Missing short spacings correspond to a lack of sensitivity to large-scale emission, meaning that an interferometer alone cannot fully recover extended flux. The resulting images may therefore underestimate the total flux or miss smooth, extended structures altogether. Physically, this reflects a key principle of interferometry: long baselines resolve small, fine structures, whereas short baselines are sensitive to large-scale, smooth emission and produce a lower-resolution image.

This is why the Total Power array exists in ALMA. The TP antennas operate as SD telescope and provide measurements at the shortest baselines in the  $(u, v)$ -plane. In this way, they recover the large-scale flux and the absolute intensity that cannot be measured by interferometric baselines alone. When TP data is combined with interferometric data, significantly more flux is recovered, especially for extended sources such as molecular clouds. Only by combining our TP and 7-m data, ALMA can fully capture emission across all relevant spatial scales. We use one technique that achieves this data combination called *feathering*, which is explained in Sec. 3.3.

The sky brightness distribution is modified by the antenna power response  $A(l, m)$ , often referred to as the primary beam. What an interferometer therefore actually measures is the Fourier transform of the sky brightness distribution multiplied by the antenna power response  $\mathcal{A}(l, m)I(l, m)$ . This can be written as:

$$\mathcal{V}(u, v) = \iint \mathcal{A}(l, m)I(l, m)e^{-2\pi i(ul+vm)}dldm, \quad (3.5)$$

compare with Eq. 3.2.

Taking both effects into account – the primary beam response in Eq. 3.5 and the incomplete  $(u, v)$ -sampling in Eq. 3.3 – the resulting dirty image  $I^D(l, m)$  can be written as the convolution of the PSF with the sky brightness distribution modified by the antenna power response:

$$I^D(l, m) = b(l, m) * I(l, m)A(l, m) = FT^{-1}\{B(u, v)\mathcal{V}(u, v)\}, \quad (3.6)$$

It is the task of **tclean** to recover the true sky brightness distribution  $I(l, m)$  as best as possible from this dirty image. The imaging process, therefore, consists of two steps: 1) creating the dirty image and then 2) cleaning the dirty image, i.e., deconvolving the dirty image to remove the effects of the dirty beam.

**tclean** handles both of these steps. **tclean** is an imaging algorithm that forms images and spectral line cubes from visibilities. It is based on the CLEAN algorithm,

introduced by Högbom [1974], which is the most popular and widely-studied method for reconstructing a model image based on interferometric data. First, `tclean` creates the dirty image using the configuration the user defined, including different parameters such as weighting schemes for the visibilities. In the subsequent deconvolution step, a model of the sky brightness distribution is constructed iteratively from  $I^D(l, m)$  and  $b(l, m)$ . This model is subtracted from  $I^D(l, m)$ , and the procedure is repeated until the residuals reach a threshold [Petty et al., 2020]. This process is called deconvolution, because formally Eq. 3.6 is being inverted to solve for  $I(l, m)$ .

When creating the dirty beam, the incomplete sampling of the  $(u, v)$ -plane is directly reflected in the PSF. Gaps in the  $(u, v)$ -plane lead to sidelobes in the dirty beam. The strength and structure of these sidelobes depend on the extents and locations of the gaps in the  $(u, v)$ -plane, as well as on the brightness of emission on the sampled spatial scales. Since the sky brightness is convolved with the PSF during each iteration in `tclean`, these sidelobes redistribute flux across the image: at every identified emission peak, a scaled version of the PSF is applied. As a result, the reconstructed image can have significant artifacts that emerge only because of the data reduction process and do not correspond to real physical features in the true sky brightness distribution.

In the final step of image processing, the image has to be divided by the antenna power response  $A(l, m)$ . This step is called primary beam correction, and when feathering, it should be performed before the images are combined.

## 3.3. Data Combination

### 3.3.1. The Common Astronomy Software Applications Package

The Common Astronomy Software Applications Package CASA [CASA Team et al., 2022] is the main software package in astronomy to reduce interferometric data, in particular ALMA data. In this work, we used CASA to do both data reduction and data combination. CASA can process data from both single-dish and interferometric telescopes, which is particularly important for this work. It also supports image combination. Here, the high-resolution data were combined with the low-resolution data using the `feather` method, within the Model Assisted Clean + Feather method.

CASA provides a comprehensive framework for all major steps of radio data analysis, including data inspection, flagging, calibration, imaging, and visualization. Its task-based structure allows users to apply well-established interferometric techniques in a consistent and reproducible manner, while also allowing workflow customization through Python scripting. CASA is developed and maintained by an international collaboration of institutes involved in major radio observatories, including the National Radio Astronomy Observatory (NRAO), the European Southern Observatory (ESO), and the National Astronomical Observatory of Japan (NAOJ).

We used CASA version 6.5 on the server *licancabur* of the Institute for Astrophysics. Each cleaning process takes between 30 min to 90 min, depending on the settings, such as the number of iterations or the noise threshold.

#### 3.3.2. Model Assisted CLEAN + Feather

Starting from the dirty image, `tclean` identifies the location of the strongest emission, assumes it represents the true sky signal, and subtracts a scaled version of the PSF centered at that position. At each iteration, it removes a fraction of the flux in the brightest pixel within a defined region of the current “dirty” image and adds this to the model image. Each iteration updates the model image that progressively approximates the true sky brightness distribution, while the residual image contains the remaining, not yet deconvolved signal and noise.

One can help `tclean` find the brightest pixels by providing a model at the start of the cleaning process. In this work, the SD image (in units of Jy/px) is used as the starting model image. The image is then invoked via the parameter `startmodel`. Using the SD image as a starting model has the advantage that the extended emission is already provided in the model image, before interferometer-based clean components are added. The short-spacing flux, even though initialized with the use of the SD image, is not constrained by `tclean`, so the result after model assisted clean should be considered as an intermediate image only. The next step is then to combine this intermediate image with the SD image using `feather`. This combined method is therefore called Model Assisted CLEAN + Feather (MACF).

The model image must be in units Jy/px, and it must also match the spectral resolution of the interferometric image. This can be achieved by regridding the SD image to the interferometric grid with `imregrid`. It must also have the same spectral resolution, which can be achieved with `imreframe`. In addition, the model must have the same axis order (position, position, velocity), which can be done with the CASA task `imrtrans`. For the conversion to Jy/px we used the scripts<sup>1</sup> by Plunkett et al. [2023].

In general, feathering is a technique used to combine a SD image (or, in our case, a TP image) with an interferometric image (in our case, 7m) of the same field. The goal of this process is to reconstruct the source emission across all spatial scales, from the small spatial scales measured by the interferometer to the large-scale structure measured by the single-dish. Bonanomi et al. [2024] show that data combination is crucial when working with ALMA data, since otherwise key parameters like angular resolution or flux recovery may not meet the required specifications.

To do this, `feather` forms a weighted sum of the Fourier transforms of TP and the 7m image. The input images or cubes are first Fourier transformed to the spatial frequency domain. The single-dish data are down-weighted at high spatial frequencies, while the interferometric data are down-weighted at low spatial frequencies. The weighted Fourier components are then added together and inverse Fourier transformed to obtain a combined image or cube. Feathering two images effectively improves the  $(u, v)$ -coverage by filling up the central gap in the  $(u, v)$ -plane with the TP data.

The quantity measured with an interferometer is the flux density  $S_\nu$  and is measured

---

<sup>1</sup><https://github.com/teuben/DataComb>

in Jy, which can be converted to brightness temperature  $T$  [Cortes et al., 2024] with:

$$\left(\frac{T}{1 \text{ K}}\right) = \left(\frac{S_\nu}{1 \text{ Jy}}\right) \left[13.6 \left(\frac{300 \text{ GHz}}{\nu}\right)^2 \left(\frac{1''}{\theta_{maj}}\right) \left(\frac{1''}{\theta_{min}}\right)\right], \quad (3.7)$$

where  $\nu$  is the observing frequency, and  $\theta_{max}$  and  $\theta_{min}$  are the major and minor axes of the synthesized beam.

#### 3.3.3. Image Quality Assessment

To assess the quality of the cleaned image  $I$  created by `tclean`, the resulting map has to be compared to a reference  $R$ . Since the true sky brightness distribution is unknown, the TP image is used as an approximation. For the quality assessment, we adopt the "accuracy parameter"  $A_\nu(x, y)$  [Plunkett et al., 2023]:

$$A_\nu(x, y) = \frac{I_\nu(x, y) - R_\nu(x, y)}{|R_\nu(x, y)|} \quad (3.8)$$

where  $-\infty < A_\nu(x, y) < \infty$ . A-par quantifies the relative pixel-by-pixel difference in flux between the input image  $I_\nu(x, y)$  and reference image  $R_\nu(x, y)$  at each frequency  $\nu$ .

The key advantage of A-par over other image quality metrics, such as the "fidelity parameter", is that it represents the signed relative error of the recovered emission with respect to the reference image. The sign indicates whether the emission is underestimated ( $A_\nu < 0$ ) or overestimated ( $A_\nu > 0$ ). A perfect match between the input  $I_\nu(x, y)$  and reference  $R_\nu(x, y)$  images thus corresponds to minimizing A-par in all pixels, resulting in a distribution of  $A_\nu$  centered around 0.



## 4. Observational Data, Set-Up in CASA, and Resulting Maps

### 4.1. Observations

Nessie was observed during ALMA Cycle 6 in Band 3 in 2019 (project ID: 2018.1.00697.S; PI: Hacar). As Nessie spans over  $2.5^\circ$  in the sky, the observations were divided into nine separate fields to cover the spine of the filament, where in previous observations extended HNC emission was found in some of the fields. The nine individual fields are shown in Fig. 4.1. The fields are not labeled in ascending or descending order; additionally, in the project, F9 was inadvertently called F1.

All fields were observed with the 7m array and the TP array. For F1, F3, F6, and F9, each comprises two 7m visibility sets; for the remaining fields, there is just one. The individual fields can later be combined through mosaicking. The calibrated 7m visibility datasets were provided via the ALMA Helpdesk. The TP data were downloaded from the ALMA archive. The observational details for all fields can be found in Tab. 4.1.

Each field was observed in five molecular tracers; however, only  $\text{N}_2\text{H}^+$  and HNC were used for this analysis.  $\text{N}_2\text{H}^+$  was observed with a spectral resolution of 30.5 kHz or  $0.098 \text{ km s}^{-1}$ , and HNC was observed with 61.0 kHz or  $0.202 \text{ km s}^{-1}$ . The spectral resolution in frequency  $\delta\nu$  can be converted to a velocity resolution  $\delta v$  in  $\text{km s}^{-1}$  by the simple conversion

$$\delta v = c \cdot \frac{\delta\nu}{\nu_0}, \quad (4.1)$$

where  $\nu_0$  is the rest frequency and  $c$  is the speed of light.

Before cleaning or combining the data, we had to subtract the continuum in every calibrated visibility set with the CASA command `uvcontsub`. For fields with two visibility sets (F1, F3, F6, F9), the data were concatenated with `concat`. The TP data were provided as a `.fits` file, and in order to process them in CASA, we had to convert it to `.image` format using `importfits`.

### 4.2. MACF

#### 4.2.1. Issues in (u,v)-Plane

As described in Sec. 3.2, the PSF corresponds to the inverse Fourier transform of the  $(u, v)$ -sampling. Incomplete sampling of the  $(u, v)$ -plane leads to sidelobes in the PSF, with larger gaps between sampled tracks producing more prominent sidelobes. During deconvolution, flux from real emission is redistributed into these sidelobes. In the case

#### 4. Observational Data, Set-Up in CASA, and Resulting Maps

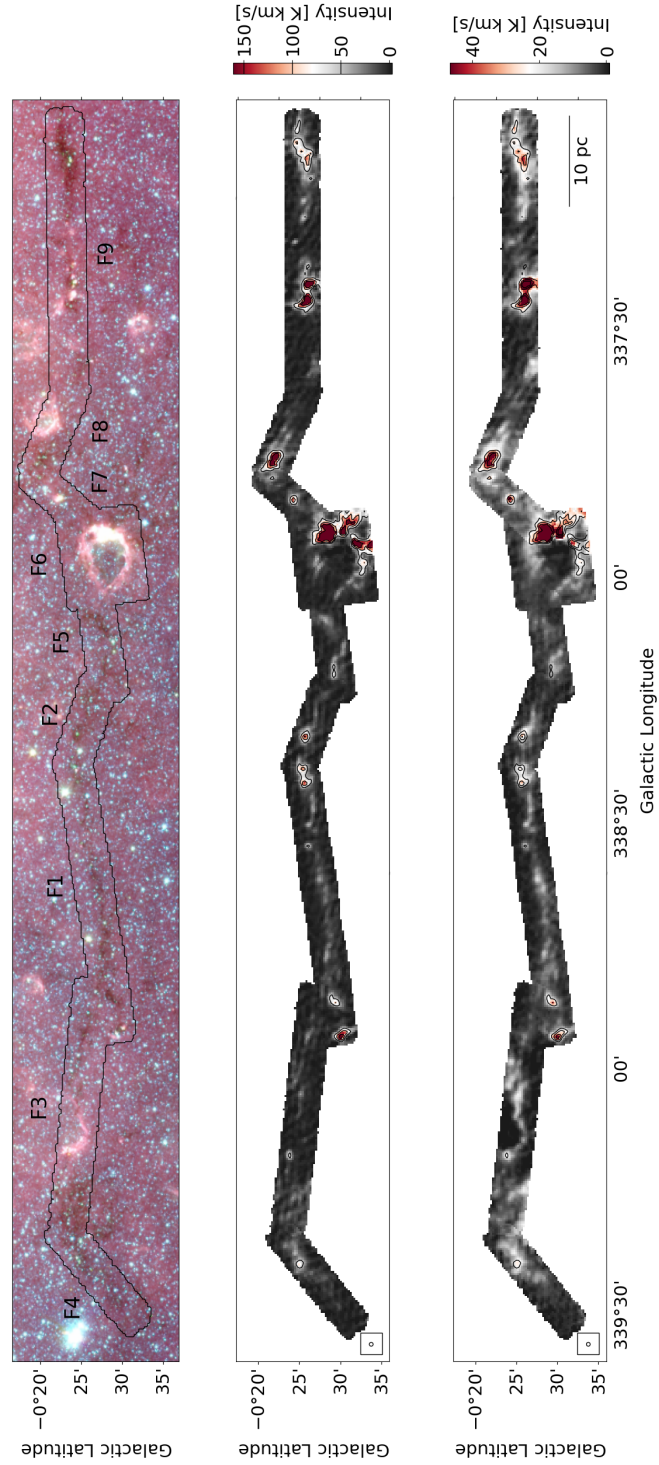


Figure 4.1.: *Left*: False-colour *Spitzer* image, with red at  $8\mu\text{m}$ , green at  $4.5\mu\text{m}$ , and blue at  $3.6\mu\text{m}$ . The individual fields are outlined by the  $\text{N}_2\text{H}^+$  footprints. *Middle*:  $\text{N}_2\text{H}^+$  integrated intensity map, with contour levels at 50, 100, and  $150\text{ K km s}^{-1}$ . *Right*: HNC integrated intensity map with the same contour levels as the  $\text{N}_2\text{H}^+$  map.



| Field | Sky Position                                    |                        | Galactic Coords. |           | Map size               | No. of Pointings |
|-------|-------------------------------------------------|------------------------|------------------|-----------|------------------------|------------------|
|       | $\alpha(\text{J2000})$                          | $\delta(\text{J2000})$ | Gal. long.       | Gal. lat. | [arcsec <sup>2</sup> ] |                  |
| F1    | 16 <sup>h</sup> 43 <sup>m</sup> 57 <sup>s</sup> | −46°32′47″             | 338.67°          | −0.45°    | 2100″×300″             | 147              |
| F2    | 16 <sup>h</sup> 42 <sup>m</sup> 28 <sup>s</sup> | −46°48′14″             | 338.31°          | −0.42°    | 900″×300″              | 55               |
| F3    | 16 <sup>h</sup> 45 <sup>m</sup> 22 <sup>s</sup> | −46°12′41″             | 339.08°          | −0.41°    | 2100″×300″             | 131              |
| F4    | 16 <sup>h</sup> 46 <sup>m</sup> 50 <sup>s</sup> | −45°58′23″             | 339.43°          | −0.45°    | 1200″×300″             | 67               |
| F5    | 16 <sup>h</sup> 41 <sup>m</sup> 59 <sup>s</sup> | −46°57′2″              | 338.15°          | −0.46°    | 900″×300″              | 55               |
| F6    | 16 <sup>h</sup> 41 <sup>m</sup> 17 <sup>s</sup> | −47°5′8″               | 337.97°          | −0.46°    | 1000″×600″             | 127              |
| F7    | 16 <sup>h</sup> 40 <sup>m</sup> 29 <sup>s</sup> | −47°6′54″              | 337.85°          | −0.38°    | 900″×300″              | 49               |
| F8    | 16 <sup>h</sup> 39 <sup>m</sup> 51 <sup>s</sup> | −47°12′58″             | 337.71°          | −0.36°    | 900″×300″              | 55               |
| F9    | 16 <sup>h</sup> 38 <sup>m</sup> 31 <sup>s</sup> | −47°30′15″             | 337.34°          | −0.39°    | 2100″×300″             | 147              |

Table 4.1.: Observational details for each field in Nessie. The listed coordinates are the center of the observed field. The map size helps in choosing the `imsize` parameter for the cleaning process of the respective field.

of large sidelobes, this redistribution can generate significant artifacts in the resulting map, i.e., features that do not represent real emission but are introduced in the cleaning process. This is tentatively visible in Fig. 4.1 as structures parallel to extended regions of intermediate emission, though their nature remains ambiguous. A definitive identification would require comparing their brightness relative to the peak emission against the first-sidelobe response of the PSF, as well as inspecting their coherence across velocity channels in the spectral cube.

To avoid that issue, the data can be tapered in the  $(u, v)$ -plane, and essentially exclude the outer baselines. In our case, this corresponds to baselines at 30 m, or at  $10\text{ k}\lambda^1$ . At this baseline length, the sampled tracks exhibit large gaps, as can be seen in Fig. 4.2a, which requires `tclean` to interpolate more extensively, which in turn enhances deconvolution artifacts. Those baselines are thus problematic, and tapering excludes those effectively. The plane is sampled inhomogeneously, and in Fig. 4.2a the  $(u, v)$ -coverage is almost oval. This means, the outermost baselines can be excluded with only minor loss of information, but still sacrificing resolution.

Fig. 4.2a further indicates that the observation strategy was not ideal, as some tracks in the  $(u, v)$ -plane overlap. This implies that observing time was not used ideally: when overlapping regions were sampled, alternative  $(u, v)$ -spacings could have been covered instead. Observing the target at different hour angles would have allowed additional regions of the  $(u, v)$ -plane to be filled, resulting in a more homogeneous coverage. Note the big gap in the center of the  $(u, v)$ -plane is filled by the TP observations.

This assessment is also reflected in the 2D-filling fraction in Fig. 4.2b, which was created with the `assess_ms2` tool [Petry et al., 2020]. The filling fraction is defined as the ratio of observed and expected visibility seconds in an individual baseline bin [Petry et al.,

<sup>1</sup> $\lambda = c/\nu = c/93.173\,764\text{ GHz} \approx 0.3\text{ cm} \Rightarrow 10000 \cdot \lambda = 30\text{ m} \simeq 10\text{ k}\lambda$

<sup>2</sup><https://zenodo.org/records/17293466>

#### 4. Observational Data, Set-Up in CASA, and Resulting Maps

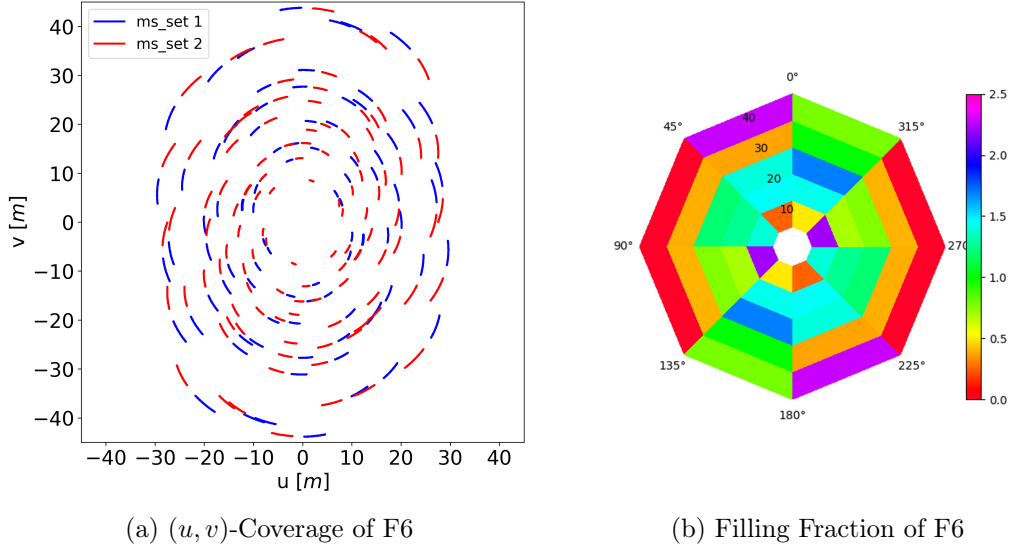


Figure 4.2.: *Left:* The  $(u, v)$ -coverage of F6 is problematic. Some tracks overlap, indicating that the observing time was not used optimally, while large gaps between tracks increase the need for interpolation. *Right:* This is also reflected in the 2D filling fraction plot, where especially the outer left and right parts were sparsely sampled (red). The oval shape of the  $(u, v)$ -coverage also shows up in the filling fraction plot, as the lower and upper parts of the plane were exposed as expected (green), or overexposed (purple).

2024], and it is calculated in each<sup>3</sup> 45° azimuthal sector. For an optimal observation, this value should be around 1 for all sectors, which means that the  $(u, v)$ -plane was sampled as prescribed. If the segment is red or orange, the corresponding part in the  $(u, v)$ -plane was undersampled. Green and blue sectors correspond to appropriate sampling and good sensitivity. Purple regions indicate oversampling, i.e., more integration time than expected in those sectors.

Almost half of the segments are either over- or undersampled, particularly at the shortest and longest baselines. At baselines larger than 30 k $\lambda$  we can see the sectors on the left and right in red and orange. This connects to Fig. 4.2a, as there are no tracks in the left and right outermost parts of the  $(u, v)$ -plane. It creates a PSF with large sidelobes, which in turn produces large artifacts in the cleaned map.

The image quality of the maps produced from the tapered measurement sets are similar to those generated using the full set of baselines. This can be seen in Tab. 4.2, where the image quality of the resulting maps is compared using A-Par (see Sec. 3.3.3) and the total flux recovered (FR). We compare maps generated using MACF (detailed description see Sec. 4.2.2), feather, and the 7m data alone, each produced from both tapered and

<sup>3</sup>Only half of the sectors, from 0° to 180°, actually need to be calculated. The rest is a mirror copy of the first half, shown for convenience.

untapered measurement sets.

The comparison of tapered and untapered data shows that tapering leads to virtually no loss of flux. Using all baselines, MACF achieves a FR of 99.2 %, and only feather has a FR of 89.9 %. If only baselines  $b < 10k\lambda$  are used in the cleaning process, FR stays virtually unchanged at 99.1 % and 89.7 %, respectively. Tapering therefore primarily affects spatial resolution rather than flux recovery in this case. The 7m-only case presents a different result: the untapered map recovers 20.8 % of the flux, compared to 17.6 % for the tapered case, a difference of  $\sim 3$  %. Regardless, both values highlight the necessity of data combination, as the 7m-only maps recover almost no flux without the short-spacing information provided by the total-power data.

On the other hand, the advantage of MACF is evident, as we can recover up to 10 % more flux when invoking a model. These results highlight both, the necessity of data combination and the effectiveness of MACF as a combination method.

| F6                  | parameter | MACF              | only feather       | only 7m            |
|---------------------|-----------|-------------------|--------------------|--------------------|
| all baselines       | A-Par     | $0.024 \pm 0.115$ | $-0.067 \pm 0.166$ | $-0.818 \pm 0.155$ |
|                     | FR        | 99.2 %            | 89.9 %             | 20.8 %             |
| taper 10k $\lambda$ | A-Par     | $0.026 \pm 0.126$ | $-0.660 \pm 0.175$ | $-0.842 \pm 0.181$ |
|                     | FR        | 99.1 %            | 89.7 %             | 17.6 %             |

Table 4.2.: The best results are obtained when using MACF. In both configurations, all baselines and taper, we can recover almost 100 % of the flux from the TP map when invoking a model. Feathering alone still yields good results, but MACF is preferable. In all cases, the need for data combination is evident, as using only 7m data cannot recover enough flux.

#### 4.2.2. MACF Set-Up

To create the clean image of each field, we used the MACF method, as described in Sec. 3.3. This yields the best result in terms of spatial resolution and flux recovery, see Tab. 4.2. The spatial resolution of the images with all the baselines is around  $16''$ , for the case of tapering at  $10k\lambda$  the resolution degrades by about 6 % to  $17''$ . For an overview of how we implemented the MACF workflow, see Fig. 4.3.

For the model assisted clean, the SD image in Jy/px is used as model. This model can be created from the TP observations of Nessie. The TP data must have the same number of pixels in the spatial plane, the same spectral resolution, and the same axis ordering as the 7m data. To achieve this, we used the CASA methods `imregrid`, `imreframe`, and `imtrans` respectively, and used the 7m image as a template. This aligned TP image can then be converted from units of Jy/beam to Jy/px and subsequently used as the model to clean the 7m data.

The `tclean` task provides several parameters that control the imaging process. An extensive testing phase showed that the following configuration leads to optimal results in terms of a narrow A-par distribution and a high FR:

#### 4. Observational Data, Set-Up in CASA, and Resulting Maps

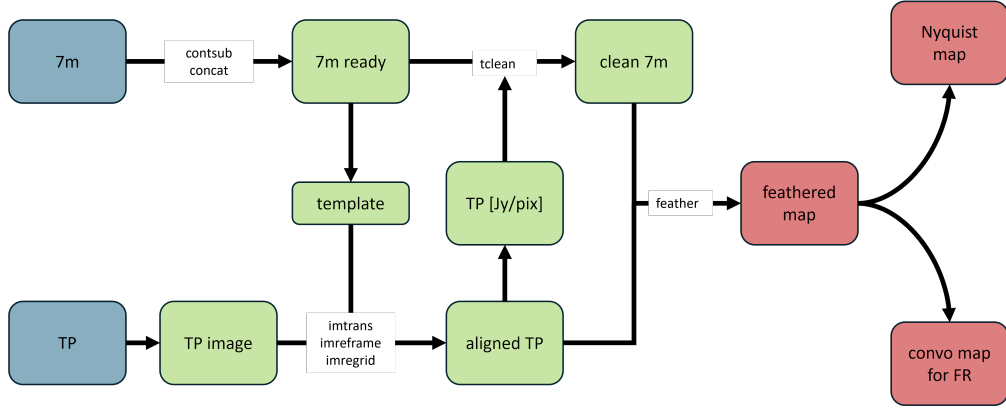


Figure 4.3.: The two observations from the 7m and TP array must be cleaned using `tclean`. When using MACF, the TP data has to be prepared accordingly, e.g., the units must be changed from Jy/beam to Jy/px. The resulting map can then be convolved to the TP beam for flux recovery assessments, or Nyquist sampled and used for spectral line fitting.

- **imsize**: The image size in number of pixels. Large images need a larger number of pixels. The number of pixels must be even and factorizable by 2, 3, and 5 only<sup>4</sup>, to ensure an optimal performance of `tclean`. This parameter was set individually for each visibility.
- **cell**: The pixel size, in arcsecond. It must oversample the synthesized beam; typically, 4 - 5 pixels per beam. The cell size was 3'' for all fields.
- **threshold**: The stopping threshold is in Jy ( $= 10^{-26} \text{ W Hz}^{-1} \text{ m}^{-2} = 10^{-23} \text{ erg s}^{-1} \text{ Hz}^{-1} \text{ cm}^{-2}$ ). A threshold of 0.5 Jy was sufficient for our case, which is twice the residual background noise.
- **niter**: The maximum number of cleaning iterations, if the threshold is not reached earlier. To allow the algorithm to reach the threshold of 0.5 Jy, we set **niter** = 100000.
- **specmode**: Since our data are spectral cubes, we used `cube`. This requires several additional parameters that are specific to the target:
  - **nchan**: Number of output channels  $n$ , to fully cover all the emission peaks in the spectrum. We used 400 for  $\text{N}_2\text{H}^+$  and 200 for HNC to have enough line-free channels to estimate the background later in the fitting procedure. As the channel width for HNC is almost double than the channel width from

<sup>4</sup>See CASA Docs for details: <https://casadocs.readthedocs.io/en/stable/api/tt/casatasks.imaging.tclean.html#imsize>

$\text{N}_2\text{H}^+$ , the number of channels can be reduced by a factor of 2, and still cover the same range of the spectrum.

- **width**: Channel width in  $\text{km s}^{-1}$ , corresponding to the spectral resolution  $\delta v$ . For  $\text{N}_2\text{H}^+$  it is  $0.098 \text{ km s}^{-1}$ , for HNC  $0.202 \text{ km s}^{-1}$ , almost double.
- **restfreq**: The rest frequency of the observed spectral line. For  $\text{N}_2\text{H}^+$  it is  $93.173\,764 \text{ GHz}$ , for HNC  $90.663\,564 \text{ GHz}$  [Hacar et al., 2024].
- **start**: Velocity of the first channel in  $\text{km s}^{-1}$ , defined as  $v_{\text{start}} = v_{\text{sys}} - \frac{n \cdot \delta v}{2}$ , where  $v_{\text{sys}} = -38 \text{ km s}^{-1}$  is the velocity of the observed system with respect to the local standard of rest.
- **gridder**: `mosaic`, which performs convolutional resampling of the visibilities onto a regular  $(u, v)$ -grid and is appropriate for mosaicked observations.
- **deconvolver**: `multiscale` is recommended for extended objects.
- **weighting**: `briggs` weighting and `robust` = 0.5. For most standard imaging `robust` = 0.5 is recommended, as it provides a trade-off between angular resolution and sensitivity [CASA Team et al., 2022].
- **usemask** = `auto-multithresh` with:<sup>5</sup>
  - `sidelobethreshold` = 1.25,
  - `noisethreshold` = 5.0,
  - `lownoisethreshold` = 2,
  - `negativethreshold` = 0.0,
  - `psfcutoff` = 0.25.

The resulting 7m image was primary-beam corrected using the corresponding `.pb` file that is created automatically by `tclean`. The 7m image was then combined with the TP image using the `feather` task. The final MACF maps achieve sensitivities between  $0.2 \text{ Jy/beam}$  to  $0.4 \text{ Jy/beam}$  at the native beam, which are brightness temperatures of  $0.036 \text{ K}$  to  $0.072 \text{ K}$  for  $\text{N}_2\text{H}^+$  and  $0.038 \text{ K}$  to  $0.076 \text{ K}$  for HNC, applying Eq. 3.7.

#### 4.2.3. Quality Assessment after MACF

To quantify the quality of the maps after the feathering process, we calculated the recovered flux with A-par, see Sec. 3.3.3. A-par compares the input image  $I$  with the reference image  $R$  at positions  $(x, y)$  using the relation shown in Eq. 3.8, thus quantifying the flux recovery. In this case, the MACF map is the input image and the TP map is the reference image [Plunkett et al., 2023]. Note that by this proceeding we implicitly assume that the TP map represents the true sky brightness distribution.

<sup>5</sup>In [https://casaguides.nrao.edu/index.php/Automasking\\_Guide\\_CASA\\_6.5.4](https://casaguides.nrao.edu/index.php/Automasking_Guide_CASA_6.5.4) the values for several auto-multithresh sub-parameters can be found. We adopted the values recommended for extended emission at millimeter wavelengths, as appropriate for the large angular extent of Nessie.

#### 4. Observational Data, Set-Up in CASA, and Resulting Maps

| Parameter      | F1    | F2    | F3    | F4           | F5    | F6    | F7    | F8    | F9     |
|----------------|-------|-------|-------|--------------|-------|-------|-------|-------|--------|
| A-par          | 0.031 | 0.033 | 0.010 | -0.016       | 0.009 | 0.024 | 0.096 | 0.013 | -0.001 |
| $\pm$          | 0.116 | 0.120 | 0.144 | 0.087        | 0.102 | 0.115 | 0.167 | 0.095 | 0.115  |
| FR [%]         | 101.8 | 98.8  | 90.6  | 95.1         | 99.7  | 99.2  | 103.1 | 100.4 | 99.8   |
| $\theta_{maj}$ | 17.39 | 19.49 | 16.88 | <b>23.29</b> | 20.06 | 16.27 | 19.61 | 17.50 | 18.60  |
| $\theta_{min}$ | 11.46 | 10.77 | 11.91 | 10.56        | 11.34 | 11.43 | 10.95 | 10.43 | 10.22  |
| A-par          | 0.025 | 0.012 | 0.040 | 0.022        | 0.029 | 0.020 | 0.049 | 0.046 | 0.034  |
| $\pm$          | 0.098 | 0.118 | 0.115 | 0.112        | 0.116 | 0.124 | 0.135 | 0.121 | 0.110  |
| FR [%]         | 101.0 | 98.3  | 97.6  | 96.2         | 98.9  | 99.4  | 100.5 | 101.1 | 101.3  |
| $\theta_{maj}$ | 17.53 | 20.03 | 17.24 | <b>23.73</b> | 20.32 | 16.71 | 20.31 | 18.02 | 19.08  |
| $\theta_{min}$ | 11.70 | 10.98 | 12.31 | 10.98        | 11.75 | 11.90 | 11.19 | 10.89 | 10.44  |

Table 4.3.: MACF allows us to recover close to 100 % of the TP reference image in most fields, in  $\text{N}_2\text{H}^+$  (*above*) as well as in HNC (*below*). F3 in  $\text{N}_2\text{H}^+$  is the main exception, which is likely because of the sparse  $(u, v)$ -coverage. Some fields recover more than 100 % flux, which is within scatter. All A-par values show that the flux is not significantly over- or underestimated in any field, although a small systematic overestimation is present. The largest major beam gives the beam size for the common beam that all fields have been convolved to.

Before performing the comparison, the high-resolution MACF map must be convolved to the beam size of the TP map to ensure consistency. This can be done with `imsmooth`. This task needs `.fits` data, therefore we converted the `.image` file for each field inside CASA with `exportfits`. The intensity threshold in Eq. 3.8 is set to 10 Jy for all maps, ensuring that only regions of bright emission contribute to the fit, and excluding faint or noisy regions that could bias the result.

The flux recovery of F6 is displayed in Fig. 4.4. In the following, we focus on F6 as a representative case, as it received the most extensive testing throughout this work. The remaining fields, except F3 and F4, yield qualitatively consistent results, with F5 constituting an equally suitable example, and we do not expect the choice of F6 to affect the conclusions drawn here. A-par is distributed around 0 with a narrow standard deviation,  $A_{F6} = 0.024 \pm 0.115$ . Tab. 4.3 summarizes the flux recovery of all nine fields in both molecules, as well as the total flux recovered FR. The overall assessment is consistent for all fields: MACF performs reliably and ensures an excellent flux recovery, exceeding 90 % in all cases and reaching almost 100 % in most cases. In some fields, FR slightly exceeds 100 %; but this is within the expected measurement uncertainties.

Two fields have to be mentioned here explicitly, as they exhibit lower A-par values than the others: F3, with  $A_{F3} = 90.600 \pm 0.144$ , and F4, with  $A_{F4} = -0.016 \pm 0.087$ . These two fields are distinctive because their  $(u, v)$ -coverage is particularly sparse in comparison to the other fields, and tracks overlap even more in the  $(u, v)$ -plane than showcased for F6.

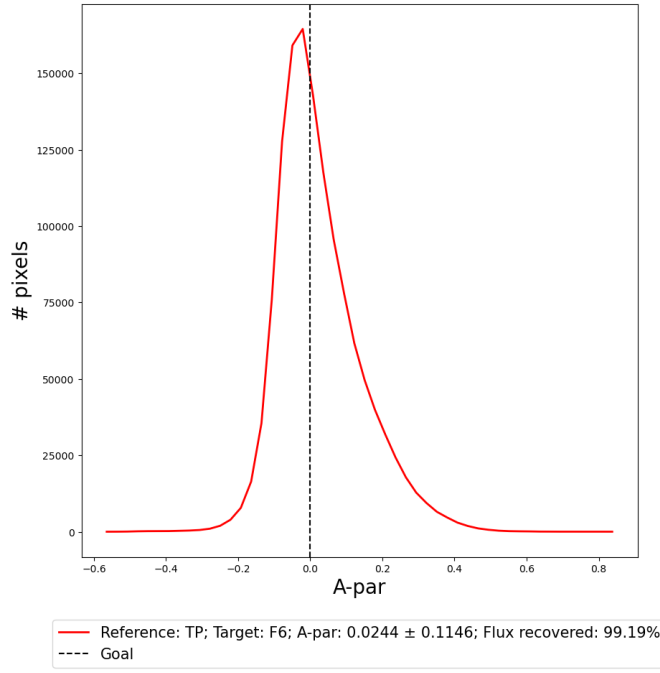


Figure 4.4.: The distribution of A-par values for F6 in  $\text{N}_2\text{H}^+$  is centered around 0, with a narrow standard deviation. This indicates an almost perfect flux recovery and therefore a successful MACF process.

#### 4.2.4. Nyquist sampling

The pixel scale of each map is chosen to fulfill the Nyquist criterion. At this stage, the pixel size still oversamples the beam, and oversampling beyond the Nyquist limit does only increase the computational cost. Each beam should have at least two pixels per FWHM, in order to reliably distinguish two emission peaks, the spatial pixels therefore have to be binned. Reducing the number of pixels does not result in a loss of information, as the data remain above the Nyquist limit. The spatial extent of two Nyquist-sampled pixel is the width of the synthesized beam, and thus gives the effective final resolution of a map. This means that the final spectral cubes have to be sampled every half beam, i.e.,  $\theta/2$ .

Nyquist sampling can be achieved either by using an image of the same field as template – provided it already has the required cell size – and regridding the image to the Nyquist cell size using `imregrid`, or by first creating a template image with half the needed cell size, and then rebinning the pixels in a second step, using `imrebin` in CASA. Both approaches yield identical results, but for some fields, only one of the two methods was applicable.

Since all individual fields are combined into one large mosaic, a common angular resolution for all fields is required. Unfortunately, we cannot use the best angular resolution for each field, but rather have to adopt the poorest resolution of all fields. Each map, which at this stage still has its native beam size, has to be convolved to the common beam of  $28''$ . Only after this homogenization step, each map can be Nyquist sampled using either `imregrid` or `imrebin` approach. The resolutions of each field after MACF are given in Tab. 4.3.

The largest major axis among the individual beams is  $23.73''$ , and to make the  $\text{N}_2\text{H}^+$  and HNC map comparable, both maps were convolved to the same common beam. Therefore, each map was convolved to a common beam of  $28'' \times 28''$ , which corresponds to the final spatial resolution. A common beam of 1.2 times the largest beam ensures a minimization of flux losses due to resampling. The native beams are highly asymmetrical because of the  $(u, v)$ -sampling. Adopting a circular beam avoids direction-dependent distortions. The resulting Nyquist-sampled pixel size is therefore  $14'' \times 14''$ .

### 4.3. Re-Projection to Galactic Coordinates and Mosaicing

If all those steps are performed, the maps of all individual fields can be combined in a final mosaic. The resulting integrated intensity maps for the two spectral windows of interest are shown in Fig. 4.1. All fields are now cleaned, feathered with the TP image, convolved to a common beam size, and Nyquist sampled.

To combine the individual fields, we used the `reproject` package from Robitaille et al. [2020]. The implemented methods allow both the construction of a mosaic out of the individual tiles and, at the same time, the re-projection of the data into the Galactic coordinate system.

The `reproject` method `find_optimal_celestial` creates a new world coordinate



#### 4.4. Resulting Maps in Two Molecular Tracers

system (WCS) object derived from the individual WCSs. In the same step, the coordinate frame is transformed from the equatorial coordinate system to the Galactic coordinate system. The resulting Galactic coordinates of each field are listed in Tab. 4.1. Nessie is a bone of the Milky Way [Goodman et al., 2014], and such structures lie in the Galactic plane [Zucker et al., 2015]. Re-projecting the data to Galactic coordinates serves two purposes: it highlights Nessie’s location within the Galactic plane and allows for a more efficient use of page space when presenting the mosaic.

This optimal WCS is then used in the next step, when `reproject_and_coadd` reprojects all individual fields onto the common coordinate grid and combines them into a single mosaic. To minimize discontinuities in the overlapping regions, background matching was enabled with `match_background=True`. This ensures a smooth transition between the individual fields and prevents individual fields from appearing artificially enhanced due to elevated background emission. A residual overlap between F7 and F6 is visible in the integrated intensity map of HNC shown in Fig. 4.1.

#### 4.4. Resulting Maps in Two Molecular Tracers

In Fig. 4.1, three different maps of the Nessie cloud are shown, detailed section images are provided in the Appendix A.1. To date, these are the highest spatial and spectral resolution maps of the full Nessie Extended region in HNC and  $\text{N}_2\text{H}^+$ , as previous observations at comparable or higher resolution have been limited to Nessie Classic [Jackson et al., 2010], to other wavelengths [Goodman et al., 2014, Mattern et al., 2018], or alternative molecular lines in individual subregions [Jackson et al., 2024]. The first panel displays a false-colour image based on *Spitzer* observations, where red corresponds to  $8.0\,\mu\text{m}$ , green to  $4.5\,\mu\text{m}$ , and blue to  $3.6\,\mu\text{m}$ . The footprint of the whole observation is overlaid, and the individual fields are labelled accordingly.

Both molecular maps have been converted to units of  $\text{K km s}^{-1}$  using Eq. 3.7 to ensure consistency in the display of the results. In both tracers, the emission traces a contiguous filamentary structure extending across all observed fields, reflecting the large-scale spatial coherence of Nessie. The filamentary emission aligns closely with the infrared absorption features visible in the *Spitzer* maps, suggesting that the molecular line emission originates from the same coherent dust structure seen in absorption. The two tracers, however, reveal complementary aspects of the filament’s structure: in  $\text{N}_2\text{H}^+$ , the emission is more concentrated in bright, spatially compact clumps that stand out clearly against a faint background, whereas in HNC the clumps appear embedded in more extended, diffuse emission that connects them along the filament spine, giving the impression of pearls on a string.

The middle panel shows the integrated intensity map of  $\text{N}_2\text{H}^+$ . The brightest emission originates from the teardrop-shaped HII shell at  $l \approx 338^\circ$ , the *Nessie Bubble* [Jackson et al., 2024], located roughly in the center of the map. Three distinct intensity peaks exceeding  $150\,\text{K km s}^{-1}$  are visible within and around the bubble, arranged in a trident-like configuration, with the largest of the three spanning approximately 12 pixels or  $\sim 2.4\,\text{pc}$  in extent. Beyond the bright clumps, the emission drops off sharply: clump

#### 4. Observational Data, Set-Up in CASA, and Resulting Maps

envelopes fall below  $80 \text{ K km s}^{-1}$ , while the gas connecting adjacent clumps along the spine reaches values below  $50 \text{ K km s}^{-1}$ . Despite this contrast, the clumps are not entirely isolated – faint emission lanes loosely connect them along the filament spine, though this interconnection is more clearly visible in HNC as described in Jackson et al. [2010]. The clumps themselves appear roughly regularly spaced along the spine, a property that has been linked to fragmentation via sausage instability in Mattern et al. [2018] (see chapter 2). The filament spine is only marginally visible in  $\text{N}_2\text{H}^+$ , as this tracer preferentially highlights the dense, compact regions rather than the more diffuse inter-clump gas. Further emission maxima are present at  $l \approx 339^\circ$  and  $l \approx 337.75^\circ$  (shown in more detail in Fig. A.1 and A.2). At  $l \approx 337.5^\circ$ , two more extended, droplet-shaped clumps are visible (Fig. A.3), which appear spatially connected via a faint emission bridge.

The right panel shows the integrated intensity map of HNC. Since HNC traces more diffuse and extended gas than  $\text{N}_2\text{H}^+$ , the overall filamentary structure is more clearly visible, with the emission broadly following the infrared absorption features. The filament is not fully continuous, however, and breaks up in several places, with the brightest HNC emission preferentially coinciding with regions of strong infrared absorption. The clumps appear embedded in more extended, diffuse emission that connects them along the spine, though their sizes vary relative to the filament width: some clumps are compact and fully embedded within the spine, while others are more extended and broader than the connecting emission. As in  $\text{N}_2\text{H}^+$ , the clumps appear roughly regularly spaced along the spine. The dynamic range in HNC is smaller than in  $\text{N}_2\text{H}^+$ , with peak intensities reaching only  $\sim 40 \text{ K km s}^{-1}$ . Consequently, some features prominent in  $\text{N}_2\text{H}^+$  are significantly weaker or absent in HNC, for example, one of the three peaks of the trident-like structure near the *Nessie Bubble* almost completely disappears. In F3, a nearly semicircular structure with an angular diameter of approximately  $8.4'$  to  $10.8'$ , corresponding to a physical extent of 8 pc to 10 pc at a distance of 3.1 kpc, is traced by comparatively strong HNC emission, and is associated with a nearby expanding supernova remnant. The two droplet-shaped clumps seem to be even more connected in HNC.

In both maps, prominent sidelobe patterns are visible, appearing as faint structures parallel to regions of bright emission. These artifacts, introduced during the cleaning process, are particularly visible in fields with comparatively weak or diffuse emission, where they are not masked by the bright emission of dense clumps. In  $\text{N}_2\text{H}^+$ , the patterns are most prominent in F3, where structures parallel to the brightest emission sources are clearly visible. Similar artifacts are present in F2, where structures parallel to the filament spine can be identified in both molecular line maps, and in F8, where stratified parallel structures are visible. As discussed in Sect. 3.3, confirming whether all such features are artifacts or whether some represent real emission would require a more detailed analysis of the spectral cubes.

#### 4.5. Spectral Line Fitting

To characterize the kinematic and physical properties of the filament – in particular the centroid velocity, line width, and peak intensity of the gas – spectral line fitting

was performed on the science-ready spectral cubes. The fitting procedure was done on *licancabur* with GILDAS-CLASS<sup>6</sup>. These quantities form the observational basis for the subsequent analysis of the fragmentation properties of the cloud.

The Grenoble Image and Line Data Analysis Software GILDAS is a comprehensive software package designed for (sub-)millimeter radio astronomy applications [Gildas Team, 2013]. One of its components, the Continuum and Line Analysis Single-dish Software CLASS [Pety, 2005], can be used to analyse spectral cubes and fit spectral lines. For this, both maps have to be converted to a CLASS-readable `.gdf` file.

For the fitting procedure, the two tracers have to be treated differently. The  $\text{N}_2\text{H}^+$  data has to be converted to  $\text{K km s}^{-1}$ , whereas the HNC data has to be left in  $\text{Jy/beam}$  for the automated fitting procedure. For visualization purposes, however, we converted the intensities shown in Fig. 4.1 and 4.5 to Kelvin using Eq. 3.5, allowing for a more direct comparison between the two molecular tracers.

#### 4.5.1. $\text{N}_2\text{H}^+$ Line Fitting

$\text{N}_2\text{H}^+(1-0)$  shows 7 emission centered around 93.173 764 GHz. To fit the spectrum, we used a semi-automated CLASS routine provided by Andrea Socci. In itself it uses a hyperfine structure data file for  $\text{N}_2\text{H}^+(1-0)$ . In this file, the relative positions and intensities of the  $\text{N}_2\text{H}^+(1-0)$  emission lines are provided, enabling a simultaneous multi-component fit. The routine takes a box of  $5 \times 5$  spectral pixels, and produces an averaged spectrum of this box. This increases the signal-to-noise ratio, which helps to identify the emission peaks and to distinguish between multiple velocity components.

In the example spectrum in Fig. 4.5, an isolated hyperfine component is visible at  $-50 \text{ km s}^{-1}$ . This peak can be used to distinguish between multiple velocity components, i.e., distinct gas structures at different radial velocities that would otherwise appear blended in a composite spectrum. If this isolated line shows a single peak, the spectrum most likely contains only one velocity component, whereas a double peak suggests the presence of two spectrally distinct components that are not fully resolved. However, only if the two peaks are clearly separated, i.e., the separation of the velocity centroids is larger than their FWHM, do we attribute two lines. The peaks for each velocity component need to be indicated manually, so that the routine can automatically fit a combined Gaussian to the spectrum. This allows for a standardized characterization of spectral lines across the 25 combined spectra, as the fitting routine takes the provided initial guesses to fit spectral lines in each pixel. In the case of Nessie, only  $\approx 5\%$  of the total 14123 pixels showed two velocity components. All the rest only showed one velocity component.

Several parameters of the fitting routine were adjusted to match the properties of our final data products. The common beam size is  $28''$ , yielding a cell size of  $14''$  following the Nyquist criterion. A  $5 \times 5$  pixel box corresponds to a spatial extent of  $5 \times 14'' = 70''$  on a side, corresponding to approximately 2.5 beam widths and ensuring that each averaged spectrum samples a well-defined, beam-scale region. We set a brightness temperature threshold of 0.2 K, in order to exclude noise-dominated spectra, chosen to be safely

---

<sup>6</sup><http://www.iram.fr/IRAMFR/GILDAS>

#### 4. Observational Data, Set-Up in CASA, and Resulting Maps

above the typical noise level while retaining faint but real emission. The fitting routine was only initiated if at least 8 spectra, i.e., roughly a third of the 25, exceeded the threshold, ensuring that the averaged spectrum represents genuine emission rather than a noise-dominated region..

The routine performs an initial non-linear least-squares fit to all 25 spectra in the box with the specified number of velocity components. Those initial fit parameters are then refined by running multiple iterations of the algorithm. Only if the displayed fit matches the spectrum, the final fit parameters are then saved manually through a `.write` file together with the position. This also serves as a checkpoint, where the routine can continue should the fitting procedure be interrupted. The routine then moves on to the next box, and the procedure is repeated.

In total, 6319 spectra were successfully fitted. This is almost 45 % of the total 14123. The spectral line fitting procedure provides key parameters, such as centroid velocity  $v_{LSR}$ , opacity  $\tau$ , antenna temperature  $T_{ant} \cdot \tau$ , and noise level  $rms$ , which we used to compute the signal-to-noise ratio  $SNR$ :

$$SNR = \frac{T_{ant} \cdot \tau}{rms} \cdot (1 - e^{-\tau}) > 3. \quad (4.2)$$

We only retained pixels with  $SNR > 3$  for further analysis. This criterion excluded 789 data points, leaving us with 5530 or 40 % of the original pixels for the final analysis.

##### 4.5.2. HNC Line Fitting

The HNC(1-0) spectrum shows in most cases only one single spectral line, as illustrated in Fig. 4.5. From the  $N_2H^+(1-0)$  spectrum, we already know that in this observation of Nessie, only a small fraction of pixels have multiple velocity components; almost all spectra only showed one velocity component. Therefore, we could use a fully automated CLASS script to fit the spectra, provided by Francesca Bonanomi. 6379 of the total 14286 pixels were above the noise level of  $1 \text{ Jy beam}^{-1}$  ( $\sim 0.19 \text{ K}$ ) and were successfully fitted. This is about 45 % of the original dataset retained for further analysis.

This script yields centroid velocities  $v_{LSR}$ , emission intensity  $I$ , and noise  $rms$ . The  $SNR$  for the HNC(1-0) data points is thus modified:

$$SNR = \frac{I}{rms} > 3. \quad (4.3)$$

As before, we set  $SNR > 3$  for all data points, but in the case of HNC(1-0), no data points were below this threshold, leaving us with all spectra.

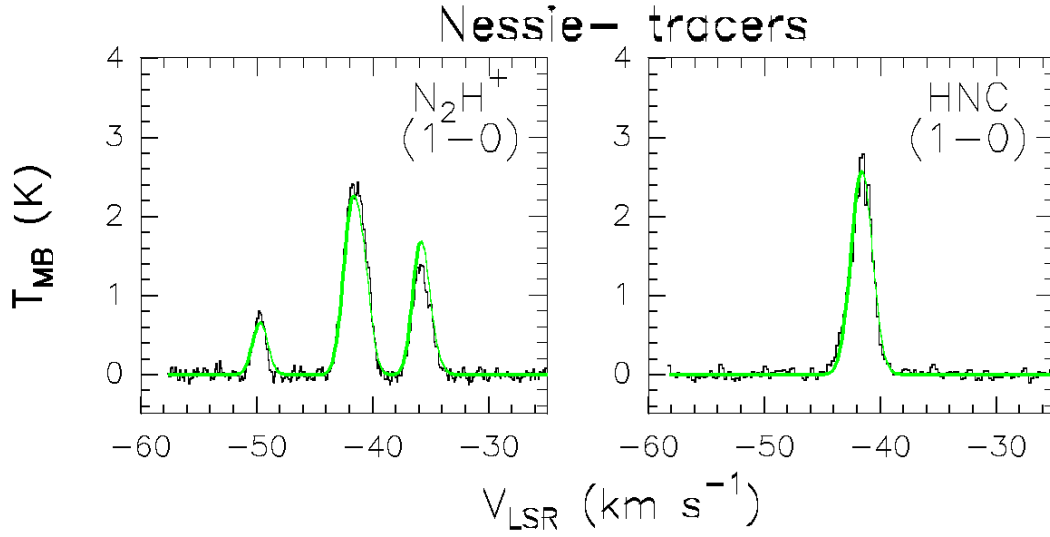


Figure 4.5.: *Left:* Example spectrum of  $\text{N}_2\text{H}^+(1-0)$ . The three peaks come from the seven hyperfine components. The leftmost peak is used to identify multiple velocity components; the other two peaks come from two triplets. The Gaussian fit is done with CLASS. *Right:* The  $\text{HNC}(1-0)$  spectrum shows a single line, with the Gaussian fit realized in CLASS. The intensity of the spectrum here is in K, whereas the automated fitting was carried out in Jy/beam.



## 5. Kinematic Analysis of the Velocity Centroids

### 5.1. Sinusoidal Undulation in Position-Velocity Space

Spectral line fitting yields, among other quantities, the velocity centroid  $v_{LSR}$  along the line of sight with respect to the local standard of rest (LSR). For each pixel that passes the threshold, the spectral line fit provides a velocity centroid. The velocity centroids after the fitting procedure range from  $-55 \text{ km s}^{-1}$  to  $-20 \text{ km s}^{-1}$ , which is well inside the limits set in the fitting procedure (compare Fig. 4.5). We excluded all points with  $v_{LSR} < -45 \text{ km s}^{-1}$  and  $v_{LSR} > 35 \text{ km s}^{-1}$ . Gas at these velocities is more likely associated with the fore- or background rather than with Nessie. A velocity map in the two molecules can be seen in Fig. 5.1.

In F9, we identify a spatially extended substructure traced by  $\text{N}_2\text{H}^+$  at  $v_{LSR} \approx -55 \text{ km s}^{-1}$  (cyan in Fig. 5.1), which we attribute to a foreground structure unassociated with Nessie. The structure spans approximately 2 pc across 29 pixels and is kinematically coherent, with a peak-to-peak velocity variation of less than  $2 \text{ km s}^{-1}$  across its full spatial extent. Another potential background structure is visible in HNC, although it is less well constrained than the foreground component.

In Fig. 5.2, we show the variation of velocity centroids of both molecules as a function of position along Nessie. Both tracers exhibit consistent velocities in the position-velocity-plot, indicating that the observed velocity variations are not confined to the dense clumps traced by  $\text{N}_2\text{H}^+$ , but are equally present in the more diffuse gas traced by HNC. This suggests that the kinematic behaviour of Nessie is coherent across different density regimes. In the leftmost region, between  $l \approx 339^\circ$  to  $339.5^\circ$ , the velocity field appears locally disturbed around  $l \approx 339.25^\circ$ , with emission spanning a significantly broader range of velocities than the surrounding filament. This kinematic disturbance may be related to the semicircular structure visible in HNC in Fig. 4.1, possibly indicating an interaction between the filament and a supernova bubble. A detailed investigation of this region goes beyond the scope of this work. Aside from this feature, no obvious patterns are evident in this part of Nessie.

Integrating the spectral cube over all velocities yields the integrated intensity map shown in Fig. 4.1. The resulting intensity profile along the filament, shown in the lower panel of Fig. 5.2, reveals several local maxima of varying prominence, the most significant of which is located at  $\sim 100 \text{ pc}$  and corresponds to the HII region in F6, the *Nessie Bubble*.

This emission maximum also marks the most prominent kinematic feature in the position-velocity diagram: the velocity field transitions smoothly from a local maximum of  $v_{LSR} \approx -42 \text{ km s}^{-1}$  to a local minimum of  $v_{LSR} \approx -38 \text{ km s}^{-1}$ , with the turning

## 5. Kinematic Analysis of the Velocity Centroids

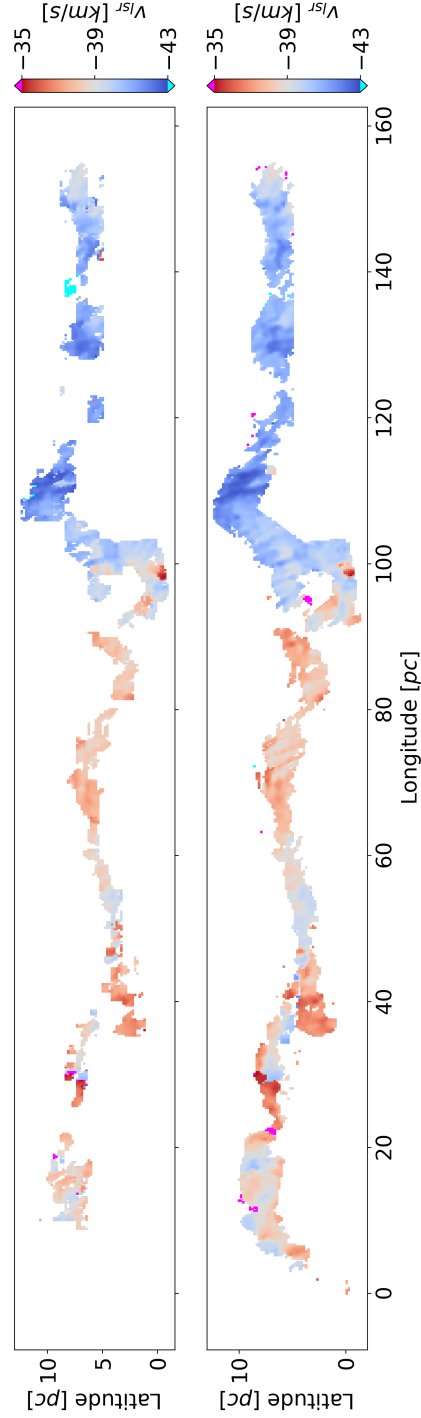


Figure 5.1.: The velocity plot shows that the right part of Nessie is more blue-shifted than the left, with F6 moving at the average velocity of  $\sim -39 \text{ km s}^{-1}$ . *Left*: velocity plot in  $\text{N}_2\text{H}^+$ . A cluster in F9 (cyan) was identified as foreground emission. *Right*: velocity plot in HNC. Both molecules show the same behaviour. Some features (magenta) may belong to emission from background structures. Residual sidelobes from the cleaning process are still visible, e.g., around 70 pc.



### 5.1. Sinusoidal Undulation in Position-Velocity Space

| Parameter                                               | N <sub>2</sub> H <sup>+</sup> | HNC                 | Combined            |
|---------------------------------------------------------|-------------------------------|---------------------|---------------------|
| $A$ [km s <sup>-1</sup> ]                               | $2.105 \pm 0.022$             | $1.978 \pm 0.024$   | $2.042 \pm 0.023$   |
| $f$ [10 <sup>-3</sup> pc <sup>-1</sup> ]                | $9.623 \pm 0.049$             | $10.752 \pm 0.060$  | $9.815 \pm 0.055$   |
| $\phi$ [°]                                              | $0.249 \pm 0.035$             | $0.059 \pm 0.041$   | $0.154 \pm 0.038$   |
| $C$ [km s <sup>-1</sup> ]                               | $-39.760 \pm 0.014$           | $-39.696 \pm 0.016$ | $-39.728 \pm 0.015$ |
| $\lambda$ [pc]                                          | $103.92 \pm 0.53$             | $99.93 \pm 0.60$    | $101.92 \pm 0.56$   |
| $\nabla v_{LSR}$ [km s <sup>-1</sup> pc <sup>-1</sup> ] | $0.1797 \pm 0.0048$           | $0.1923 \pm 0.0039$ | $0.1860 \pm 0.0044$ |

Table 5.1.: Sine fit parameters for both molecules are very similar and can therefore be combined to a mean value. The mean value of the errors was used in the MCMC estimation of the  $1\sigma$  uncertainty band shown in Fig. 5.2.

point of this sinusoidal pattern coinciding with the emission maximum of the *Nessie Bubble*. Sinusoidal velocity patterns are a commonly observed feature in filamentary molecular clouds [Henshaw et al., 2020], and the position-velocity diagram indeed reveals a prominent oscillatory pattern in both tracers across the position range 50 pc to 150 pc. A sine fit of the form

$$f(x) = A \cdot \sin(2\pi f \cdot x + \phi) + C,$$

where  $A$  is the amplitude,  $f$  the frequency (such that  $1/f = \lambda$  corresponds to the wavelength),  $\phi$  the phase, and  $C$  the offset, was applied to both oscillation patterns with `scipy.optimize.curvefit`. The resulting fit parameters are similar for both tracers, see Tab. 5.1. The uncertainties of the individual parameters are the square root of the diagonal elements of the covariance matrix.

This quantifies the properties of the oscillatory pattern, and both tracers yield consistent results. For the following, we adopt the mean values of all fit parameters. Jackson et al. [2010] classify *Nessie* as a velocity coherent structure at  $(-38.0 \pm 3.4)$  km s<sup>-1</sup>. This is reflected in the offset of  $(-39.728 \pm 0.015)$  km s<sup>-1</sup>, as well as in the derived amplitude of  $(2.042 \pm 0.023)$  km s<sup>-1</sup>. The amplitude introduces a complication when calculating the kinematic distance. If the right arm of *Nessie* is at  $-42$  km s<sup>-1</sup> and the other one at  $-38$  km s<sup>-1</sup>, the corresponding kinematic distance will differ.

Since  $1/f = \lambda$ , we can quantify the wavelength with  $(101.92 \pm 0.56)$  pc, which matches the scatter of velocity centroids within the  $1\sigma$  uncertainty band of the sine fit. We estimated the uncertainty by conducting a Markov Chain Monte Carlo (MCMC) simulation of the sine fits. A total of 5000 samples were drawn from normal distributions with means  $\mu = A, f, \phi, C$  and standard deviations  $\sigma = \sigma_A, \sigma_f, \sigma_\phi, \sigma_C$ . This procedure yields 5000 different realizations of the sine function. The  $1\sigma$  standard deviation of these realizations defines the upper and lower bounds for the fitted curve.

Since the velocity gradient varies along the filament axis, we evaluate it at the position of the turning point of the sinusoidal pattern,  $l \approx 98.94$  pc, which coincides with the emission maximum of the *Nessie Bubble* and therefore represents the most prominent and well-defined kinematic reference point along the filament. At this position, we measure  $\nabla v_{LSR} = (0.1860 \pm 0.0044)$  km s<sup>-1</sup> pc<sup>-1</sup>. This value is at the lower bound of typical

## 5. Kinematic Analysis of the Velocity Centroids

gradients for IRDCs,  $0.2 \text{ km s}^{-1} \text{ pc}^{-1}$  [Hacar et al., 2023], while GMFs typically show values not larger than  $0.08 \text{ km s}^{-1} \text{ pc}^{-1}$ .

In the lower panel of Fig. 5.2 we show the intensity profile along the filament axis, where each emission peak corresponds to a dense clump along the spine. The spatial distribution of these peaks provides a key observable for testing fragmentation models, as gravitational fragmentation via sausage instability predicts a characteristic clump separation [Mattern et al., 2018].

To identify the emission peaks, the integrated intensity is projected onto the principal axis of the filament, without correction for its bending. We then detect the peaks using `scipy.signal.find_peaks`, with a minimum integrated intensity threshold of  $100 \text{ K km s}^{-1}$  applied to the  $\text{N}_2\text{H}^+$  map. Since the intensity peaks of HNC consistently correspond to those of  $\text{N}_2\text{H}^+$  but are always less prominent, we only identify the peaks in  $\text{N}_2\text{H}^+$  to avoid missing fainter peaks in the more diffuse tracer. All detected peaks are marked with  $\times$  in Fig. 5.2.

The mean separation between the identified peaks is  $(3.54 \pm 1.71) \text{ pc}$ , consistent with the value of  $3.5 \text{ pc}$  reported by Mattern et al. [2018]. This is also in agreement with the expected mean separation of cores resulting from sausage instability, for which Eq. 2.5 yields  $3.32 \text{ pc}$ .

## 5.2. New Distance Estimate to Nessie

If an extended object is placed in three-dimensional space, artificial wave patterns can arise in a position-velocity diagram purely due to projection effects of the velocities along the line of sight [Burton, 1971]. These apparent waves can have large amplitudes [Soler et al., 2022], comparable to what we observe in Nessie (Fig. 5.2). It is therefore necessary to assess the potential contribution of projection effects associated with the spatial orientation of the filament. If the filament, for example, lies perpendicular to the line of sight, the observed velocities can only track the gas flow in the line-of-sight direction (see Sec. 5.3). To ensure that the observed velocity patterns do not arise purely from projection effects, but reflect the intrinsic kinematic structure of the filament, we need to constrain the position angle (PA) under which we observe Nessie. Since the PA depends on the assumed distance to the filament, we first derive an improved distance estimate in the following section before returning to the deprojection in Sec. 5.3.

In order to de-project the observed velocities, the distance to the object is needed. Current values in the literature range from  $3.1 \text{ kpc}$  [Jackson et al., 2010] to  $(3.5 \pm 0.5) \text{ kpc}$  [Mattern et al., 2018] for the entire filament, and  $3.23 \text{ kpc}$  [Jackson et al., 2024] for the *Nessie Bubble*. The values span a total range of  $\sim 1 \text{ kpc}$  when accounting for uncertainties, which is not sufficiently precise for our purposes. We therefore attempted to better constrain the distance to Nessie using two independent methods: the kinematic distance and the break point in the extinction curve.

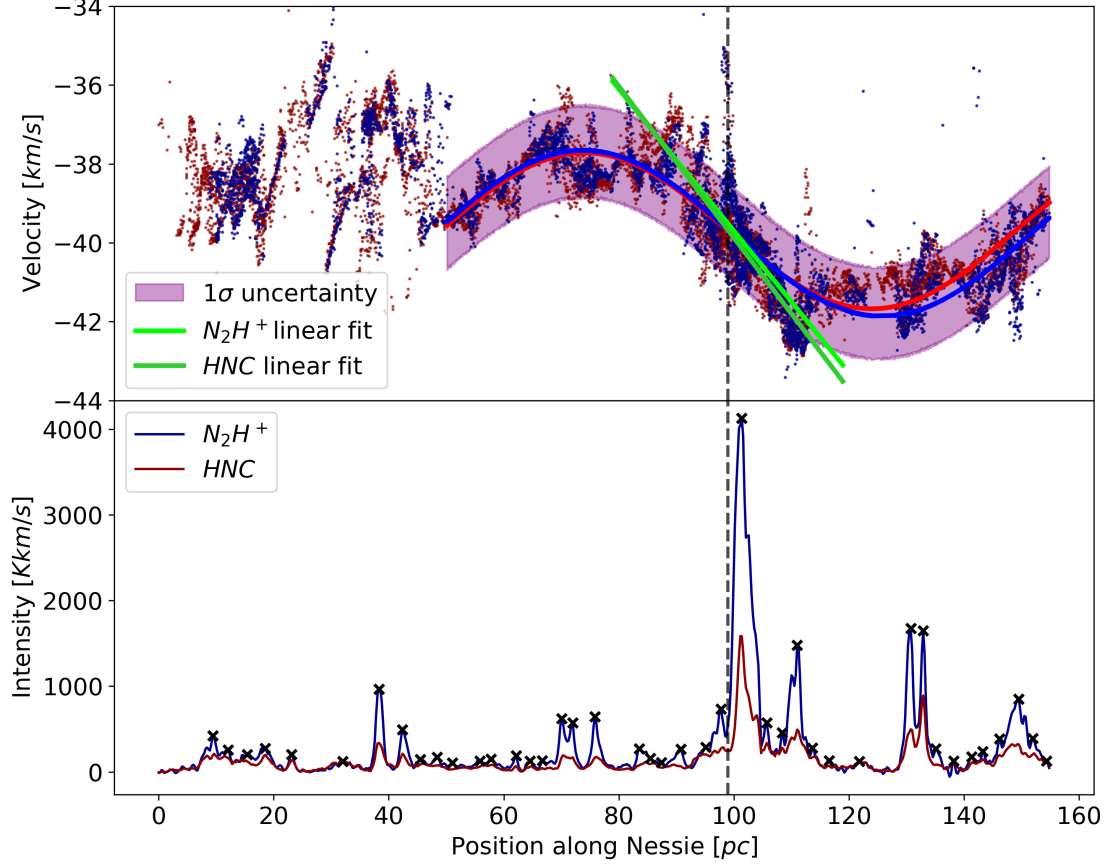


Figure 5.2.: *Top:*  $N_2H^+$  (blue) and HNC (red) velocities track the same behaviour in the pv-plot. There is a prominent wave pattern around the emission maximum at 100 pc cloud length. The wave in both tracers can be described by a sinusoidal function. A linear fit (green) to the central parts (dashed line) of the two sinusoids yields  $\Delta v_{LSR}$ . *Bottom:* The integrated intensity along the spine shows many emission peaks, marked with  $\times$ , and traces the position of clumps in Nessie. The mean separation is  $(3.54 \pm 1.71)$  pc.

## 5. Kinematic Analysis of the Velocity Centroids

### 5.2.1. Distance Estimation Based on Kinematic Distance

Assuming a rotation curve  $V(R)$ , one can calculate the LOS distance  $d$  to an object. This requires the velocity with respect to the local standard of rest  $v_{LSR}$ , and the position of the object in Galactic coordinates  $l, b$ . These quantities can be re-projected to Galactocentric coordinates, yielding the Galactocentric radius  $R_{gal}$ :

$$R_{gal} = R_0 \sin l \frac{V(R_{gal})}{v_{LSR} + V_0 \sin l}, \quad (5.1)$$

where  $R_0$  and  $V_0$  are the Galactocentric radius and tangent velocity of the solar orbit around the Galactic center, respectively.  $V(R_{gal})$  is the rotation curve evaluated at the position  $R_{gal}$ .

The distance  $d$  is then given by:

$$d = R_0 \cos l \pm \sqrt{R_{gal}^2 - R_0^2 \sin^2 l}. \quad (5.2)$$

In the inner Galaxy ( $R_{gal} < R_0$ ), Eq. 5.2 yields two solutions, a near and a far kinematic distance,  $d_{near}$  and  $d_{far}$ .

These equations can be solved using the python package **Kinematic Distance Utilities**<sup>1</sup> [Wenger et al., 2018], where we used the rotation curve from Reid et al. [2019]. The code takes  $v_{LSR}$  obtained from spectral line fitting and the Galactic coordinates  $l, b$  for each pixel and returns the corresponding near and far kinematic distances along the line of sight.

We applied the routine to all pixels with a successful spectral fit in  $N_2H^+$  that pass the threshold. The tolerance to determine a match between the rotation curve and the measured LSR velocity is 0.1. The resulting histogram of  $d_{near}$  is shown in Fig. 5.3. We decided to take  $d_{near}$ , since all distances  $d_{far}$  are above 12 kpc, which is inconsistent with the values from the literature. The resulting kinematic distances follow a normal distribution, which we fit using `scipy.stats.norm.fit`. The fitted Gaussian is centered at 2.77 kpc with a standard deviation of 0.075 kpc. This yields a kinematic distance  $d_{kinematic}$  of  $(2.77 \pm 0.15)$  kpc, where the uncertainty is equal to the  $2\sigma$  interval of the fit. The distribution shows a second peak, which can be attributed to the two segments of the wave.

### 5.2.2. Distance Estimation Based on the Break Point in *Gaia* Colour Excess

Molecular clouds scatter light from background stars, causing them to appear less luminous or to be completely obscured. This effect is quantified either by the visual extinction  $A_V$  or by the colour excess  $E(B - V)$ , defined as the difference between observed and intrinsic  $(B - V)$  magnitudes [Draine, 2011]. This phenomenon motivates an independent, photometric method to constrain the distance to Nessie: the identification of a break point in the colour excess  $E(B - V)$  [Cardelli et al., 1989]. Stars located in front of the cloud are only weakly affected, whereas stars behind the cloud experience additional

<sup>1</sup><https://zenodo.org/records/1166001>

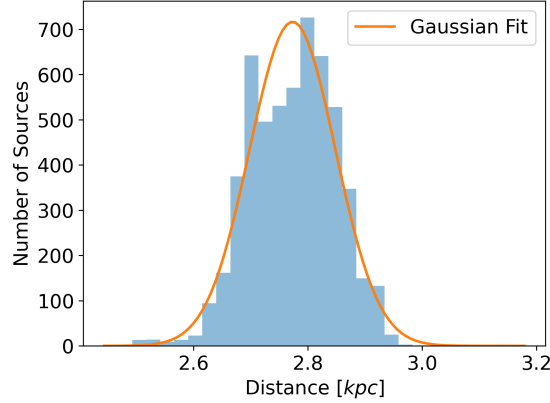


Figure 5.3.: For each pixel and its corresponding  $v_{LSR}$  we calculated the kinematic distance. The histogram can be described by a Gaussian, which peaks at 2.77 kpc with a standard deviation of 0.075 kpc.

reddening. Consequently, a jump in colour excess  $E(B - V)$  is expected at a distance  $d_{extinction}$ , where the cloud is located [Schlafly et al., 2014, Zucker et al., 2019, Mei et al., 2024].

*Gaia* was a space observatory operated by ESA and primarily designed for high-precision astrometry. It has made more than 3 trillion observations of 2 billion stars and other objects throughout our Milky Way galaxy and beyond, mapping their motions, luminosity, temperature, and composition. For stars up to 1 kpc the measured parallaxes are of unprecedented accuracy [Gaia Collaboration et al., 2021].

*Gaia* therefore provides both needed quantities for this method: distances  $d$  via parallaxes  $\varpi$  with

$$d = \frac{1}{\varpi}, \quad (5.3)$$

and the colour excess in the optical. In the case of *Gaia*, the photometric bands are *BP* (400 nm to 650 nm), and *G* (400 nm to 900 nm), which are the blue and broad-optical bands of the *Gaia* photometric system [Gaia Collaboration et al., 2021], roughly analogous to the classical *B* and *V* Johnson bands, respectively. The colour excess  $E(BP - G)$  therefore serves the same diagnostic role as  $E(B - V)$ . It is derived from these bands and provided directly in the *Gaia* data product. The error in colour excess  $\sigma(E(BP - G))$  is not given explicitly in the catalogue, but it can be calculated from the reported flux uncertainties [Gaia Collaboration, 2020] following the Gauss' law of error propagation:

$$\sigma(E(BP - G)) = \frac{2.5}{\ln 10} \sqrt{\left(\frac{\sigma_{BP}}{F_{BP}}\right)^2 + \left(\frac{\sigma_G}{F_G}\right)^2}, \quad (5.4)$$

where  $F_{BP}$  and  $F_G$  are the measured fluxes and  $\sigma_{BP}$  and  $\sigma_G$  their respective uncertainties in the *BP* and *G* bands.

At distances of  $\sim 3$  kpc, the individual stellar parallax uncertainties from *Gaia* are substantial, with distance errors of order several hundred parsecs per star. In principle,

## 5. Kinematic Analysis of the Velocity Centroids

these uncertainties affect the inferred cloud distance, as they smear out the location of the extinction jump across distance bins. However, since the jump is identified from the ensemble behaviour of many stars across multiple bins rather than from any individual parallax measurement, the position of the break point is determined statistically, and its uncertainty is driven primarily by the bin width and the sharpness of the jump rather than by individual distance errors. The dominant source of uncertainty on the mean colour excess per bin is therefore the scatter in  $E(BP - G)$  among stars within that bin, propagated from the reported flux uncertainties as described above.

We looked only at stars that are 1) within 6 kpc and 2) lie within the field of Nessie, traced by the footprint of the ALMA observations. After applying this selection criteria, 41657 stars remained that all have measured  $E(BP - G)$ , as well as  $\sigma_{BP}$  and  $\sigma_G$ . To reduce the scatter in the data, we calculated the mean value of all stars that are within a 100 pc bin over the range 0 pc to 6000 pc. The uncertainty of each bin was estimated from the mean of  $\sigma(E(BP - G))$  in each bin.

The resulting profile in Fig. 5.4a shows a clear increase in colour excess beginning at 2.8 kpc. At intermediate distances in the foreground, between 1 kpc to 2.5 kpc,  $E(BP - G)$  remains approximately constant. Interestingly,  $E(BP - G)$  at 0 pc is higher than at 1 kpc, where it stays constant. This leads to a curved behaviour of the  $E(BP - G)$  profile.

The colour excess increases from 2.8 kpc to 3.5 kpc with an approximately constant slope, in good agreement with the literature, yet more constrained. This suggests a foreground level of  $E(BP - G) = 0.95$  mag. After 3.5 kpc however,  $E(BP - G)$  flattens out and sets at a background level of  $E(BP - G) \approx 1.2$  mag.

Overall, the  $E(BP - G)$  profile resembles a broad "U"-shaped curve: relatively higher values at very small distances, a flatter minimum at intermediate distances, and a systematic increase at larger distances due to the additional extinction introduced by the cloud.

### 5.2.3. Distance Estimation Based on the Break Point in 2MASS Extinction

Infrared light is less affected by dust scattering and extinction than optical light. It therefore gives a measure of extinction independent from optical *Gaia* colour excess. Comparing infrared- and optical-based extinction measurements, therefore, allows us to assess whether the break in  $E(BP - G)$  is robust or influenced by wavelength-dependent systematics [Maheswar et al., 2010].

We applied the same methodology as described above to data from the Two Micron All-Sky Survey (2MASS), a ground-based astronomical survey of the whole sky in infrared light. It was conducted at three distinct infrared bands,  $J$  (1.25  $\mu\text{m}$ ),  $H$  (1.65  $\mu\text{m}$ ), and  $K_s$  (2.16  $\mu\text{m}$ ) [Skrutskie et al., 2006], and is therefore well suited for our purposes. We cross-matched the *Gaia* and 2MASS catalogues using a matching radius of 1". This allowed us to keep the parallaxes from *Gaia*, while adopting the near-infrared colour information from 2MASS, in particular the  $H - K_s$  colour excess. In total, we obtained 5828 stars from this cross-match.

The uncertainty of the colour excess  $\sigma(E(H - K_s))$  is:

$$\sigma(E(H - K_s)) = \sqrt{\sigma_H^2 + \sigma_{K_s}^2}, \quad (5.5)$$

where  $\sigma_H$  and  $\sigma_{K_s}$  are the photometric uncertainties of the magnitudes in the respective bands, as provided in the 2MASS data product.

The infrared colour excess can be converted to visual extinction  $A_V$  using the standard reddening law of Rieke and Lebofsky [1985]:

$$A_V = \frac{0}{0.112} \cdot E(H - K_s). \quad (5.6)$$

This converts the infrared excess to a visual extinction in magnitudes, which gives it a clearer physical interpretation. The general picture stays the same: The behaviour of  $A_V$  depends on the distance. Within the first 2 kpc  $A_V$  has a similar curved U-profile as  $E(BP - G)$ , although in the infrared data it is even more pronounced than in the optical data. Stars located in front of the cloud have a typical visual extinction  $A_V \approx 2.5$  mag, while stars behind the cloud have  $A_V \approx 4$  mag. As in the  $E(BP - G)$  profile, the increase in  $A_V$  begins at 2.8 kpc, becoming steeper at around 3.1 kpc. This is again consistent with the distance estimates in Mattern et al. [2018] or Jackson et al. [2010]. The  $A_V$  profile levels off again at 3.5 kpc, consistent with the flattening observed in the optical data. This value as well is in good agreement with the literature values presented above, but placing the cloud to closer distances.

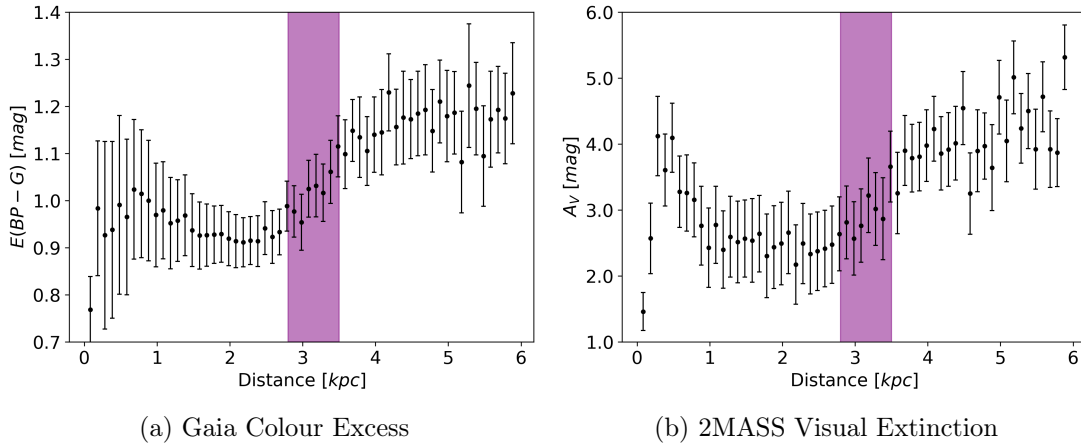


Figure 5.4.: *Left:* The Gaia colour excess increases at about 2.8 kpc and at flattens beyond 3.5 kpc. *Right:* Also the 2MASS extinction shows an increase at around 2.8 kpc, becoming more pronounced at 3.1 kpc, and similarly flattens beyond 3.5 kpc.

## 5. Kinematic Analysis of the Velocity Centroids

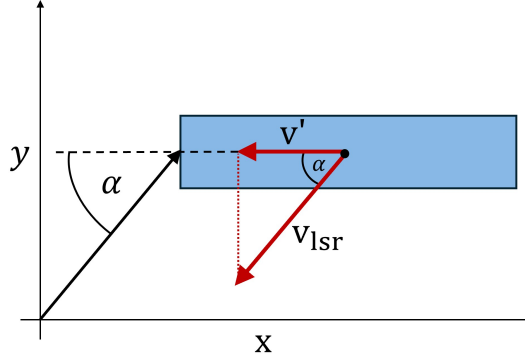


Figure 5.5.: The LOS velocity  $v_{LSR}$  is the quantity observed from Earth. The velocity along the spine of the filament  $v'$  can be recovered using the position angle  $\alpha$  of the filament.

### 5.3. (No) De-Projection of Velocities

The observed LOS velocities  $v_{LSR}$  are given in the LSR reference frame. Therefore, they do not provide information about the actual gas flow along the spine of the filament. With two simplifying assumptions, we can attempt to constrain this motion.

The first assumption is that the gas flows *only* along the spine, i.e., the principal axis of the filament. This is a clear oversimplification of the true structure, which in reality is highly unlikely to have a uniform gas flow in just one principal direction. However, in the absence of additional information, we have to assume a well-behaved gas motion in the cloud.

The second assumption is that the filament is not bent, but can be approximated as a straight structure. This simplifies the modelling by allowing us to adopt a linear increase in PA for the whole structure. With this assumption, we might miss critical information on the actual geometry of the entire filament, but as current data give no possibility to reconstruct the three-dimensional shape of the object, we have to employ a simplified description, which might not satisfy reality.

What we can therefore derive is the velocity  $v'$  projected onto the spine of the filament. The geometry of this configuration is shown in Fig. 5.5. The filament is placed in the Galactic plane, with the Sun located at the origin of the sketch. The projected velocity  $v'$  depends on the position angle  $\alpha$ , which is the angle of the filament to the line of sight, and can therefore vary by its placement in the Galactic plane.  $v'$  is connected to  $v_{LSR}$  via the cosine of  $\alpha$ :

$$v' = \frac{v_{LSR}}{\cos \alpha}.$$

The position angle of the filament is approximately  $77^\circ$ ; the effect on the position-velocity diagram is twofold.

First, the de-projection substantially modifies the absolute velocity scale. For a PA of  $77^\circ$  to the near end of the filament,  $1/\cos \alpha$  is equal to 4.44. Nessie has a POS angle



of about  $2.5^\circ$ , and the far end of the filament is consequently at  $79^\circ$ , where  $1/\cos \alpha$  is equal to 5.24. The velocities are therefore enlarged more than five times, setting the de-projected values to  $-200 \text{ km s}^{-1}$ .

Second, the de-projection alters the shape and amplitude of the sinusoidal pattern. The amplitude increases from approximately around  $2 \text{ km s}^{-1}$  before to  $10 \text{ km s}^{-1}$  after de-projection. This yields a key result: the sine wave still appears even after de-projecting. The observed oscillatory pattern is therefore intrinsic to the cloud and cannot be explained solely as a projection effect.

Because the de-projection relies on several simplifying assumptions and yields unphysical results, while not altering the qualitative behaviour, we decided not to de-project the observations, as this would introduce new complications. The de-projection would change the x-axis or the position of pixels along the spine, and in consequence, also lead to a larger separation of the clumps along Nessie.

## 5.4. Power Spectrum of PV-Space

The velocity field of Nessie, as described in Sec. 5.1, contains a prominent undulation at the scale of the filament alongside smaller-scale velocity variations. Whether these smaller-scale variations are driven by gravitational collapse of individual clumps or by turbulent motions is not immediately apparent from the position-velocity diagram alone. To disentangle the large-scale coherent motion from clump-scale and turbulent contributions, we compute the power spectrum  $P(k)$ , which decomposes the velocity field into its constituent spatial frequencies [Elmegreen and Scalo, 2004]. For purely turbulent, incompressible flow,  $P(k)$  follows the scale-free Kolmogorov power law [Kolmogorov, 1941], with no preferred spatial scale. A distinct peak at a specific wavenumber, for example corresponding to  $\lambda$ , would therefore indicate a physical process injecting energy at that scale, standing out above the turbulent background.

To analyse the pv-diagram for periodicity, including the large-scale wave with  $\lambda \approx 102 \text{ pc}$ , we computed the power spectrum  $P(k)$  of the velocity profile. Let the distribution in pv-space be described by the observable  $v(x)$ . Then the corresponding power spectrum  $P(k)$  is defined as the Fourier transform  $FT\{v(x)\} = \hat{v}(k)$  times its complex conjugate  $\hat{v}(k)^*$  [Elmegreen and Scalo, 2004]:

$$P(k) \equiv \hat{v}(k)\hat{v}(k)^* = |\hat{v}(k)|^2. \quad (5.7)$$

For incompressible turbulence, the Kolmogorov power spectrum follows a power-law behaviour of the form  $P \propto k^{-5/3}$  [Kolmogorov, 1941]. In Fig. 5.6, the power spectrum of the pv-diagram is shown. To compute it, we divided the Galactic longitude axis into 2000 equally spaced bins, each of width  $\Delta l = L/2000$  where  $L$  is the total length of the filament, and computed the mean velocity within each bin. The number of contributing data points varies between bins depending on the local sampling density. Bins without data points were filled by interpolation to obtain a uniformly sampled velocity profile. This resulting cleaned pv-profile was Fourier transformed using `numpy.fft.fft`, the corresponding Fourier frequency values  $k$  were computed with `numpy.fft.fftfreq`. The

## 5. Kinematic Analysis of the Velocity Centroids

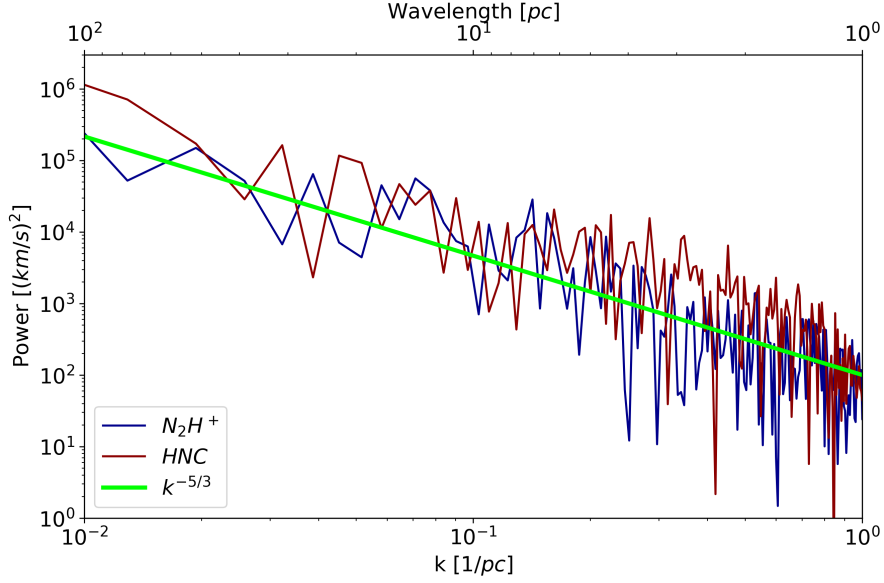


Figure 5.6.: The power spectra of the observed positions and velocities of HNC (red) and  $N_2H^+$  (blue) follow a Kolmogorov slope (green).

power spectrum in Fig. 5.6 follows a power law  $\propto k^{-5/3}$ , as predicted from theory. This result must be interpreted carefully, though.

In a turbulent ISM, energy is transferred from large to small scales in a Kolmogorov-type cascade [Kolmogorov, 1941, Larson, 1981, Elmegreen and Scalo, 2004], producing a power spectrum with no preferred spatial scale. A power spectrum that follows  $k^{-5/3}$ , such as the power spectrum in Fig. 5.6, is therefore commonly interpreted as evidence for turbulence, implying a disordered velocity field without periodicity.

This would suggest that the large scale velocity pattern in Fig. 5.2 originates from turbulence. We show, however, that this interpretation is not necessarily true: we show that Kolmogorov-like power spectra can also result from a superposition of periodic signals in pv-space, and therefore do not necessarily imply turbulent, chaotic motion.

In Fig. 5.7, we present a series of sinusoidal functions of increasing complexity, constructed to progressively approximate the observed velocity profile in Fig. 5.2, together with their corresponding power spectra. This allows us to demonstrate how a superposition of periodic signals builds up a Kolmogorov-like power spectrum. This approach is analogous to Hacar et al. [2016], where they found that a 2-point correlation function is similar to Larson’s  $\sigma$ - $L$  relationship. The first row shows a simple sine function

$$g_1(x) = A \cdot \sin f \cdot x + C, \quad (5.8)$$

with  $A = 2$ ,  $f = \pi/10$ , and  $C = -39$ . The peak in the power spectrum is placed at the corresponding wavelength  $\lambda = 2\pi/f = 2\pi/\pi/10 = 20$  pc.

The linear function in the second row is

$$g_2(x) = kx + C - k \cdot 80, \quad (5.9)$$

with  $k = -0.04$ . The constant term  $k \cdot 80$  shifts the function such that it lies between  $g_2(0) = -36$  and  $g_2(150) = -42$ . The third function is a linear combination of the two previous functions:

$$g(x) = g_1(x) + g_2(x) - C. \quad (5.10)$$

In the fourth row, we construct a piecewise function  $h(x)$  that reproduces the observed velocity profile of Nessie as closely as possible. For  $x > 50$ ,  $h(x)$  combines a large-scale wave with several superimposed smaller-scale oscillations, while for  $x < 50$  it accounts for the more irregular, less periodic behaviour seen in that region of the position-velocity diagram:

$$h(x) = \begin{cases} \sin \left[ \frac{\pi}{2} \cdot \frac{L}{\lambda_2 - \lambda_1} \cdot \ln \left( \lambda_1 + \frac{x}{L}(\lambda_2 - \lambda_1) \right) \right] + C, & \text{for } x < 50 \\ A \cdot \sin \left( \frac{\pi}{50}x - 3 \right) + \frac{1}{2} \cdot \sin \frac{\pi}{10}x + \frac{1}{2} \cdot \sin \frac{\pi}{4}x + C, & \text{for } x \geq 50. \end{cases} \quad (5.11)$$

The parameters were found by trial and error to approximate the measurements closely, with  $L = 150$ ,  $\lambda_1 = 0.2$ , and  $\lambda_2 = 8$ . In the fifth and final row, we add Gaussian noise to  $h(x)$ , with  $\mu = 0$  and  $\sigma = 0.2$ .

This demonstrates that a purely periodic, non-chaotic velocity field can produce a power spectrum consistent with a Kolmogorov-type slope. Consequently, the power spectrum in Fig. 5.6 alone cannot be taken as definitive evidence for a turbulent cascade in Nessie, and by extension, cannot be used to argue that the large-scale oscillation identified in Fig. 5.2 originates from turbulent motions. Other physical explanations for the observed velocity pattern therefore remain viable.

## 5. Kinematic Analysis of the Velocity Centroids

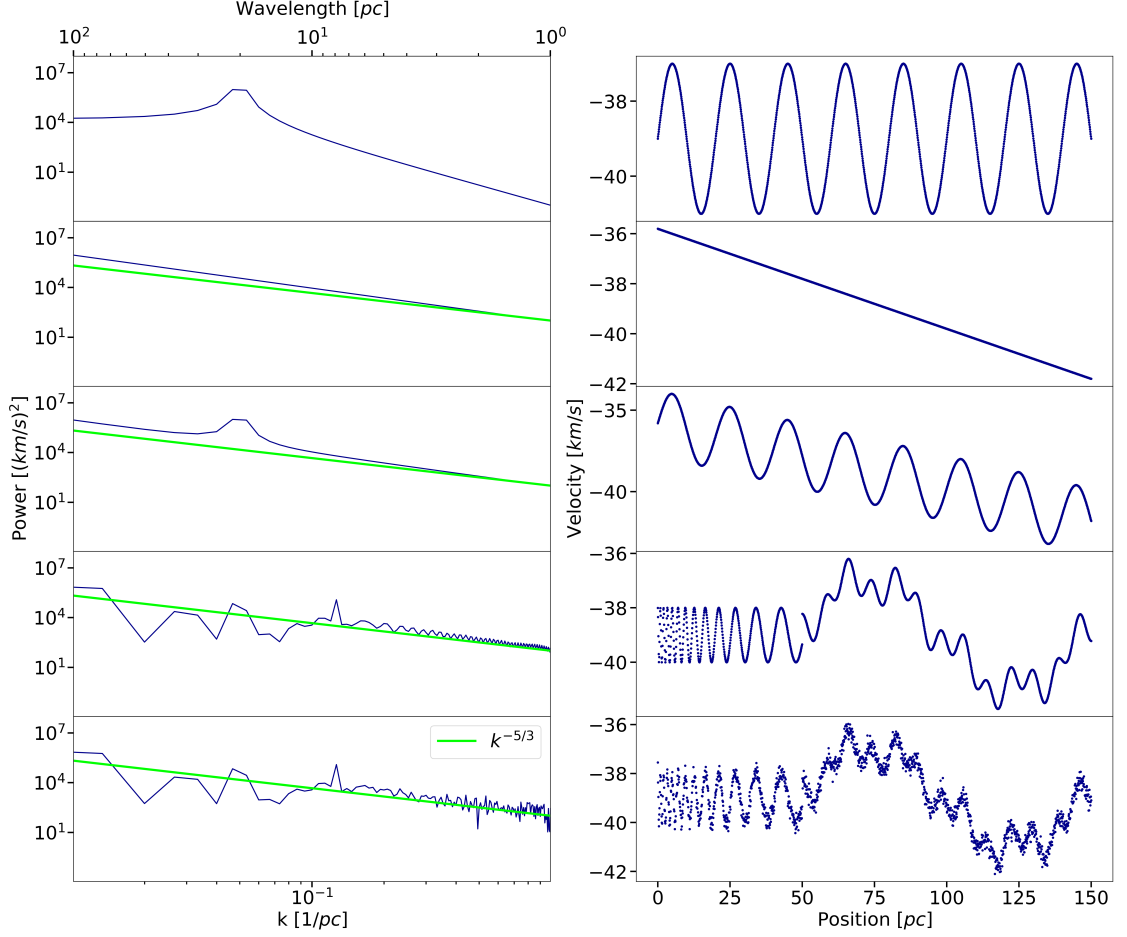


Figure 5.7.: The power spectra (*left*) of the corresponding analytic functions (*right*), i.e., the pv-plot, follow a Kolmogorov  $-5/3$  slope (green). With each step, the analytic function more closely resembles the observations, and also the power spectrum increasingly matches the power spectrum obtained from the real observations in Fig. 5.6.

## 6. Discussion

The high-resolution interferometric maps of Nessie in the two molecular tracers  $\text{N}_2\text{H}^+$  and HNC allowed us to investigate the kinematic behaviour of this GMF in unprecedented detail (see Fig. 4.1, with  $28''$  resolution). Our investigation showed that there is a coherent velocity wave going through Nessie with a wavelength of  $(101.92 \pm 0.56)$  pc, see Fig. 5.2. We quantified this pattern by fitting a sine curve to the position-velocity distribution. To estimate the  $1\sigma$ -uncertainty band, we conducted an MCMC simulation of 5000 samples based on the uncertainties of the best-fit parameters to this wave.

Large scale undulating patterns and velocity gradients over large angles are an inherent property of large-scale structures and are found everywhere in the ISM, but also in extra-Galactic observations [Henshaw et al., 2020]. In the GMF CFG024.00+0.48, He et al. [2026] found wave-like patterns, but with considerably shorter wavelengths at around 25 pc. The HI filament Maggie shows oscillations with wavelengths  $(502 \pm 113)$  pc [Syed et al., 2022], and in the IRDC 28.37+0.07 Beuther et al. [2026] found a west-east velocity gradient in the cloud.

The sinusoidal velocity pattern in Nessie is present in both molecular tracers, indicating that the dense gas component traced by  $\text{N}_2\text{H}^+$  follows the same large-scale kinematic behaviour as the more diffuse gas traced by HNC. The undulating velocity structure is also evident in the velocity map of Fig. 5.1, where the turning point of the sinusoidal pattern coincides with the emission maximum of the *Nessie Bubble*, the teardrop-shaped HII region at the center of the filament [Jackson et al., 2024]. Fitting a linear regression to that region yields a velocity gradient  $\nabla v_{LSR} = (0.1860 \pm 0.0044) \text{ km s}^{-1} \text{ pc}^{-1}$ .

It should be emphasized that the work we present here is purely observational and has not been corrected for projection effects. In the case of extended sources, apparent POS longitudinal velocity patterns can arise just from projection effects to the line of sight Burton [1971] and therefore require a careful de-projection Soler et al. [2022]. In an attempt to de-project the data, we found out that the PA under which Nessie is observed,  $\approx 77^\circ$ , introduces a substantial offset to the measured velocity centroids if we de-projected the data. However, the sinusoidal pattern is not removed by this correction. On the contrary, the de-projection significantly amplifies the wave, enlarging the amplitude by a factor of 5. We therefore concluded that the sinusoidal pattern we found reflects a real physical behaviour rather than a geometric effect caused by projection along the LOS.

The investigation of the projection effects also needed a careful revision of the literature values for the LOS distance to Nessie, which allowed us to place constraints on the literature values. First, we derived the kinematic distance  $d_{kinematic}$ , based on the measured velocity centroids we obtained from our spectral line fitting. Second, we determined an extinction-based distance  $d_{extinction}$  by identifying breaks in  $E(BP - G)$  or  $A_V$  as a function of distance. We applied this method to two data sets, one from *Gaia*

## 6. Discussion

( $E(BP - G)$ ) and a *Gaia*-2MASS cross-matched catalogue ( $A_V$ ).

For the kinematic distance, we used the `astropy` package `KDUtils` Wenger et al. [2018], adopting a rotation curve by Reid et al. [2019]. This method yields a distance of  $(2.77 \pm 0.15)$  kpc, which is systematically lower than the values in the literature. Jackson et al. [2010] give 3.1 kpc (without uncertainty), Mattern et al. [2018] give  $(3.5 \pm 0.5)$  kpc, and Jackson et al. [2024] places the central HII region of Nessie at 3.23 kpc, based on the work of Allingham et al. [2024].

There are several caveats to those values. First, Jackson et al. [2010] use a rotation curve from Clemens [1985]. Given that Reid et al. [2019] provide a more refined rotation curve that incorporates more recent parallax measurements and revised Galactic parameters, this different methodology may explain the difference between our derived value and the value from Jackson et al. [2010].  $d_{kinematic}$  is not agreement with the value from Mattern et al. [2018], who derive the distance using the foreground star density. Allingham et al. [2024] found two  $\text{NH}_3(3,3)$  masers in Nessie F6, and give 3.23 kpc and 12.66 kpc (without uncertainties) as distance estimates to Nessie. Allingham et al. [2024] get these values from Whitaker et al. [2017], who calculate kinematic distances to star-forming clumps. The large value 12.66 kpc is more likely to be  $d_{far}$ , arising from the near-far ambiguity of Eq. 5.2, rather than two physically distinct solutions for Nessie.

The increase in  $E(BP - G)$  starts at the same distance as in  $A_V$ , and the overall profile of both extinction-based measurements is very similar. In both data sets, *Gaia* and 2MASS, we observe an increase in colour excess or extinction at a distance closer to us than 3.1 kpc. The increase in both data sets started at 2.8 kpc, and flattens beyond 3.5 kpc. From this behaviour, we therefore derive an extinction-based distance  $d_{extinction} = (3.1 \pm 0.3)$  kpc. This value is consistent with the value from Jackson et al. [2010], while now providing an explicit uncertainty range. Within this uncertainty range, it is plausible, and particularly when considered together with our  $d_{kinematic}$ , even likely, that Nessie is located closer than previously thought.

Taken together, the kinematic distance and the two extinction-based estimates are mutually consistent, all pointing to a distance in the range 2.8 kpc to 3.1 kpc. This places Nessie systematically closer than the values adopted in most of the literature, and the agreement between the three methods strengthens this conclusion. Crucially, our extinction-based estimate provides an explicit uncertainty range of  $(3.1 \pm 0.3)$  kpc for a distance that was previously quoted without uncertainty. We therefore adopt  $(3.1 \pm 0.3)$  kpc as our best estimate for the distance to Nessie, while noting that the true distance may be somewhat lower, closer to 2.8 kpc, as suggested by the kinematic estimate.

The physical origin of the observed sinusoidal velocity pattern remains unclear; nonetheless, some possible explanations can be ruled out. In particular, Jeans fragmentation cannot explain the undulation we see in pv-space. If local overdensities are caused by Jeans instability, the wavelength in velocity space would follow the Jeans length. Following Eq. 2.1, the wavelength in velocity space should be 0.95 pc, which is not evident from our data. Jeans instability thus also fails to explain the mean separation of clumps. This discrepancy is not unexpected, as the Jeans instability model relies on several idealized

assumptions. It assumes a spherical, homogeneous, and isothermal cloud in equilibrium; assumptions clearly not met in the case of Nessie. The filamentary morphology alone violates the assumption of a spherical cloud. The presence of over-dense clumps, the differing spatial distribution of HNC and  $\text{N}_2\text{H}^+$  gas, and the velocity patterns at large and small scales indicate that the cloud is neither uniform nor in global equilibrium. Adopting the more adequate description of a filament by Ostriker [1964] and Stodólkiewicz [1963], the predicted mean separation of clumps, based on Eq. 2.3, is 1.18 pc, which is not the separation of the clumps observed in Nessie [Mattern et al., 2018].

Sausage instability can, however, possibly explain the mean separation of clumps [Jackson et al., 2010, Mattern et al., 2018]. Using the same assumptions as for the Jeans instability, Eq. 2.5 yields a wavelength of  $\lambda_S = 3.32$  pc. This value is consistent with the mean separation of the intensity peaks,  $(3.54 \pm 1.71)$  pc, measured in Fig. 5.2, where the peaks are marked with  $\times$  in the POS longitudinal intensity profile. Given that Nessie is observed under a large positional angle to the line of sight, the true separation of clumps has to be larger than the projected value here and in Mattern et al. [2018] (3.5 pc). As with Jeans' instability, also sausage instability relies on highly idealized assumptions. While Nessie appears to be elongated enough to be approximated by an infinitely long cylinder, the assumption is not strictly satisfied. However, Sausage instability cannot provide a viable explanation for the large-scale velocity pattern. In this sense, a purely gravitational origin for the sinusoidal wave can be ruled out.

In a turbulent and chaotic environment, the power spectrum of the velocity field is expected to follow a Kolmogorov cascade [Kolmogorov, 1941], where the power-law scales with  $k^{3/5}$ . The power spectrum derived from the position-velocity distribution shows this trend, which could suggest that the velocity pattern arises from turbulent, stochastic motions alone. However, we demonstrate that the inverse of this statement is not necessarily true; a power spectrum that scales  $\propto k^{3/5}$  is not, by itself, sufficient evidence for purely turbulent behaviour.

In Fig. 5.7, we present several analytical functions together with their respective power spectra. In every case, the power spectrum follows a Kolmogorov turbulent cascade, even though the position-velocity space is sampled by analytical functions, which are by construction well-behaved, periodic, and therefore non-chaotic.

One possible explanation we found for the large-scale oscillation comes from the dynamical description of a cold fluid disk. Toomre's instability works on large-scale dynamics and is therefore more suited to explain large-scale structures that theories based on gravitational fragmentation cannot explain.

Nessie is located at  $R_{gal} = 5267$  pc, see Fig. A.4. Following the IAU recommendation [Whitaker et al., 2017], we adopt the circular velocity  $V_c = 220 \text{ km s}^{-1}$ , and obtain  $\kappa = 59 \text{ km s}^{-1} \text{ pc}^{-1}$  for the ISM at a Galactocentric distance of Nessie. With these values, Eq. 2.7 yields a critical Toomre length of  $\lambda_T \approx 630$  pc. Perturbations in the ISM with wavelengths  $\lambda < 630$  pc can therefore be amplified [Li et al., 2017].

$\lambda_T$  is strongly dependent on  $V_c$ . If a different rotation curve is assumed, e.g.  $240 \text{ km s}^{-1}$  by [Reid et al., 2014], the value decreases to  $\lambda_{T_2} \approx 530$  pc.

$\lambda_{char} = \lambda_T / 2\pi$  [Li et al., 2017] gives the scaling relation of  $\lambda_T$  and sets an even lower

## 6. Discussion

limit of  $\approx 100$  pc. Coincidentally, it matches fairly well with the wavelength in Nessie. This value  $\lambda_{char}$  should be interpreted with care. A scaling relation of this kind provides only an order-of-magnitude estimate, and agreement within a factor of a few is not sufficient to uniquely identify Toomre instability as the driving mechanism.  $\lambda_{char}$  also does not necessarily reflect the true maximum wavelength in the environment of Nessie, which is twice the critical wavelength anyway, but rather gives a characteristic length. First, the value itself *strongly* depends on the assumed rotation curve and can yield very different results depending on the chosen  $V_c$ . Second,  $\lambda_T$  also strongly depends on  $\Sigma$ , and if it is increased,  $\lambda_T$  also increases.

The agreement between  $\lambda$  and  $\lambda_T$  suggests that the large-scale oscillation may have a Galactic dynamical origin, as the Toomre wavelength characterizes the scale at which self-gravity overcomes differential rotation and turbulent pressure in a differentially rotating disk. The wavelength we find is  $(101.92 \pm 0.56)$  pc and lies well below the upper limit of  $\lambda_T$ . It is thus consistent with an instability described by a fluid disk in Toomre [1964], and linked to large-scale Galactic dynamics, rather than to local gravitational fragmentation.

The comparison of different imaging procedures, see Tab. 4.2, confirms a well-established picture: When working with ALMA data, data combination is essential to achieve a high flux recovery, and MACF is preferable to feathering alone. Using MACF, we could recover up to 10 % more flux, while systematic under- or overestimations are reduced, as reflected in the A-par distribution that is more tightly centered around 0. In comparison to MACF, imaging based solely on 7m data results in a flux recovery that is not insufficient for quantitative scientific analysis.

### 6.1. Outlook

Nessie was observed in 5 molecular tracers. This leaves us with 3 additional molecules to study, one of which is  $\text{HC}_3\text{N}$ , which can be used to probe the densest parts of star forming clouds [Bergin and Tafalla, 2007].

Comparable objects show similar patterns in pv-space. In the case of Nessie, the observed pattern may extend beyond the region analysed here and could continue into the adjacent part of the filament, which has not been observed in the required resolution in those tracers yet. It may therefore be necessary to observe "Nessie Optimistic" to determine whether this wave-like pattern is confined to the region analysed here or represents a coherent dynamical feature spanning the entire cloud. If the oscillation persists over the full extent of Nessie, this would strongly support a Galactic-scale driving mechanism such as Toomre instability, as the coherence length would far exceed what local gravo-turbulent fragmentation models can produce. If the pattern is absent in the unobserved sections, a more localised dynamical origin would need to be considered. In general, Nessie needs to be observed in higher resolution to resolve filamentary structures on core-scales.

Besides the large-scale sinusoidal pattern, several small-scale oscillatory features can be seen in Fig. 5.2. These velocity undulations coincide spatially with intensity peaks in  $\text{N}_2\text{H}^+$ , and thus with identified clumps, for example at around 70 pc and 130 pc. The



investigation of the origin of these velocity fluctuations is beyond this work, but we speculate that the origin of these local velocity fluctuations arise from gas infall motions associated with the individual clumps. Similar oscillatory signatures have been found in dense cores [Hacar and Tafalla, 2011], and IRDCs in general [Schneider et al., 2015, Henshaw et al., 2014]. The presence of such features in Nessie could indicate gravitational collapse in the clumps, but this has to be investigated, as the aforementioned examples are at much smaller scales.

One open question concerns the true extent of Nessie, which remains uncertain. The analysis of Goodman et al. [2014] relied on CO data for the large-scale morphology and  $\text{NH}_3$  observations at only a sparse set of positions. Whether the region that constitutes Nessie Optimistic is truly velocity-coherent and physically associated with Nessie Extended remains unconfirmed and is even contradictory in the literature: Zucker et al. [2015] lists the filament at  $l \approx 335^\circ$  as a separate bone. To understand this one could inspect the  $^{13}\text{CO}$  data of the SEDIGISM survey Schuller et al. [2021], which provides continuous coverage at  $30''$  resolution and would allow a systematic search for velocity-coherent emission along the full proposed extent of the filament.

The distance estimate based on photometry could also be further refined. Above 1 kpc the parallaxes provided by *Gaia* have larger uncertainties than at closer distances. In our approach, this might lead to stars being assigned to the wrong distance bin. The jump in  $E(BP - G)$  might be more pronounced when implementing distances with smaller uncertainties, such as from Bailer-Jones et al. [2021].



## 7. Conclusion

In the course of this work, we have successfully created high-resolution ( $28''$ ) spectral maps of the GMF Nessie in two different molecular tracers, HNC and  $\text{N}_2\text{H}^+$ . The gas line-of-sight velocity, extracted by spectral line fitting, showed a particular feature in the position-velocity space – an unusually large sinusoidal pattern – spanning over  $(101.92 \pm 0.56)$  pc, with an amplitude of  $(2.042 \pm 0.023)$   $\text{km s}^{-1}$ . This wave was still apparent, even after the quantities were corrected for projection effects.

This de-projection needed a careful investigation of the estimated distance of Nessie. Two methods gave results consistent with one another. The kinematic distance  $d_{\text{kinematic}} = (2.77 \pm 0.15)$  kpc, and a photometry-based approach gave  $d_{\text{extinction}} = (3.1 \pm 0.3)$  kpc. We therefore conclude that Nessie is located at a distance  $d_{\text{Nessie}} = (3.1 \pm 0.3)$  kpc. This value is consistent with literature values and with  $d_{\text{kinematic}}$ . The lower bound of 2.8 kpc that we provide suggests that the cloud might be closer than previously thought.

We characterized the properties of the wave and investigated possible fragmentation scenarios and how they relate to star formation. The most likely cause for this undulation is Toomre’s instability, and thus a dynamical process on the Galactical scale. The critical Toomre length in Nessie’s environment is  $\approx 630$  pc, and  $\lambda$  is well below that value. Different assumptions can yield reduced values of  $\lambda_T$ , but the characteristic length scale is always in the order of several 100 pc.



# Bibliography

- Y. Aikawa, N. Ohashi, S.-i. Inutsuka, E. Herbst, and S. Takakuwa. Molecular Evolution in Collapsing Prestellar Cores. *ApJ*, 552(2):639–653, May 2001. doi: 10.1086/320551.
- D. Allingham, J. M. Jackson, T. Hogge, J. S. Whitaker, P. Patterson, N. Killerby-Smith, J. Askew, and T. Vandenberg. The Curated ATCA Census of High-Mass Clumps (CACHMC) Legacy Survey. *PASA*, 41:e079, Oct. 2024. doi: 10.1017/pasa.2024.68.
- P. André, A. Men’shchikov, S. Bontemps, V. Könyves, F. Motte, N. Schneider, P. Didelon, V. Minier, P. Saraceno, D. Ward-Thompson, J. di Francesco, G. White, S. Molinari, L. Testi, A. Abergel, M. Griffin, T. Henning, P. Royer, B. Merín, R. Vavrek, M. Attard, D. Arzoumanian, C. D. Wilson, P. Ade, H. Aussel, J. P. Baluteau, M. Benedettini, J. P. Bernard, J. A. D. L. Blommaert, and Cambrésy. From filamentary clouds to prestellar cores to the stellar IMF: Initial highlights from the Herschel Gould Belt Survey. *A&A*, 518:L102, July 2010. doi: 10.1051/0004-6361/201014666.
- P. André, J. Di Francesco, D. Ward-Thompson, S. I. Inutsuka, R. E. Pudritz, and J. E. Pineda. From Filamentary Networks to Dense Cores in Molecular Clouds: Toward a New Paradigm for Star Formation. In H. Beuther, R. S. Klessen, C. P. Dullemond, and T. Henning, editors, *Protostars and Planets VI*, pages 27–51, Jan. 2014. doi: 10.2458/azu\_uapress\_9780816531240-ch002.
- D. Arzoumanian, P. André, P. Didelon, V. Könyves, N. Schneider, A. Men’shchikov, T. Sousbie, A. Zavagno, S. Bontemps, J. di Francesco, M. Griffin, M. Hennemann, T. Hill, J. Kirk, P. Martin, V. Minier, S. Molinari, F. Motte, N. Peretto, S. Pezzuto, L. Spinoglio, D. Ward-Thompson, G. White, and C. D. Wilson. Characterizing interstellar filaments with Herschel in IC 5146. *A&A*, 529:L6, May 2011. doi: 10.1051/0004-6361/201116596.
- Astropy Collaboration, A. M. Price-Whelan, P. L. Lim, N. Earl, N. Starkman, L. Bradley, D. L. Shupe, A. A. Patil, L. Corrales, C. E. Brasseur, M. Nöthe, A. Donath, E. Tollerud, B. M. Morris, A. Ginsburg, E. Vaher, B. A. Weaver, J. Tocknell, W. Jamieson, M. H. van Kerkwijk, T. P. Robitaille, B. Merry, M. Bachetti, H. M. Günther, Astropy Project Contributors, and others. The Astropy Project: Sustaining and Growing a Community-oriented Open-source Project and the Latest Major Release (v5.0) of the Core Package. *ApJ*, 935(2):167, Aug. 2022. doi: 10.3847/1538-4357/ac7c74.
- C. A. L. Bailer-Jones, J. Rybizki, M. Fouesneau, M. Demleitner, and R. Andrae. Estimating Distances from Parallaxes. V. Geometric and Photogeometric Distances to

## Bibliography

- 1.47 Billion Stars in Gaia Early Data Release 3. *AJ*, 161(3):147, Mar. 2021. doi: 10.3847/1538-3881/abd806.
- E. A. Bergin and M. Tafalla. Cold Dark Clouds: The Initial Conditions for Star Formation. *ARA&A*, 45(1):339–396, Sept. 2007. doi: 10.1146/annurev.astro.45.071206.100404.
- H. Beuther, C. Gieser, H. Linz, Q. Zhang, S. Feng, A. Ahmadi, J. D. Soler, D. Semenov, M. R. A. Wells, and S. Reyes-Reyes. The converging gas flow around the infrared dark cloud G28.3. *arXiv e-prints*, art. arXiv:2601.07407, Jan. 2026. doi: 10.48550/arXiv.2601.07407.
- J. Binney and S. Tremaine. *Galactic Dynamics: Second Edition*. Princeton University Press, 2008.
- F. Bonanomi, A. Hacar, A. Socci, D. Petry, and S. Suri. Emergence of high-mass stars in complex fiber networks (EMERGE). II. The need for data combination in ALMA observations. *A&A*, 688:A30, Aug. 2024. doi: 10.1051/0004-6361/202348920.
- W. B. Burton. Galactic Structure Derived from Neutral Hydrogen Observations Using Kinematic Models Based on the Density-wave Theory. *A&A*, 10:76, Jan. 1971.
- G. Busquet, R. Estalella, A. Palau, H. B. Liu, Q. Zhang, J. M. Girart, I. de Gregorio-Monsalvo, T. Pillai, G. Anglada, and P. T. P. Ho. What Is Controlling the Fragmentation in the Infrared Dark Cloud G14.225-0.506?: Different Levels of Fragmentation in Twin Hubs. *ApJ*, 819(2):139, Mar. 2016. doi: 10.3847/0004-637X/819/2/139.
- J. A. Cardelli, G. C. Clayton, and J. S. Mathis. The Relationship between Infrared, Optical, and Ultraviolet Extinction. *ApJ*, 345:245, Oct. 1989. doi: 10.1086/167900.
- CASA Team, B. Bean, S. Bhatnagar, S. Castro, J. Donovan Meyer, B. Emonts, E. Garcia, R. Garwood, K. Golap, J. Gonzalez Villalba, P. Harris, Y. Hayashi, J. Hoskins, M. Hsieh, P. Jagannathan, W. Kawasaki, A. Keimpema, M. Kettenis, J. Lopez, J. Marvil, J. Masters, A. McNichols, D. Mehringer, R. Miel, G. Moellenbrock, F. Montesino, T. Nakazato, J. Ott, D. Petry, M. Pokorny, R. Raba, U. Rau, D. Schiebel, N. Schweighart, S. Sekhar, K. Shimada, D. Small, J.-W. Steeb, K. Sugimoto, V. Suoranta, T. Tsutsumi, I. M. van Bemmelen, M. Verkouter, A. Wells, W. Xiong, and others. CASA, the Common Astronomy Software Applications for Radio Astronomy. *PASP*, 134(1041): 114501, Nov. 2022. doi: 10.1088/1538-3873/ac9642.
- P. Caselli, P. C. Myers, and P. Thaddeus. Radio-astronomical Spectroscopy of the Hyperfine Structure of  $\text{N}_2\text{H}^+$ . *ApJL*, 455:L77, Dec. 1995. doi: 10.1086/309805.
- S. Chandrasekhar and E. Fermi. Problems of Gravitational Stability in the Presence of a Magnetic Field. *ApJ*, 118:116, July 1953. doi: 10.1086/145732.
- D. P. Clemens. Massachusetts-Stony Brook Galactic plane CO survey: the galactic disk rotation curve. *ApJ*, 295:422–436, Aug. 1985. doi: 10.1086/163386.

- P. C. Cortes, C. Vlahakis, A. Hales, J. Carpenter, W. Dent, S. Kamen, R. Loomis, B. Vila-Vilaro, K. Immer, C. Vlahakis, J. Law, F. Stoecklin, and K. Saini. Alma technical handbook, alma doc. 11.3, ver. 1.4, Mar. 2024.
- T. M. Dame, D. Hartmann, and P. Thaddeus. The Milky Way in Molecular Clouds: A New Complete CO Survey. *ApJ*, 547(2):792–813, Feb. 2001. doi: 10.1086/318388.
- B. T. Draine. *Physics of the Interstellar and Intergalactic Medium*. 2011.
- B. G. Elmegreen and J. Scalo. Interstellar Turbulence I: Observations and Processes. *ARA&A*, 42(1):211–273, Sept. 2004. doi: 10.1146/annurev.astro.41.011802.094859.
- A. Faure, H. N. Varambhia, T. Stoecklin, and J. Tennyson. Electron-impact rotational and hyperfine excitation of HCN, HNC, DCN and DNC. *MNRAS*, 382(2):840–848, Dec. 2007. doi: 10.1111/j.1365-2966.2007.12416.x.
- J. Feng, R. J. Smith, A. Hacar, S. E. Clark, and D. Seifried. On the evolution of the observed mass-to-length relationship for star-forming filaments. *MNRAS*, 528(4): 6370–6387, Mar. 2024. doi: 10.1093/mnras/stae407.
- Gaia Collaboration. Gaia DR2 Primer. Technical report, ESA, 2020. URL <https://www.cosmos.esa.int/documents/29201/1773953/Gaia+DR2+primer+version+1.3.pdf>.
- Gaia Collaboration, A. G. A. Brown, A. Vallenari, T. Prusti, J. H. J. de Bruijne, C. Babusiaux, M. Biermann, O. L. Creevey, D. W. Evans, L. Eyer, A. Hutton, F. Jansen, C. Jordi, S. A. Klioner, U. Lammers, L. Lindegren, X. Luri, F. Mignard, C. Panem, D. Pourbaix, S. Randich, P. Sartoretti, C. Soubiran, N. A. Walton, F. Arenou, C. A. L. Bailer-Jones, and others. Gaia Early Data Release 3. Summary of the contents and survey properties. *A&A*, 649:A1, May 2021. doi: 10.1051/0004-6361/202039657.
- Gildas Team. GILDAS: Grenoble Image and Line Data Analysis Software. Astrophysics Source Code Library, record ascl:1305.010, May 2013.
- A. A. Goodman, J. Alves, C. N. Beaumont, R. A. Benjamin, M. A. Borkin, A. Burkert, T. M. Dame, J. Jackson, J. Kauffmann, T. Robitaille, and R. J. Smith. The Bones of the Milky Way. *ApJ*, 797(1):53, Dec. 2014. doi: 10.1088/0004-637X/797/1/53.
- A. Hacar and M. Tafalla. Dense core formation by fragmentation of velocity-coherent filaments in L1517. *A&A*, 533:A34, Sept. 2011. doi: 10.1051/0004-6361/201117039.
- A. Hacar, J. Kainulainen, M. Tafalla, H. Beuther, and J. Alves. The Musca cloud: A 6 pc-long velocity-coherent, sonic filament. *A&A*, 587:A97, Mar. 2016. doi: 10.1051/0004-6361/201526015.
- A. Hacar, S. E. Clark, F. Heitsch, J. Kainulainen, G. V. Panopoulou, D. Seifried, and R. Smith. Initial Conditions for Star Formation: a Physical Description of the Filamentary ISM. In S. Inutsuka, Y. Aikawa, T. Muto, K. Tomida, and M. Tamura, editors, *Protostars and Planets VII*, volume 534 of *Astronomical Society of the Pacific Conference Series*, page 153, July 2023. doi: 10.48550/arXiv.2203.09562.

## Bibliography

- A. Hacar, A. Socci, F. Bonanomi, D. Petry, M. Tafalla, D. Harsono, J. Forbrich, J. Alves, J. Grossschedl, J. R. Goicoechea, J. Pety, A. Burkert, and G. X. Li. Emergence of high-mass stars in complex fiber networks (EMERGE). I. Early ALMA Survey: Observations and massive data reduction. *A&A*, 687:A140, July 2024. doi: 10.1051/0004-6361/202348565.
- C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant. Array programming with NumPy. *Nature*, 585(7825):357–362, Sept. 2020. doi: 10.1038/s41586-020-2649-2. URL <https://doi.org/10.1038/s41586-020-2649-2>.
- Q.-R. He, W.-J. Kim, G. A. Fuller, A. Traficante, S. D. Clarke, Y. Gao, X.-P. Chen, M. Fang, K. Wang, E. Chen, T. Baug, X.-L. Wang, C. Wang, and Y.-X. Wang. The Structure of an 80 pc Long Massive Filament. *arXiv e-prints*, art. arXiv:2601.13616, Jan. 2026. doi: 10.48550/arXiv.2601.13616.
- F. Heitsch, L. W. Hartmann, and A. Burkert. Fragmentation of Shocked Flows: Gravity, Turbulence, and Cooling. *ApJ*, 683(2):786–795, Aug. 2008. doi: 10.1086/589919.
- J. D. Henshaw, P. Caselli, F. Fontani, I. Jiménez-Serra, and J. C. Tan. The dynamical properties of dense filaments in the infrared dark cloud G035.39-00.33. *MNRAS*, 440(3):2860–2881, May 2014. doi: 10.1093/mnras/stu446.
- J. D. Henshaw, P. Caselli, F. Fontani, I. Jiménez-Serra, J. C. Tan, S. N. Longmore, J. E. Pineda, R. J. Parker, and A. T. Barnes. Investigating the structure and fragmentation of a highly filamentary IRDC. *MNRAS*, 463(1):146–169, Nov. 2016. doi: 10.1093/mnras/stw1794.
- J. D. Henshaw, J. M. D. Kruijssen, S. N. Longmore, M. Riener, A. K. Leroy, E. Rosolowsky, A. Ginsburg, C. Battersby, M. Chevance, S. E. Meidt, S. C. O. Glover, A. Hughes, J. Kainulainen, R. S. Klessen, E. Schinnerer, A. Schrubba, H. Beuther, F. Bigiel, G. A. Blanc, E. Emsellem, T. Henning, C. N. Herrera, E. W. Koch, J. Pety, S. E. Ragan, and J. Sun. Ubiquitous velocity fluctuations throughout the molecular interstellar medium. *Nature Astronomy*, 4:1064–1071, July 2020. doi: 10.1038/s41550-020-1126-z.
- J. A. Högbom. Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines. *A&AS*, 15:417, June 1974.
- J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007. doi: 10.1109/MCSE.2007.55.
- S.-I. Inutsuka and S. M. Miyama. Self-similar Solutions and the Stability of Collapsing Isothermal Filaments. *ApJ*, 388:392, Apr. 1992. doi: 10.1086/171162.



- J. M. Jackson, S. C. Finn, E. T. Chambers, J. M. Rathborne, and R. Simon. The “Nessie” Nebula: Cluster Formation in a Filamentary Infrared Dark Cloud. *ApJL*, 719(2): L185–L189, Aug. 2010. doi: 10.1088/2041-8205/719/2/L185.
- J. M. Jackson, J. S. Whitaker, E. Chambers, R. Simon, C. Guevara, D. Allingham, P. Patterson, N. Killerby-Smith, J. Askew, T. Vandenberg, H. A. Smith, P. Sanhueza, I. W. Stephens, L. Bonne, F. L. Polles, A. Schmiedeke, N. Honigh, and M. Justen. Absorption and Self-absorption of [C II] and [O I] Far Infrared Lines toward a Bright Bubble in the Nessie Infrared Dark Cloud. *ApJ*, 965(2):187, Apr. 2024. doi: 10.3847/1538-4357/ad2849.
- J. H. Jeans. The Stability of a Spherical Nebula. *Philosophical Transactions of the Royal Society of London Series A*, 199:1–53, Jan. 1902. doi: 10.1098/rsta.1902.0012.
- J. Kainulainen, J. Alves, H. Beuther, T. Henning, and F. Schuller. Mass reservoirs surrounding massive infrared dark clouds. A view by near-infrared dust extinction. *A&A*, 536:A48, Dec. 2011. doi: 10.1051/0004-6361/201117757.
- R. S. Klessen, J. Ballesteros-Paredes, Y. Li, and M.-M. Mac Low. Gravoturbulent Star Cluster Formation. In H. J. G. L. M. Lamers, L. J. Smith, and A. Nota, editors, *The Formation and Evolution of Massive Young Star Clusters*, volume 322 of *Astronomical Society of the Pacific Conference Series*, pages 299–308, Dec. 2004. doi: 10.48550/arXiv.astro-ph/0403469.
- A. Kolmogorov. The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds’ Numbers. *Akademiia Nauk SSSR Doklady*, 30:301–305, Jan. 1941.
- R. B. Larson. Turbulence and star formation in molecular clouds. *MNRAS*, 194:809–826, Mar. 1981. doi: 10.1093/mnras/194.4.809.
- R. B. Larson. Cloud fragmentation and stellar masses. *MNRAS*, 214:379–398, June 1985. doi: 10.1093/mnras/214.3.379.
- G.-X. Li, F. Wyrowski, and K. Menten. Revealing a spiral-shaped molecular cloud in our galaxy: Cloud fragmentation under rotation and gravity. *A&A*, 598:A96, Feb. 2017. doi: 10.1051/0004-6361/201628251.
- M.-M. Mac Low and R. S. Klessen. Control of star formation by supersonic turbulence. *Reviews of Modern Physics*, 76(1):125–194, Jan. 2004. doi: 10.1103/RevModPhys.76.125.
- G. Maheswar, C. W. Lee, H. C. Bhatt, S. V. Mallik, and S. Dib. A method to determine distances to molecular clouds using near-IR photometry. *A&A*, 509:A44, Jan. 2010. doi: 10.1051/0004-6361/200912594.

## Bibliography

- M. Mattern, J. Kainulainen, M. Zhang, and H. Beuther. Structure and fragmentation of a high line-mass filament: Nessie. *A&A*, 616:A78, Aug. 2018. doi: 10.1051/0004-6361/201731778.
- J. Mei, Z. Chen, Z. Jiang, S. Zheng, and H. Feng. New method for estimating molecular cloud distances based on Gaia, 2MASS, and the TRILEGAL galaxy model. *A&A*, 685: A39, May 2024. doi: 10.1051/0004-6361/202347952.
- P. C. Myers. Dense cores in dark clouds. III. Subsonic turbulence. *ApJ*, 270:105–118, July 1983. doi: 10.1086/161101.
- P. C. Myers. Filamentary Structure of Star-forming Complexes. *ApJ*, 700(2):1609–1625, Aug. 2009. doi: 10.1088/0004-637X/700/2/1609.
- J. Ostriker. The Equilibrium of Polytropic and Isothermal Cylinders. *ApJ*, 140:1056, Oct. 1964. doi: 10.1086/148005.
- P. Padoan, M. Juvela, A. A. Goodman, and Å. Nordlund. The Turbulent Shock Origin of Proto-Stellar Cores. *ApJ*, 553(1):227–234, May 2001. doi: 10.1086/320636.
- L. Pagani, F. Daniel, and M.-L. Dubernet. On the frequency of  $\text{N}_2\text{H}^+$  and  $\text{N}_2\text{D}^+$ . *A&A*, 494(2):719–727, Feb. 2009. doi: 10.1051/0004-6361:200810570.
- N. Peretto, G. A. Fuller, A. Duarte-Cabral, A. Avison, P. Hennebelle, J. E. Pineda, P. André, S. Bontemps, F. Motte, N. Schneider, and S. Molinari. Global collapse of molecular clouds as a formation mechanism for the most massive stars. *A&A*, 555: A112, July 2013. doi: 10.1051/0004-6361/201321318.
- D. Petry, M. Díaz Trigo, R. Kneissl, I. Toledo, and S. Facchini. New methods for ALMA angular-scale based observation scheduling, quality assessment, and beam shaping. In D. S. Adler, R. L. Seaman, and C. R. Benn, editors, *Observatory Operations: Strategies, Processes, and Systems VIII*, volume 11449 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 114491D, Dec. 2020. doi: 10.1117/12.2557596.
- D. Petry, M. Díaz Trigo, R. Kneissl, I. Toledo, A. Miyazaki, T. Takagi, A. Barnes, and F. Bonanomi. New methods for ALMA angular-scale based observation scheduling, quality assessment, and beam shaping II: refinements. *arXiv e-prints*, art. arXiv:2406.13199, June 2024. doi: 10.48550/arXiv.2406.13199.
- J. Pety. Successes of and Challenges to GILDAS, a State-of-the-Art Radioastronomy Toolkit. In F. Casoli, T. Contini, J. M. Hameury, and L. Pagani, editors, *SF2A-2005: Semaine de l’Astrophysique Française*, page 721, Dec. 2005.
- J. Pety, V. V. Guzmán, J. H. Orkisz, H. S. Liszt, M. Gerin, E. Bron, S. Bardeau, J. R. Goicoechea, P. Gratier, F. Le Petit, F. Levrier, K. I. Öberg, E. Roueff, and A. Sievers. The anatomy of the Orion B giant molecular cloud: A local template for studies of nearby galaxies. *A&A*, 599:A98, Jan. 2017. doi: 10.1051/0004-6361/201629862.

- A. Plunkett, A. Hacar, L. Moser-Fischer, D. Petry, P. Teuben, N. Pingel, D. Kunneriath, T. Takagi, Y. Miyamoto, E. Moravec, S. Suri, K. M. Hess, M. Hoffman, and B. Mason. Data Combination: Interferometry and Single-dish Imaging in Radio Astronomy. *PASP*, 135(1045):034501, Mar. 2023. doi: 10.1088/1538-3873/acb9bd.
- C. R. Purcell, S. N. Longmore, A. J. Walsh, M. T. Whiting, S. L. Breen, T. Britton, K. J. Brooks, M. G. Burton, M. R. Cunningham, J. A. Green, L. Harvey-Smith, L. Hindson, M. G. Hoare, B. Indermuehle, P. A. Jones, N. Lo, V. Lowe, C. J. Phillips, M. A. Thompson, J. S. Urquhart, M. A. Voronkov, and G. L. White. The H<sub>2</sub>O Southern Galactic Plane Survey: NH<sub>3</sub> (1,1) and (2,2) catalogues. *MNRAS*, 426(3):1972–1991, Nov. 2012. doi: 10.1111/j.1365-2966.2012.21800.x.
- J. M. Rathborne, J. M. Jackson, and R. Simon. Infrared Dark Clouds: Precursors to Star Clusters. *ApJ*, 641(1):389–405, Apr. 2006. doi: 10.1086/500423.
- J. M. Rathborne, J. S. Whitaker, J. M. Jackson, J. B. Foster, Y. Contreras, I. W. Stephens, A. E. Guzmán, S. N. Longmore, P. Sanhueza, F. Schuller, F. Wyrowski, and J. S. Urquhart. Molecular Line Emission Towards High-Mass Clumps: The MALT90 Catalogue. *PASA*, 33:e030, July 2016. doi: 10.1017/pasa.2016.23.
- M. J. Reid, K. M. Menten, A. Brunthaler, X. W. Zheng, T. M. Dame, Y. Xu, Y. Wu, B. Zhang, A. Sanna, M. Sato, K. Hachisuka, Y. K. Choi, K. Immer, L. Moscadelli, K. L. J. Rygl, and A. Bartkiewicz. Trigonometric Parallaxes of High Mass Star Forming Regions: The Structure and Kinematics of the Milky Way. *ApJ*, 783(2):130, Mar. 2014. doi: 10.1088/0004-637X/783/2/130.
- M. J. Reid, K. M. Menten, A. Brunthaler, X. W. Zheng, T. M. Dame, Y. Xu, J. Li, N. Sakai, Y. Wu, K. Immer, B. Zhang, A. Sanna, L. Moscadelli, K. L. J. Rygl, A. Bartkiewicz, B. Hu, L. H. Quiroga-Núñez, and H. J. van Langevelde. Trigonometric Parallaxes of High-mass Star-forming Regions: Our View of the Milky Way. *ApJ*, 885(2):131, Nov. 2019. doi: 10.3847/1538-4357/ab4a11.
- G. H. Rieke and M. J. Lebofsky. The interstellar extinction law from 1 to 13 microns. *ApJ*, 288:618–621, Jan. 1985. doi: 10.1086/162827.
- T. Robitaille, C. Deil, and A. Ginsburg. reproject: Python-based astronomical image reprojection. Astrophysics Source Code Library, record ascl:2011.023, Nov. 2020.
- E. F. Schlafly, G. Green, D. P. Finkbeiner, H.-W. Rix, E. F. Bell, W. S. Burgett, K. C. Chambers, P. W. Draper, K. W. Hodapp, N. Kaiser, E. A. Magnier, N. F. Martin, N. Metcalfe, P. A. Price, and J. L. Tonry. A Large Catalog of Accurate Distances to Molecular Clouds from PS1 Photometry. *ApJ*, 786(1):29, May 2014. doi: 10.1088/0004-637X/786/1/29.
- N. Schneider, T. Csengeri, S. Bontemps, F. Motte, R. Simon, P. Hennebelle, C. Federrath, and R. Klessen. Dynamic star formation in the massive DR21 filament. *A&A*, 520:A49, Sept. 2010. doi: 10.1051/0004-6361/201014481.

## Bibliography

- N. Schneider, T. Csengeri, R. S. Klessen, P. Tremblin, V. Ossenkopf, N. Peretto, R. Simon, S. Bontemps, and C. Federrath. Understanding star formation in molecular clouds. II. Signatures of gravitational collapse of IRDCs. *A&A*, 578:A29, June 2015. doi: 10.1051/0004-6361/201424375.
- F. Schuller, J. S. Urquhart, T. Csengeri, D. Colombo, A. Duarte-Cabral, M. Mattern, A. Ginsburg, A. R. Pettitt, F. Wyrowski, L. Anderson, F. Azagra, P. Barnes, M. Beltran, H. Beuther, S. Billington, L. Bronfman, R. Cesaroni, C. Dobbs, D. Eden, M.-Y. Lee, S.-N. X. Medina, K. M. Menten, T. Moore, F. M. Montenegro-Montes, S. Ragan, A. Rigby, M. Riener, D. Russeil, E. Schisano, A. Sanchez-Monge, A. Traficante, and others. The SEDIGISM survey: First Data Release and overview of the Galactic structure. *MNRAS*, 500(3):3064–3082, Jan. 2021. doi: 10.1093/mnras/staa2369.
- M. F. Skrutskie, R. M. Cutri, R. Stiening, M. D. Weinberg, S. Schneider, J. M. Carpenter, C. Beichman, R. Capps, T. Chester, J. Elias, J. Huchra, J. Liebert, C. Lonsdale, D. G. Monet, S. Price, P. Seitzer, T. Jarrett, J. D. Kirkpatrick, J. E. Gizis, E. Howard, T. Evans, J. Fowler, L. Fullmer, R. Hurt, R. Light, E. L. Kopan, K. A. Marsh, H. L. McCallon, R. Tam, S. Van Dyk, and S. Wheelock. The Two Micron All Sky Survey (2MASS). *AJ*, 131(2):1163–1183, Feb. 2006. doi: 10.1086/498708.
- L. E. Snyder and D. Buhl. Detection of several new interstellar molecules. *Annals of the New York Academy of Sciences*, 194(1):17–24, May 1972. doi: 10.1111/j.1749-6632.1972.tb12687.x.
- J. D. Soler, M.-A. Miville-Deschênes, S. Molinari, R. S. Klessen, P. Hennebelle, L. Testi, N. M. McClure-Griffiths, H. Beuther, D. Elia, E. Schisano, A. Traficante, P. Girichidis, S. C. O. Glover, R. J. Smith, M. Sormani, and R. Trefß. The Galactic dynamics revealed by the filamentary structure in atomic hydrogen emission. *A&A*, 662:A96, June 2022. doi: 10.1051/0004-6361/202243334.
- S. W. Stahler and F. Palla. *The Formation of Stars*. Viley-VCH Verlag, Weinheim, 2004.
- J. S. Stodólkiewicz. On the Gravitational Instability of Some Magneto-Hydrodynamical Systems of Astrophysical Interest. Part III. *ACTAA*, 13:30–54, Jan. 1963.
- J. Syed, J. D. Soler, H. Beuther, Y. Wang, S. Suri, J. D. Henshaw, M. Riener, S. Bialy, S. Rezaei Kh., J. M. Stil, P. F. Goldsmith, M. R. Rugel, S. C. O. Glover, R. S. Klessen, J. Kerp, J. S. Urquhart, J. Ott, N. Roy, N. Schneider, R. J. Smith, S. N. Longmore, and H. Linz. The “Maggie” filament: Physical properties of a giant atomic cloud. *A&A*, 657:A1, Jan. 2022. doi: 10.1051/0004-6361/202141265.
- B. Thomasson, I. Joncour, E. Moraux, F. Motte, F. Louvet, M. González, and T. Nony. A stochastic and analytical model of hierarchical fragmentation: The fragmentation of gas structures into young stellar objects in the interstellar medium. *A&A*, 689:A133, Sept. 2024. doi: 10.1051/0004-6361/202449649.

- A. Toomre. On the gravitational stability of a disk of stars. *ApJ*, 139:1217–1238, May 1964. doi: 10.1086/147861.
- B. E. Turner. U93.174: a New Interstellar Line with Quadrupole Hyperfine Splitting. *ApJL*, 193:L83, Oct. 1974. doi: 10.1086/181638.
- E. Vázquez-Semadeni, A. Palau, J. Ballesteros-Paredes, G. C. Gómez, and M. Zamora-Avilés. Global hierarchical collapse in molecular clouds. Towards a comprehensive scenario. *MNRAS*, 490(3):3061–3097, Dec. 2019. doi: 10.1093/mnras/stz2736.
- P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.
- T. V. Wenger, D. S. Balser, L. D. Anderson, and T. M. Bania. Kinematic Distances: A Monte Carlo Method. *ApJ*, 856(1):52, Mar. 2018. doi: 10.3847/1538-4357/aaaec8.
- J. S. Whitaker, J. M. Jackson, J. M. Rathborne, J. B. Foster, Y. Contreras, P. Sanhueza, I. W. Stephens, and S. N. Longmore. MALT90 Kinematic Distances to Dense Molecular Clumps. *AJ*, 154(4):140, Oct. 2017. doi: 10.3847/1538-3881/aa86ad.
- J. P. Williams, L. Blitz, and C. F. McKee. The Structure and Evolution of Molecular Clouds: from Clumps to Cores to the IMF. In V. Mannings, A. P. Boss, and S. S. Russell, editors, *Protostars and Planets IV*, page 97, May 2000. doi: 10.48550/arXiv.astro-ph/9902246.
- C. Zucker, C. Battersby, and A. Goodman. The Skeleton of the Milky Way. *ApJ*, 815(1):23, Dec. 2015. doi: 10.1088/0004-637X/815/1/23.
- C. Zucker, J. S. Speagle, E. F. Schlafly, G. M. Green, D. P. Finkbeiner, A. A. Goodman, and J. Alves. A Large Catalog of Accurate Distances to Local Molecular Clouds: The Gaia DR2 Edition. *ApJ*, 879(2):125, July 2019. doi: 10.3847/1538-4357/ab2388.



# A. Appendix

## A.1. Detailed Spitzer, $\text{N}_2\text{H}^+$ , and HNC Maps

The final maps in Fig. 4.1 show a *Spitzer* false-colour image together with the integrated intensity maps of  $\text{N}_2\text{H}^+$  and HNC. Nessie spans over  $2.5^\circ$  on the sky, which is reflected in the large extent of the image. The level of details this high-resolution maps provide cannot be fully appreciated at this scale. We therefore present three section details of Nessie in figures A.1 to A.3.

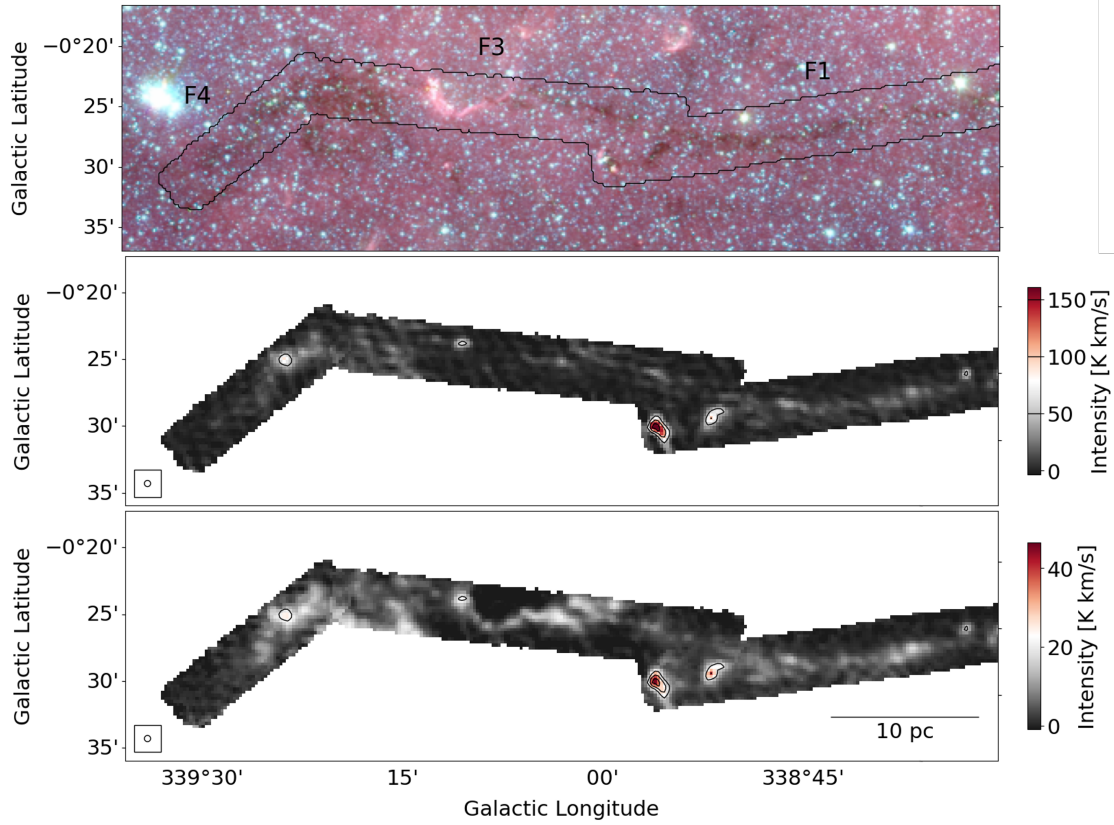


Figure A.1.: First section detail of Nessie. *Top*: *Spitzer* false-colour image, red at  $8\,\mu\text{m}$ , green at  $4.5\,\mu\text{m}$ , blue at  $3.6\,\mu\text{m}$ . *Middle*:  $\text{N}_2\text{H}^+$  Map, with contour lines at 50, 100, and  $150\,\text{K km s}^{-1}$ . *Bottom*: HNC map with  $\text{N}_2\text{H}^+$  contours.

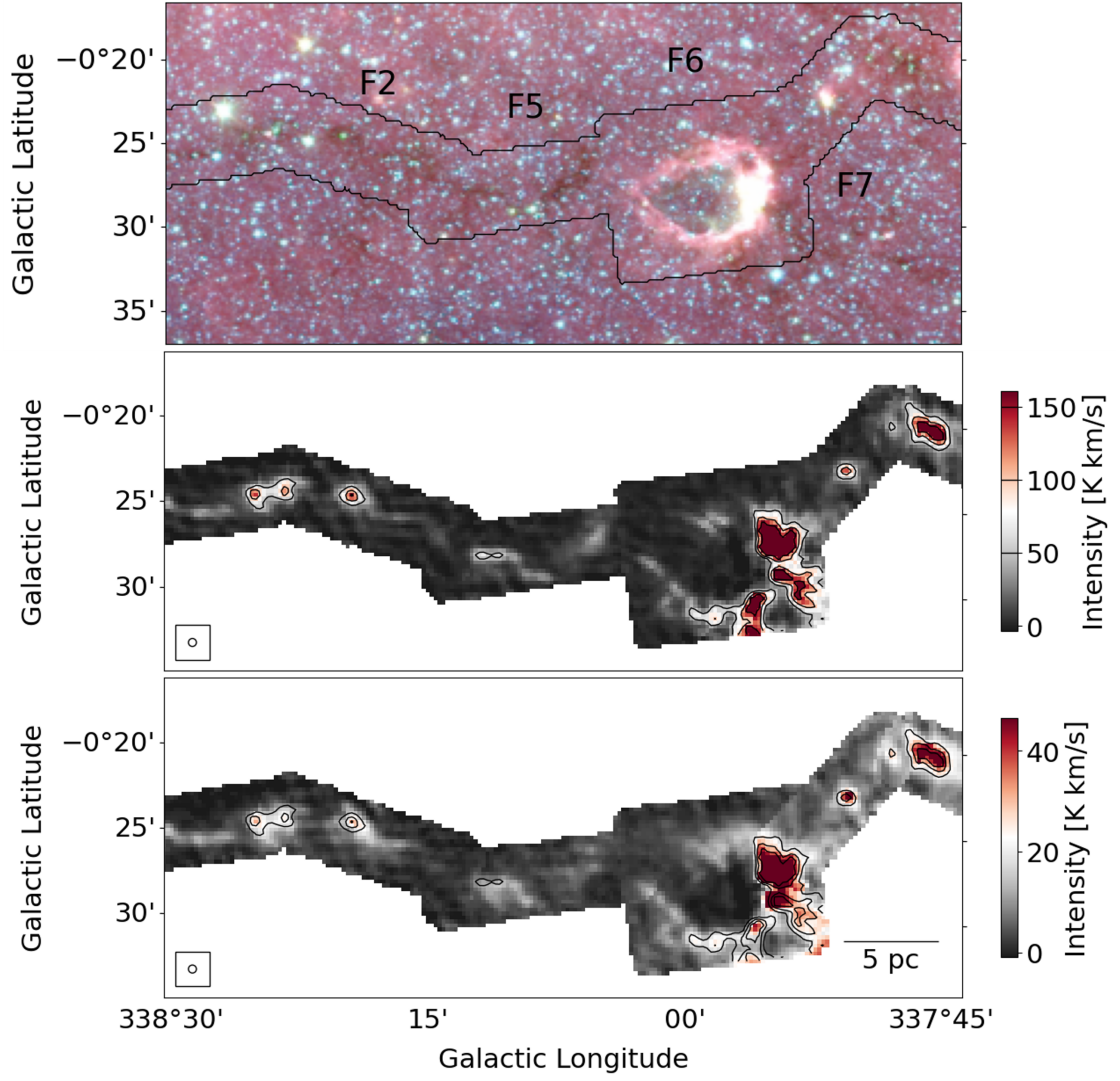


Figure A.2.: Second section detail of Nessie. *Top*: *Spitzer* false-colour image, red at  $8\,\mu\text{m}$ , green at  $4.5\,\mu\text{m}$ , blue at  $3.6\,\mu\text{m}$ . *Middle*:  $N_2H^+$  Map, with contour lines at 50, 100, and  $150\,\text{K km s}^{-1}$ . *Bottom*: HNC map with  $N_2H^+$  contours.



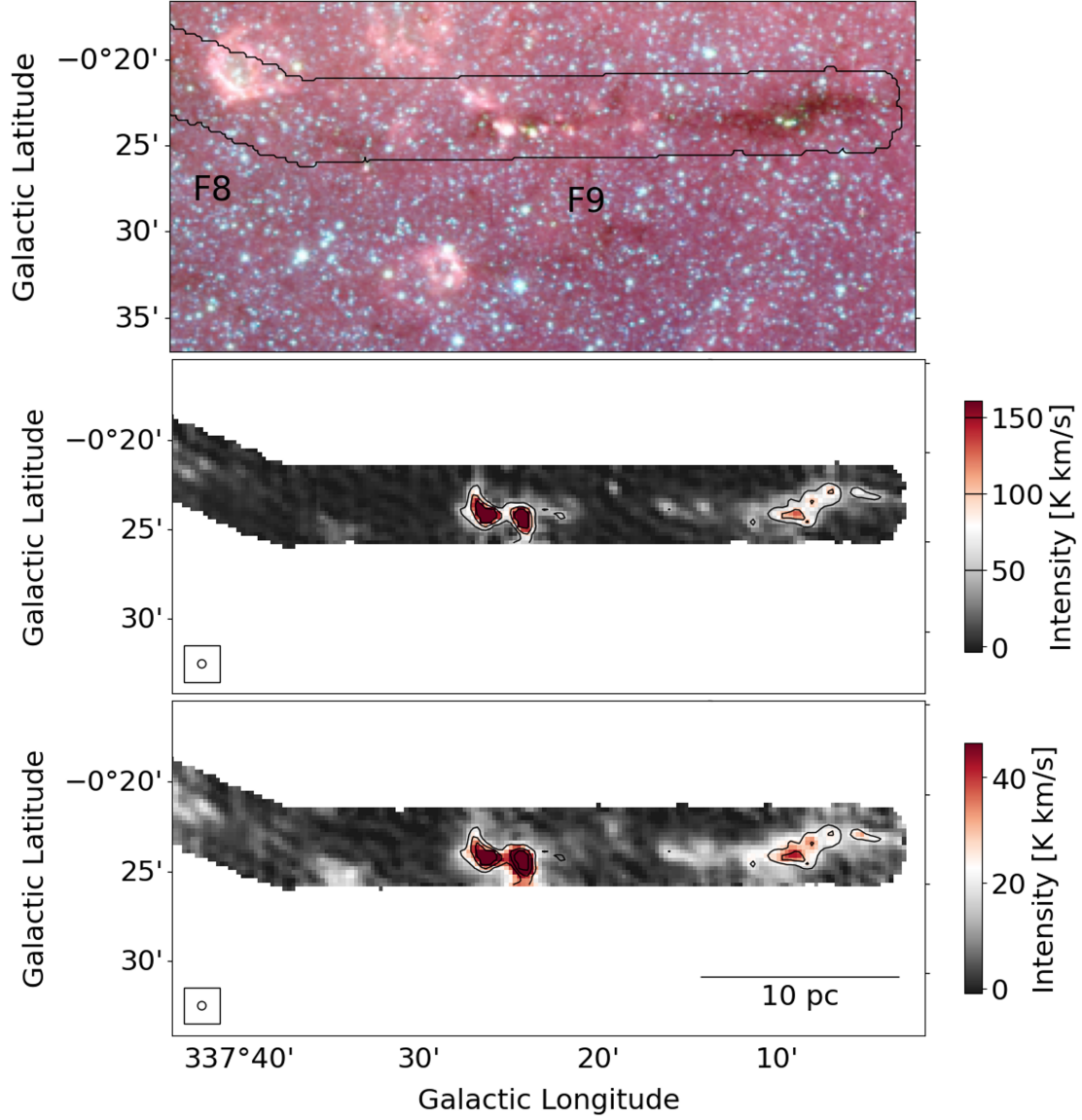


Figure A.3.: Third section detail of Nessie. *Top*: *Spitzer* false-colour image, red at 8  $\mu\text{m}$ , green at 4.5  $\mu\text{m}$ , blue at 3.6  $\mu\text{m}$ . *Middle*:  $N_2H^+$  Map, with contour lines at 50, 100, and 150  $\text{K km s}^{-1}$ . *Bottom*: HNC map with  $N_2H^+$  contours.

## A.2. Comparison of Cloud Placement with Goodman et al. 2014

To obtain the position angle  $\alpha$  needed to de-project the velocity centroids in Eq. 5.3, we placed Nessie in the Galactic plane, as sketched in Fig. 5.5. To ensure that this coordinate transformation from Galactic to Galactocentric ( $R, z, \varphi$ ) coordinates was performed correctly, we overlaid the result with the Galactic map of the Nessie neighbourhood from Goodman et al. (2014), see Fig. A.4. The agreement is satisfactory in the sense that, for a LOS distance of 3.1 kpc, Nessie aligns well with the structure reported by Goodman et al. (2014).

This comparison, however, serves only as a consistency check. While it shows that our reconstruction is consistent with values adopted in the literature, it tells nothing about the true three-dimensional structure of Nessie. The kinematic distance remains unclear, as shown in the comparison of literature values, and the overlap of our structure with Goodman et al. (2014) partly results from adopting the same LOS distance. If the bending of the filament were taken into account, a different geometry could emerge, with one end being potentially at a closer distance than the other. The actual distance and structure of Nessie therefore remain ambiguous and will most likely remain so until precise maser parallaxes are measured.

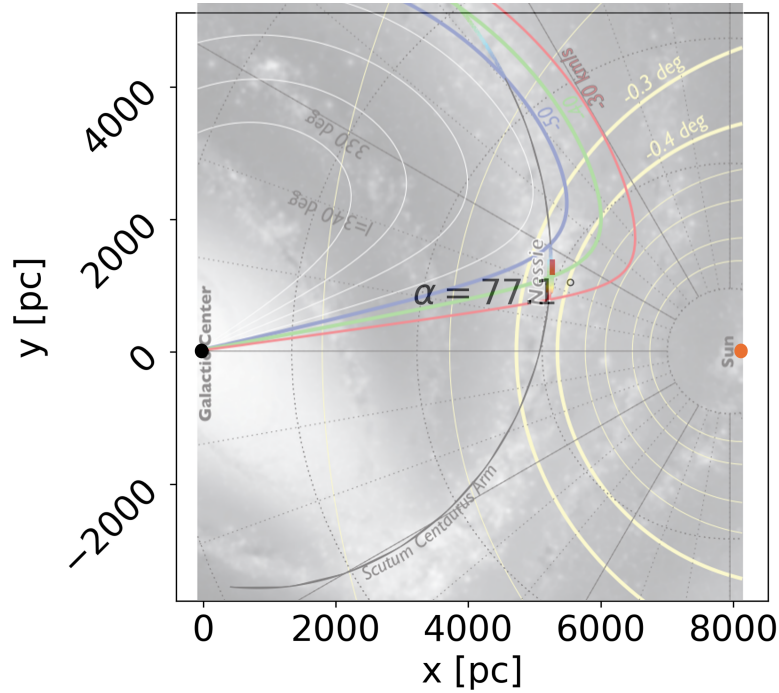


Figure A.4.: Nessie lies at a Galactocentric distance  $R_{gal}=5267$  pc and has a position angle of about  $77^\circ$  to the LOS. Our reconstructed filament (red dash) matches with the overlaid structure in Goodman et al. (2014).