The Atomic Interstellar Medium of M33

by

Eric William Koch

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Physics
University of Alberta

Abstract

I present new observations of 21-cm hydrogen line emission from the nearby spiral galaxy M33 (Triangulum) using the Jansky Very Large Array (VLA) in New Mexico. I have spent the majority of the past two years calibrating and imaging these new observations. This was a particularly time-consuming task due to the sheer volume of data, with the final data footprint reaching several terabytes in size. I developed an alternate approach for imaging massive interferometric datasets, which achieves a significant speed-up in computational time by splitting the data into minimally sized chunks. The analysis portion of this thesis focuses on the 21-cm spectral line of atomic hydrogen (H I), which traces the atomic medium. From this I determine the global properties of the atomic interstellar medium (ISM), fit a rotation curve to the galactic disk, and explore the relation between the atomic and molecular media, utilizing existing observations of the 2 to 1 rotational transition of carbon monoxide [CO(2-1)] line. I provide new analysis techniques for determining the co-location of CO(2-1) to the H I emission for spatially resolved studies of the extragalactic ISM. I find similarities in the spectral properties of a narrow ~ 10 km s^{-1} H I line component, likely arising from cool H I ($T \sim 100$ K), and the molecular medium traced by CO(2-1). This suggests they trace a similar kinematic portion of the

ISM. The average full-width-half-max (FWHM) of the H I filamentary network is 274.1 ± 8.8 pc and is a factor of 2.5 times brighter than a large-scale near-constant background component. The filamentary network is predicted to arise from the cool neutral medium ($T \sim 100\,\mathrm{K}$), while the large-scale component is consistent with the warm neutral medium ($T \sim 10^4\,\mathrm{K}$). Using a novel masking approach to trace the edges of bright H I, namely H I bubbles and wind-driven shells, I show that CO(2-1) emission is preferentially located a distance of $\sim$ 100 pc from the mask edges and not simply co-located with the brightest H I emission. This suggests that wind-driven shells may play a large role in the formation of molecular clouds in M33.

Preface

This thesis is original work by Eric William Koch. The telescope proposal to obtain the data used in this thesis was written and submitted by E.W. Rosolowsky. The single-dish Arecibo observations are published in Putman et al. (2009). The IRAM 30-m data of CO(2-1) emission is published in Druard et al. (2014).

Acknowledgments

Thank you to my supervisor, Erik Rosolowsky, for your constant encouragement, patience, and steady stream of interesting (and crazy) ideas. Your enthusiasm and passion for this subject make it a pleasure to work with you. I look forward to the next four years of research together and the crazy ideas that it will spawn.

Thank you to all members of the Astrophysics group for creating a wonderful work environment. All of you make it a joy to come to work everyday.

Thank you to my fellow Astrophysics graduate students for the excellent discussions I have had with so many of you, research related or not. It is a pleasure to hear your ideas and advice on so many topics — I have truly learned so much from all of you. And a special thank you to all of my office mates over the past two years for all of the ridiculous moments, of which a few particular ones will stick with me for the rest of my life.

A huge thank you to my parents and my sister for your love, constant support, and encouragement in each aspect of my life. Without your help, your lessons, and your inspiration to follow my dreams, I could not have possibly got to this point on my own.

And most of all, thank you to my wonderful Sarah. You are always there for me — to cheer for, to cheer up, and to drag me away from work when I get a little too fixated. Your constant love and support has made this process so much more enjoyable. I am glad we get to face these challenges together.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

I investigate the properties of the atomic interstellar medium (ISM) in M33 using new observations of the 21-cm neutral atomic hydrogen (H I) line. Using observations of the 2 to 1 rotational transition of carbon monoxide [CO(2-1)] from Druard et al. (2014), I explore the connection between the atomic and molecular media. The 21-cm H I line is a sensitive tracer of the neutral medium and is widespread throughout galactic disks. Since it directly traces the main component of the neutral medium — H I— the interpretation of physical properties from the line properties is rather straightforward, more-so than the proxy tracers of the molecular medium (e.g., CO).

The widespread emission across the galactic disk makes 21-cm H I emission an excellent tracer of galactic dynamics. As with all spectral lines, its Doppler shift from the rest frequency is a measure of the line-of-sight velocity. Thus it contains information pertinent to several areas of galaxy evolution: large-scale galactic dynamics like rotation and disk warping, inflows from the intergalactic medium cooling to atomic gas, the impact of small-scale disruptions like wind-driven shells, and the role of cool H I in forming H_2 (molecular hydrogen) that may subsequently host star formation. Thus high spatial and spectral observations of a galactic disk give a wealth of information on the galaxy as a whole and how the processes that occur within it shape its evolution.

1.1 Phases of the ISM

Interstellar space is pervaded by gas and dust referred to as the interstellar medium (ISM). The ISM is typically broken into *phases* to account for the massive range in its properties. These phases range from the most diffuse hot ionized medium, with temperatures exceeding 2×10^6 K resulting from sources like supernovae shocks, to the densest molecular medium, with temperatures of a few tens of Kelvin. Different numbers of phases are favoured by different models, ranging from three in the influential McKee and Ostriker (1977) model, to five or more from observationally driven models (e.g., Heiles, 2001). Here I list the properties of five of the phases, ignoring some of the more "specialized" phases (Heiles, 2001) that are not addressed in this thesis. I draw some of the physical properties of these phases from Heiles (2001) as their uncertainties are clearly pointed out in this article. All phases must fall into two broad classes: the ionized ISM and the neutral ISM, classified based on the dominant hydrogen phase at those conditions. The following list is arranged from the hottest/most diffuse, to the coldest/densest medium.

1. **Hot Ionized Medium (HIM)** — The HIM has the lowest density (0.003 cm^{-3}) and fills the largest volume fraction (~ 0.5) within the ISM. It is typically detected through low energy X-rays due to its high temperature ($\sim 10^7$ K). Around these temperatures, the peak of emission occurs around 0.25–0.75 keV, the low-end of which is severely attenuated at moderate hydrogen column densities (6×10^{19} cm^{-2}). As the gas cools, its emission becomes completely invisible within the ISM, and its properties must be studied through absorption lines in the UV. This gas reaches such high temperatures due to shocks from supernovae (SN), but is also found within super-bubble structures (the result of winds driven by clusters of massive stars; see §1.3). The blow-outs through the galactic disk from these energetic events typically funnel the HIM into the halo of the

galaxy (Tenorio-Tagle and Bodenheimer, 1988).

2. **Warm Ionized Medium (WIM)** — The WIM is significantly cooler than the HIM, with temperatures around 8000 K and typical densities between 0.1–0.4 cm^{-3}, about two orders of magnitude greater than the HIM. This phase accounts for most of the widespread Hα[1] emission throughout the Galactic plane (Hill et al., 2012), and is also evident within H II[2] regions. The emission from H II regions can have much higher electron densities, likely due to the emission being confined within an ionized bubble embedded in the denser, neutral ISM, as is the case with hypercompact H II regions around young massive protostars (Beuther et al., 2007).

3. **Warm Neutral Medium (WNM)** — The WNM accounts for roughly half of the 21-cm H I emission. Heiles and Troland (2003) find temperatures of the WNM range from 500–5000 K. Typical volume densities range between 0.1–0.4 cm^{-3}. This phase is the second most volume filling behind the HIM, with an approximate filling fraction of 0.3 and is the large-scale and widespread source of emission across galactic disks.

4. **Cool Neutral Medium (CNM)** — This is the dense cold counterpart of the WNM in the atomic medium. Heiles and Troland (2003) find temperatures ranging from 15–80 K for this component, with a median around 50 K within the nearby (< 500 pc) ISM. Due to the large temperature difference between the CNM and WNM, the CNM is typically observed in absorption against background sources, or as H I self-absorption (HISA) features. The CNM is expected to be localized to 20–40 pc regions, near the typical scales of giant molecular clouds, based on the highest resolution studies of extragalactic H I (e.g., Braun, 2012).

[1] This emission arises from the 3 to 2 transition of atomic hydrogen.
[2] H II is ionized atomic hydrogen.

5. **Molecular Medium** — The coldest and densest regions of the ISM are comprised of molecules, which are only capable of forming when the gas is shielded from ionizing ultraviolet (UV) photons within the interstellar radiation field[3] (ISRF). Collections of the molecular medium are referred to as molecular clouds, the largest of which, giant molecular clouds (GMCs), can have masses up to a few $\sim 10^6$ $M_\odot$ (Blitz et al., 2007). These exhibit temperatures up to several tens of Kelvin and densities of $\sim 10^3$ cm^{-3}. However, nearby molecular clouds show an intricate substructure where densities increase by several orders of magnitude on ~ 0.1 pc scales (André et al., 2014). The molecular component has the smallest filling fraction. As molecular clouds are expected to be enveloped within the CNM, their filling fraction is necessarily some fraction of the CNM's.

This thesis investigates the final three ISM phases on this list using observations of the 21-cm H I line (§1.2) and CO(2-1) (§1.4).

1.2 The 21-cm Hydrogen Line

The ubiquity of neutral hydrogen (HI) throughout the ISM makes the 21 cm hydrogen line the most readily detectable line emission throughout the Universe. As such, it is the most important tracer we have of the neutral medium. The following derivation draws from three sources I find to provide clear explanations: Stahler and Palla (2005), Rohlfs and Wilson (2004), and the *Essential Radio Astronomy* lecture notes by James Condon & Scott Ransom[4].

The 21 cm ($\nu = 1420.405$ MHz) line results from a hyperfine transition in hydrogen's ground state. The spin of the proton and electron each have two possible states. Due to the interaction of their magnetic moments, this

[3]Radiation pervades the ISM. The contribution of UV photons arises mostly from the most massive stars (referred to as OB stars).

[4]`https://science.nrao.edu/opportunities/courses/era/`

causes a change in the energy when the spins are parallel versus when they are anti-parallel. The emission coefficient for a magnetic dipole is

$$A_{mn} = \frac{64\pi^4}{3\hbar c^3}\nu_{mn}^3 \left|\mu_{\rm B}\right|, \tag{1.1}$$

where $|\mu_{\rm B}| = 9.27401 \times 10^{21}$ erg/Gauss is the magnitude of the Bohr magnetron. For the 21 cm line, $A \approx 2.85 \times 10^{-15}$ s^{-1}. The inverse of the emission coefficient is the half-life for radiative transitions; this yields a large half-life around 10^7 years for this transition, and an accordingly small critical density[5]. Hydrogen atoms are then typically collisionally excited into the higher energy state, and collisionally de-excited as well. The probability of radiative de-excitation, yielding a 21-cm photon, is small, but we measure an appreciable signal due to the abundance of hydrogen atoms (Stahler and Palla, 2005).

The relative populations between the states is described by the spin temperature $T_{\rm s}$:

$$\frac{n_{\rm u}}{n_{\rm l}} = \frac{g_{\rm u}}{g_{\rm l}} \exp\left(-\frac{h\nu}{kT_{\rm s}}\right), \tag{1.2}$$

where the subscripts u and l denote the upper and lower states, g is the statistical weight, and n is the number density of hydrogen in one of the states. The statistical weights of the spin states are $g_{\rm u} = 3$ and $g_{\rm l} = 1$. Since the constant term in the exponential is small ($h\nu/k = 0.0682$ K), and the typical temperature in the CNM is $\sim 50 - 150$ K,

$$\frac{n_{\rm u}}{n_{\rm l}} \approx \frac{g_{\rm u}}{g_{\rm l}} = 3. \tag{1.3}$$

In local thermodynamic equilibrium (LTE), three hydrogen atoms are in the excited hyperfine state relative to one in the lower state. This allows the general expression of the line absorption coefficient (see Equation 11.17 in Rohlfs and

[5]Defined to be where the rate of collisional processes match that of the radiative ones.

Wilson (2004)) to be written as

$$\kappa_\nu = \frac{3c^2}{32\pi}\frac{A_{10}n_{\rm H}h}{\nu_{10}kT_{\rm s}}\phi(\nu) \tag{1.4}$$

by approximating the exponential term with the first two terms of the Taylor series (Rohlfs and Wilson, 2004). $\phi(\nu)$ is the normalized line profile. The more commonly used measure of optical depth is defined as the integral of κ_ν over the path length s,

$$\tau_\nu(s) = \int_{s_0}^{s} \kappa_\nu(s){\rm d}s \tag{1.5}$$

If the hydrogen is assumed to be isothermal along the line of sight, typically a good approximation when viewing a galactic disk face-on, $\tau_\nu \propto T_{\rm s}^{-1}$.

Centimetre wavelengths are well into the radio regime, where the Rayleigh-Jeans approximation is valid. Brightness temperatures are commonly used in this regime, since they are equivalent to intensity through the Rayleigh-Jeans law,

$$T_{\rm b} = \frac{\lambda^2}{2k}I_\nu. \tag{1.6}$$

Solving the equation of radiative transfer in terms of the brightness temperature gives

$$T_{\rm b}(v) = T_{\rm s}\left[1 - e^{-\tau(v)}\right] + T_{\rm c}e^{-\tau(v)}, \tag{1.7}$$

where $T_{\rm c}$ is the brightness temperature of a background source. In the optically thin limit ($\tau \ll 1$) with no background source, $T_{\rm b} = T_{\rm s}\tau$. The line brightness temperature is then independent of the spin temperature of the emitting gas, since $\tau_\nu \propto T_{\rm s}^{-1}$, as shown above.

This independence allows the column density $N_{\rm H}$ to be directly computed from the line brightness, integrated over some spectral extent,

$$N_{\rm H} = 1.82 \times 10^{18} \int T_{\rm b}{\rm d}v \tag{1.8}$$

where units of cm^{-2}, K, and $\mathrm{km\,s}^{-1}$ are assumed for N_{H}, T_{b}, and $\mathrm{d}v$, respectively. It is then straightforward to convert the column density into a mass,

$$M = 0.34 D^2 \theta^2 \int T_{\mathrm{b}} \mathrm{d}v \tag{1.9}$$

where M is in units of $\mathrm{M}_{\odot}$, D is the distance in Mpc and θ is the beam width in arcseconds.

The optically thin case is valid for the WNM, due to its high temperature and moderate density. The assumption of optically thin H I emission is used throughout this thesis. The reasons for ignoring opacity effects are explained more thoroughly in §3.3 and are largely due to the lack of spatial resolution in the data. Quantities related to the mass are therefore lower-limits.

The line opacity for the 21-cm line is important for the CNM component (e.g., Braun, 1997). It is typically seen in absorption against background sources, but can also show self-absorption features. H I self-absorption (HISA) occurs when a collection of the CNM, typically called an H I cloud, is located in front of a bright WNM component (Gibson et al., 2005). The in-depth study of Heiles and Troland (2003) in the local ISM demonstrates how the spectrum along the line-of-sight can be decomposed into individual components to constrain the CNM and WNM properties. Since optical depth effects are not addressed in this thesis, I do not review them here.

1.3 The Atomic ISM

This thesis primarily investigates the properties of the atomic ISM in M33 and how they relate to recent results obtained in other nearby galaxies. The largest observational survey of the 21-cm line in nearby galaxies is THINGS - The H I Nearby Galaxy Survey (Walter et al., 2008). This survey used the VLA to observe 34 galaxies, out to a maximum distance of 15 Mpc, with $7''$ angular and 5 $\mathrm{km\,s}^{-1}$ spectral resolution. One galaxy from the set, NGC 5055, is shown

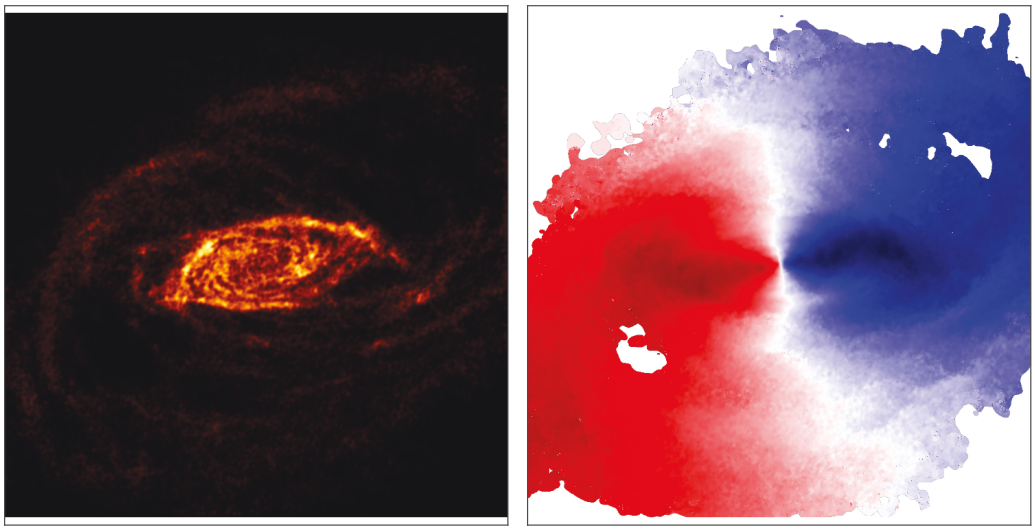


Figure 1.1: The zeroth moment (total emission; left) and centroid map (weighted average velocity; right) of NGC 5055 (Walter et al., 2008).

in Figure 1.1. The sample of galaxies was chosen to cover a wide range of star formation rates, masses, evolutionary stages and metallicity to provide a comprehensive set to investigate numerous issues on the role of the small-scale ISM, dark matter distribution, and the star formation process. The ability to compare galaxy properties in a homogeneous manner provides a crucial test of whether the observed variations from studies of single galaxies are reproducible over the entire set. Here I focus on results pertaining to H I only, and leave connections to the molecular medium to §1.4.

Rotation Curves — The large-scale morphology of 21-cm emission traces the overall shape of the galaxy. For spiral galaxies, this shape is a flat disk that is typically dominated by the spiral arms, or spiral arm fragments in the case of a flocculent spiral galaxy. Since 21-cm emission traces the entire disk, and is smooth and widespread, it is excellent for deriving the rotation curve of galaxies. The centroid map of NGC 5055 in Figure 1.1 clearly shows the rotation in the disk.

However, H I profiles often have structure that is non-Gaussian, either due to multiple components along the line-of-sight or significantly flattened peaks. Unresolved features may also show asymmetrical line profiles pointed towards the systematic velocity; this is the result of beam smearing. Using the THINGS galaxies that show significant rotation, de Blok et al. (2008) compare rotation curves derived using different estimates of velocity centres. These include fitting a single Gaussian, fitting multiple Gaussians, using the line centroid, the velocity at the peak line brightness, or fitting Hermite third-order polynomials. They prefer using the Hermite polynomials, which includes a term to account for skewness in the profile. However, galaxies with widespread H I emission show little difference in the rotation curve when using the line centroid — the weighted mean over the line profile — which does not require fitting. The connection between the Hermite coefficients and physical parameters is also uncertain (e.g., Stilp et al., 2013). Furthermore, the fits must be limited to profiles with fairly high S/N, and are not typically suitable for low-surface brightness features. All fitting approaches suffer from this issue. This is typically dealt with by smoothing the data, sometimes both spatially and spectrally, to reduce the noise level and focus on the large-scale variations in the rotation profile (Corbelli et al., 2014). However, this uncertainty may affect additional terms included to account for small-scale coherent variations in the velocity surface, such as radial flows.

Non-symmetric velocity features can be accounted for through higher-order terms in a harmonic decomposition, as was first introduced by Schoenmakers et al. (1997). Trachternach et al. (2008) performed a second order harmonic decomposition for the 19 THINGS galaxies examined by de Blok et al. (2008). They find that the effects of non-circular motions are small, typically about $\sim 10\%$ of the rotation velocity. They show that the amplitudes of these non-circular motions decreases in later Hubble types and lower luminosity galaxies. In certain systems, however, the non-circular motions are highly significant and

play an important role in the large-scale structure. For example, Colombo et al. (2014b) finds evidence of a third-order wave in the inner disk of M51, likely a result of the nuclear bar, which has a significant effect on the rotation velocity when not included in the model.

A typical approach to fitting a rotation curve is to split the velocity surface into a series of tilted-rings, which are independently fit for the rotation model, the centre, and the position angle and inclination. This is useful when changes in the disk parameters — the centre, inclination and position angle — are expected to radially vary across the galaxy. There are significant covariances, particularly the inclination, when allowing these parameters to each be free and independent in each ring. Most studies that use this approach perform a tilted-ring fit in two stages, where the second uses some form of regularized or smoothed parameterization of the inclination and position angle (e.g., Colombo et al., 2014b; Corbelli et al., 2014). Alternatively if significant deviations are not expected over the region to be fit, the entire disk can be fit simultaneously to provide better estimates of the galactic disk parameters, while using the ring approach to derive the rotation velocity, and other velocity components, only. This is the approach introduced by Spekkens and Sellwood (2007) and Sellwood and Sánchez (2010) for the DISKFIT algorithm. Sellwood and Sánchez (2010) finds that this approach provides better fit results of the asymmetrical velocity components when compared to the analysis by Trachternach et al. (2008) for five of the THINGS galaxies.

Velocity Dispersion and H I Phases — Observed line widths can be attributed to a combination of thermal broadening and turbulence. The atomic gas, namely H I, is expected to have two temperatures at which it can be in a stable thermal equilibrium (Wolfire et al., 2003); this gives rise to the observed and predicted CNM and WNM phases (§1.1). If this model holds, the two components should be in a pressure equilibrium. The thermal line widths of

these component are expected to be ~ 1 km s^{-1} and ~ 8 km s^{-1} for the CNM and WNM, respectively. When lines are observed to be broader than ~ 8 km s^{-1}, the excess is typically attributed to turbulent motion in the medium or smearing from the beam size.

Tamburro et al. (2009) investigated the source of H I velocity dispersion in a portion of the THINGS galaxies. They find a characteristic value of 10 ± 2 km s^{-1} within the optical radii of the galaxies, above the ~ 8 km s^{-1} expected from thermal effects alone. The excess is attributed to turbulence resulting primarily from supernova explosions (SNe). However, the cooling timescale in the ISM ($\sim 10^3$ yr; Wolfire et al., 2003) is significantly shorter than the expected turbulent dissipation timescale (~ 10 Myr; Mac Low, 1999), implying that energy must be continuously replenished to maintain the level of turbulence. Radial profiles of the velocity dispersion show a decrease with galactic radius in all of the samples. Based on the slope of these profiles, they argue that the velocity dispersion in the outer disk cannot be maintained by SNe injected turbulence, since the low levels of star formation cannot maintain the required SNe rate. They suggest magnetorotational instabilities (MRI; Balbus and Hawley, 1991) as an alternative energy source for driving turbulence in the outer disks. A number of other potential turbulent energy sources are discussed in Elmegreen and Scalo (2004).

A further study by Ianjamasimanana et al. (2012) examined the shapes of H I velocity profiles by utilizing a spectral stacking technique to create rotation-subtracted spectra for large regions of the THINGS galaxies. They find that these profiles are well represented by a two-Gaussian model — a wide and a narrow line width component. Two components are expected for the H I distribution, as it contains two unique phases — the CNM and WNM (§1.1). The width of these components varies, from $3.4 - 8.6$ km s^{-1} for the narrow component and $10.1 - 24.3$ km s^{-1} for the wide component. The authors find that the ratio in total emission explained by the components is $A_n/A_w \sim 0.63 \pm 0.24$,

implying that the wide component is dominant on large-scales. A correlation is found with the ratio between the surface density of molecular hydrogen (from CO) and H I ($\Sigma_{\rm H_2}/\Sigma_{HI}$), and A_n/A_w. Assuming that the molecular medium is created from the CNM, they argue that the narrow component results from the CNM.

Ianjamasimanana et al. (2015) examined the radial dependence of the results of Ianjamasimanana et al. (2012). They find similar results to Tamburro et al. (2009) — the velocity dispersion in the outer disks requires an alternative energy source from SNe. For the entire sample, they find an average width ratio of 0.46 ± 0.08 between the narrow and wide component. The ratio A_n/A_w is found to decrease with increasing radius, which the authors interpret as a lack of the CNM in the outer disk. This fraction is also lower in the dwarf galaxies in their sample relative to the spiral galaxies. A lack of the CNM phase may explain the typical low star formation rates observed in dwarf galaxies. Where the dwarf galaxies do have a clear narrow component, the width is consistent with the narrow component in the spiral galaxies, suggesting that conditions in the CNM environment may be similar across a range of galaxy types.

H I Morphology: Bubbles and Shells — The large-scale structure of the ISM is often dominated by the spiral structure of galaxies. On smaller scales, however, the ISM becomes highly irregular, garnering apt comparisons such as the "cosmic bubble bath" (Brand and Zealey, 1975). Names for the structures identified in this chaotic menagerie include shells, rings, holes, loops, bubbles, filaments, superbubbles, supershells, and worms, to name a few. All of these refer to regions that shows a distinct lack of interstellar material or the edge of regions where a higher density of material is observed. An example of a bubble-dominated region is shown in Figure 1.2.

The broadest definition of an H I hole is a region that persists at multiple velocities whose interior brightness is significantly lower than its surroundings.

Typically these features appear spatially as mostly complete circular or elliptical shapes, and are found to be expanding (Hu, 1981). These holes are often found surrounded by a shell of bright H I, likely swept up in the hole's expansion. This description is based on the appearance in H I; looking at other wavebands can reveal a correlated, but different morphology. H I holes have been found to surround shells of Hα emission, and/or be filled with UV and soft X-ray emission (e.g., Deul and den Hartog, 1990). Not all holes are filled with hot gas, however. For example, kpc-scale holes in the dwarf galaxy IC 1613 are remarkably empty, devoid even of a significant stellar component (Silich et al., 2006). This is a strong indication that holes up to a few hundred pc are the result of SN explosions or winds driven by massive stars, and kpc-scale holes are formed through another mechanism. Holes and shells with clear signatures of recent star formation (i.e., hot/warm gas) can be broadly classified as *wind-driven shells* (WDS), referring to the source arising from a combination of stellar winds, H II regions, and SN explosions.

For this thesis, the link between the H I shells and the location of GMCs is of particular interest. Multiple studies have found CO emission on the rims of these shells (e.g., Yamaguchi et al., 2001; Leroy et al., 2006; Dawson et al., 2011) suggesting that the compression of the neutral ISM in the shell triggers the formation of H_2 and molecular clouds. The shells shown in Figure 1.2 are an example where compression between multiple bubbles is occurring. This is consistent with the theory of Elmegreen et al. (2002), who suggest that the compression in the shell leads to a self-regulating cycle of propagating star formation. A similar, though more incomplete, idea of self-propagating star formation is outlined in Tenorio-Tagle and Bodenheimer (1988). If this is the dominant mode of molecular cloud formation throughout a galaxy, then CO should be preferentially detected near the edges of bubbles.

This compression is dependent on the overall *porosity* of the ISM. Porosity refers to the relative volume of the ISM filled by voids (Cox and Smith, 1974).

This idea is used in the McKee and Ostriker (1977) three-phase model, since the dominant form of energy input into the ISM is taken to be from supernovae. The porosity of the medium affects the spread of energy within it. Bubbles rarely interact in a low porosity medium, leaving their energy trapped within. A highly porous ISM represents a large connection of low-density bubbles and voids, whose energy is distributed more effectively through this network. More energy would then be expected to escape out of the galactic disk and into the halo. From an observational stand-point, the porosity affects how frequently "squeezing" between H I shells can occur, like those found by Leroy et al. (2006) in IC 10. The enhanced compression provides more atomic material capable of a phase transition into molecular clouds, and taking the large-scale star formation view (§1.4), has a direct effect on the resulting star formation rate.

H I Morphology: Filaments — The narrow bright H I emission is often referred to as *filamentary*, or as a *high-brightness network* (HBN; Braun, 1997). Many spiral galaxies have bright H I features arranged into similar networks, typically arising along spiral arm features. Braun (1997) finds that 60–90% of the total H I flux arises from the HBN in 11 nearby spiral galaxies. This fraction drops significantly beyond the optical disk, where H I is expected to dominate (e.g., Schruba et al., 2011). Within the HBN, the line profiles tend to be peaked, with low level extended wings. This is similar to the somewhat more resolved line profiles of Ianjamasimanana et al. (2012) and Ianjamasimanana et al. (2015), though Braun (1997) favours using a Lorentzian profile. The "peakiness" (i.e., low kurtosis, assuming a Gaussian line profile) is perhaps exaggerated in the profiles found by Braun (1997) due to the severe data masking that was applied on the low signal-to-noise data. He finds that the HBN is dominated by high-opacity H I emission with kinetic temperatures of 80–200 K; low-brightness emission is consistent with arising from the WNM with a temperature of ~ 8000 K.

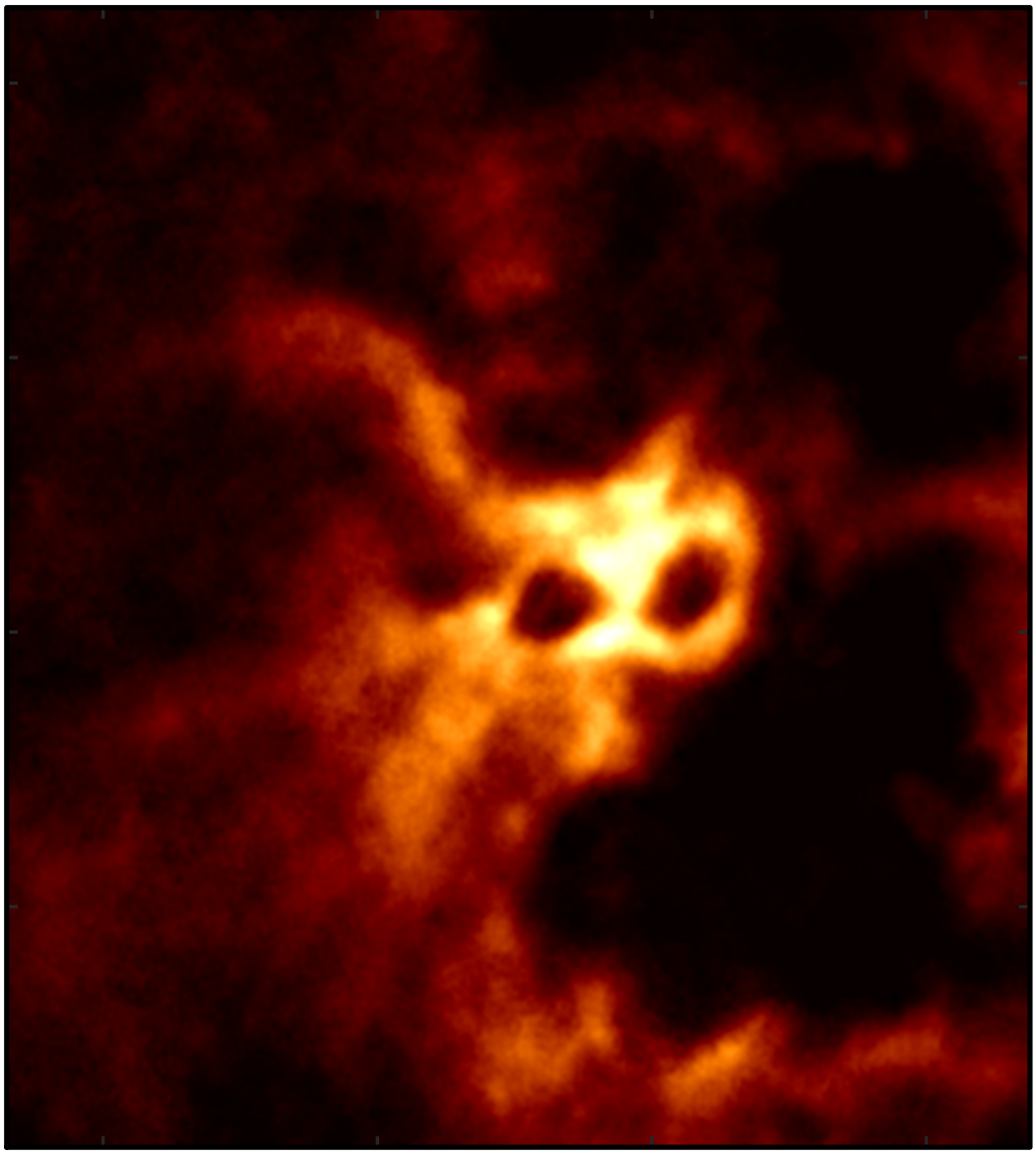


Figure 1.2: A bubble-dominated region of the dwarf galaxy IC 1613 (Hunter et al., 2012). Silich et al. (2006) find several clusters of massive stars (i.e., OB associations) consistent with the locations of the bubbles.

A recent survey by Patra et al. (2016) examined tracers of the CNM in faint dwarf galaxies. They find that decomposing spectra into a multi-Gaussian model reveals a smaller amount of cold H I than that found within the HBN, often by ~ 3 orders of magnitude. They identify any Gaussian spectral component with a line width less than 6 $\mathrm{km\,s^{-1}}$ as arising from cold H I, making direct comparisons to the stacking results of Ianjamasimanana et al. (2012) uncertain. They find that 70% of H I in the HBN is associated with recent star formation, based on Hα emission.

Systems of dark filaments — spatially coherent structures of HISA — have been noted in the Milky Way (Gibson et al., 2005) and in M31 (Andromeda; Braun et al., 2009). Gibson et al. (2005) identified HISA using the Canadian Galactic Plane Survey. They find coherent structures with an angular length of $\sim 5°$. At a distance of ~ 2 kpc, these are filaments with a length of 175 pc and may be a component of the Perseus spiral arm. Within M31, Braun et al. (2009) find kpc long complexes that appear to preferentially trace the leading edges of spiral arms. This may indicate the location of recently compressed H I by the passing of the shock front of the arm, and regions where molecular cloud formation can be triggered (Dobbs et al., 2014). Thilker et al. (2002) finds consistent preliminary results of these features in M33.

1.4 Forming the Molecular ISM

Conditions within the densest regions of the CNM allow for molecules to form. Primarily, this is a column density argument: the atomic medium requires shielding from the photoionizing far ultraviolet (FUV) photons in the ISRF to form molecules. The H I-H_2 conversion process, what induces it, and the timescales involved remain topics of active research.

Our understanding of the relation of the CNM and molecular medium to star formation largely comes from the THINGS (Walter et al., 2008) and HER-

ACLES (Leroy et al., 2009) surveys. As was also stated in §1.3, the work resulting from these studies provides a uniform treatment of the data over a wide range of galaxy types.

Several pieces of observational evidence suggest that the formation of H_2, and hence molecular clouds, is the bottleneck in the star formation process. This argument revolves around the Kennicutt-Schmidt Law (K-S Law; Schmidt, 1959; Kennicutt, 1989), which relates the surface density of the ISM ($\Sigma_{\rm gas}$) to the surface density of the star formation rate ($\Sigma_{\rm SFR}$), found using a tracer of recent star formation (e.g., Hα, FUV, and 24 μm);

$$\Sigma_{\rm SFR} \propto \Sigma_{\rm gas}^{1.4\pm0.2}. \tag{1.10}$$

The original work of Kennicutt (1989), and updated studies (e.g., Kennicutt, 1998; Leroy et al., 2008), show that there is more star formation where there is more molecular gas. This finding anchors the entire top-down view of star formation used in extragalactic observations where the physical resolution corresponds to near kpc scales. The phase of the ISM can be used to split this relation into two parts, the surface density of the H_2 and H I.

Bigiel et al. (2008) performed this separation on 18 nearby galaxies. They find that a powerlaw index of 1.0 ± 0.2 relates the surface density of star formation to that of H_2 alone. This result was found to be robust to different binning approaches and the choice of star formation tracer used. The authors interpret this result to infer that star formation in GMCs has approximately uniform properties. If this is the case, $\Sigma_{\rm H2}$ represents a measure of the volume filling factor of GMCs rather than variations within the GMC properties. There is no significant correlation of $\Sigma_{\rm SFR}$ with $\Sigma_{\rm HI}$. This lack of correlation is demonstrated in Figure 1.3, where tracers of the star formation rate correlate with the CO(2-1) emission in the inner disk, but not significantly with the H I in the outer disk. The resulting star formation efficiency (SFE), $\Sigma_{\rm SFR}/\Sigma_{\rm gas}$,

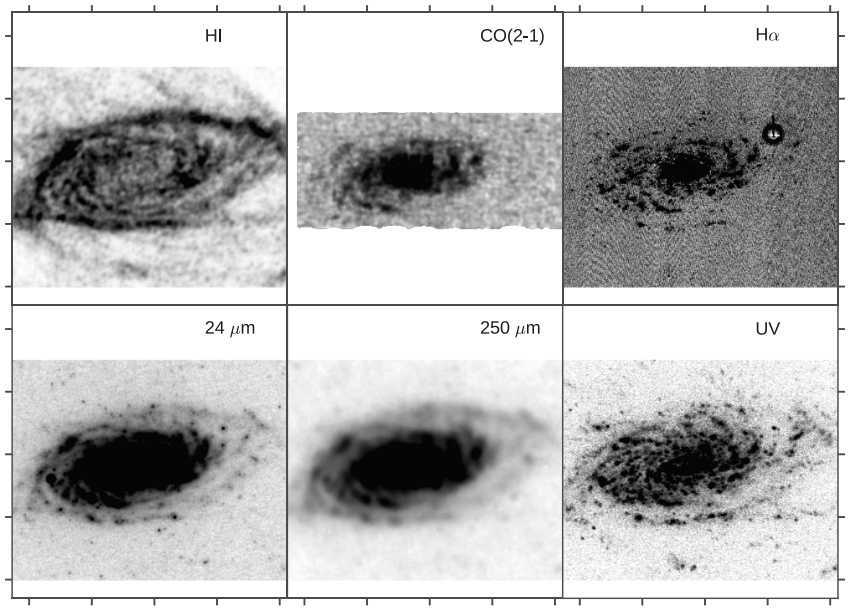


Figure 1.3: Tracers of the gas (H I, CO(2-1)) and star-formation (Hα, 24 μm, 250 μm, UV) for the spiral galaxy NGC 5055. Each image has been reprojected to match the H I map. The star-formation tracers match the emission regions of CO(2-1) better than the H I emission. The H I disk extends beyond what is shown (see Figure 1.1). The data is from the following surveys: H I— THINGS (Walter et al., 2008), CO(2-1)— HERACLES (Leroy et al., 2009), Hα — (Regan et al., 2001), 24 μm — SINGS (Kennicutt et al., 2003), 250 μm — KINGFISH (Kennicutt et al., 2011), UV — GALEX (Gil de Paz et al., 2007).

varies across the sample, showing that the SFE may be set by the local environment. Spiral galaxies show a radial decline in SFE, and the dwarf galaxies show similar properties to the outer disks of the spirals. H I surface densities are shown to saturate at ~ 9 $M_\odot\,pc^{-2}$, above which the molecular component dominates.

Furthering this result, Schruba et al. (2011) examined the star formation tracers in the outer regions of galactic disks where the atomic ISM is the dominant component. This tested whether the lack of dependence of Σ_{HI} continues into regions where only isolated CO(2-1) emission is found (Leroy et al., 2009). They find the same result: star formation is highly correlated with the molec-

ular gas, even when it is not the dominant phase in the ISM. This strongly implies that the fundamental relation is the $\Sigma_{\rm SFR} - \Sigma_{\rm H2}$ relation. The galactic star formation is suggested to occur in two steps: (1) assembly of molecular clouds; and (2) star formation from H_2.

Finally, the most inclusive study using the HERACLES and THINGS data is presented in Leroy et al. (2013b). They too find that the fundamental relation in the K-S Law is SFR and $\Sigma_{\rm H2}$, and that there are important second order effects — such as the metallicity — that play an important role in the H I dominated areas. The powerlaw here is found to follow an index of 1.0 ± 0.15, consistent with the results of Bigiel et al. (2008). Further evidence of the impact of local environmental effects on star formation is demonstrated. Real enhancements in $\Sigma_{\rm SFR}$ are seen in "starburst" environments relative to the normal H_2 surface density found for "disk" star formation.

These studies all suggest that understanding the star formation process on large scales hinges on understanding of molecular cloud formation. There are two aspects to this process: (1) how does H_2 form in the ISM, and (2) which physical mechanism within the ISM is responsible for creating the conditions needed for H_2 formation. The direct relevance for this thesis is point (1), which is expanded on below. Both points are thoroughly discussed in the recent review by Dobbs et al. (2014).

Formation of H_2 — Understanding how H_2 forms provides important constraints on the timescale to create molecular clouds. Forming H_2 in gas phase reactions is rather inefficient — the straightforward reaction of combining two hydrogen atoms into a molecule is hindered by its symmetry, since no allowed radiative transitions occur in the ground state to carry away the binding energy (Krumholz, 2014). Another gas-phase reaction is available, though it requires a population of free electron and protons to create an initial anionic or cationic hydrogen atom. The supply of free electrons in typical Milky Way ISM condi-

tions severely limits this from occurring (Wolfire et al., 2003). Instead, the bulk of H_2 is assumed to have been formed on the surface of dust grains, since they act to catalyze H_2 formation, carrying away the binding energy in formation through a vibrational lattice (Krumholz, 2014).

In understanding how H_2 exists in the ISM, it is necessary to discuss its destruction and how that sets the observed balance. The dominant destruction process is through photodissociation by FUV photons in the ISRF (§1.1). But the symmetry of H_2 again blocks the obvious destruction pathway. The binding energy of H_2 is 4.5 eV, but transitions where the molecule is directly photodissociated requires one of the hydrogen atoms to be excited. This requires a photon with an energy greater than 14.5 eV (Krumholz, 2014). Photons of this energy and greater can ionize neutral hydrogen (whose binding energy is 13.6 eV) and so are typically absent in the ISRF. The dominant mechanism is instead a two-step process, where the H_2 molecule is first excited into a higher electronic state. These transitions are in the Lyman-Werner band — the first and second excited states — with energies in the range of 11–13.6 eV. When in this excited state, there is a finite probability that the spontaneous decay into the ground state will unbind the molecule.

Krumholz (2014) shows that the equilibrium fraction between H_2 and H I, for Milky Way values, is

$$\frac{n_{\rm H2}}{n_{\rm HI}} = 8 \times 10^{-6} \left(\frac{4 \times 10^{-11}\,{\rm s}^{-1}}{\zeta_{\rm diss}} \right) \left(\frac{n_{\rm H}}{10\,{\rm cm}^{-3}} \right). \tag{1.11}$$

where $\zeta_{\rm diss}$ is the photodissociation rate set by the ISRF and $n_{\rm H}$ is the number density of free hydrogen atoms. The H_2 requires high densities or a low photodissociation rate to become dominant. Further adopting a simple slab geometry and assuming a sharp transition, Krumholz (2014) shows that the

surface density of H I depends only on the total rate coefficient $\mathcal{R}_{\rm gr}$

$$\Sigma_{\rm HI} \approx 9 \left(\frac{\mathcal{R}_{\rm gr}}{\mathcal{R}_{\rm gr,MW}} \right) {\rm M}_\odot \, {\rm pc}^{-2}, \tag{1.12}$$

where $\mathcal{R}_{\rm gr,MW} = 3 \times 10^{-17}$ cm^3 s^{-1} is appropriate for the formation rate of H_2 in the Milky Way. This explains the saturation point of $\Sigma_{HI} \sim 10-15$ M$_\odot$ pc^{-2} demonstrated across a wide range of galactic environments (e.g., Bigiel et al., 2008; Schruba et al., 2011; Leroy et al., 2013b). The assumptions in Equation 1.12 break down at low metallicities (~ 0.05 Z$_\odot$), where the dust population is dramatically decreased (Wolfire et al., 2003).

The slab geometry assumed above is rather crude, however analytic solutions quickly become unfeasible for other types of simple geometries. Krumholz et al. (2008) and Krumholz et al. (2009a) derive a semi-analytical solution for the fraction of H_2 for a spherical cloud of H_2, embedded within a layer of H I. The entire cloud is submersed inside an ISRF. They find that the fraction of molecular gas is controlled by the relation between the ISRF and molecular self-shielding. They express this relation as

$$f_{\rm H2}(\Sigma_{\rm comp}, Z') \approx 1 - \left[1 + \left(\frac{3s}{4(1+\delta)} \right) \right]^{-1/5} \tag{1.13}$$

where $s = \ln(1+0.6\chi)/(0.04\Sigma_{\rm comp,0} Z')$, $\chi = 0.77(1+3.1Z'^{0.365})$, $\delta = 0.0712(0.1s^{-1}+0.675)^{-2.8}$, and $\Sigma_{\rm comp,0} = \Sigma_{\rm comp}/1$ M$_\odot$ pc^{-2} is the surface density averaged over a 100 pc scale. The metallicity Z' is normalized to solar metallicity. These values come from Krumholz et al. (2008) and Krumholz et al. (2009a), though the approximation changes slightly in Krumholz et al. (2009b). The cloud is assumed to be uniform and layered, such that middle is dominated by the molecular gas and the outer region is an atomic envelope. The entire modeling builds upon the semi-analytic model of Wolfire et al. (2003), who has shown the assumptions made break down at metallicities lower than $Z' = 0.05$. The parameter $\Sigma_{\rm comp}$ is the surface density of a $\sim$100 pc atomic-molecular complex.

Such a definition is ill-defined observationally, so Krumholz et al. (2009b) introduces a clumping factor[6] C defined by $\Sigma_{\rm comp} = C\Sigma_g$. Σ_g represents a large-scale average of the gas surface density on ∼kpc scales, as is appropriate for most extragalactic observations. As the resolution of the data reaches 100 pc, this clumping factor should converge to one, as the model is defined at a 100 pc scale.

Schruba et al. (2011) showed that using large-scale radial bins to find the H_2-H I ratio provides a decent match to the Krumholz et al. (2009b) model with a clumping factor of 1–2. However there is up to an order of magnitude scatter about the model. Variations in the metallicity were ignored, which may explain some of this scatter.

Observational results based on CO(2-1) require a caution, however. H_2 is not directly detectable in molecular cloud conditions; it lacks rotational transitions at low temperatures due to its symmetry. Instead, it is most common to use CO as a proxy for H_2, since the former is an asymmetrical simple rotor with rotational levels that can be excited at T $\sim$ 15K. The largest issue with inferring H_2 properties from CO is the difference in their chemistry. CO is more easily photodissociated than H_2, so the outer regions of molecular clouds are "CO-dark". A detailed study of the nearby Perseus molecular cloud by Lee et al. (2012) found that 30% of H_2 is CO-dark. This implies that the entire population of H_2 is not traced by CO(2-1), and a population of CO-dark molecular gas would flatten the rapid radial drop of $\Sigma_{\rm H2}$ observed by Schruba et al. (2011). Analytical modeling by Krumholz et al. (2011), using the model of Krumholz et al. (2008) & Krumholz et al. (2009a), suggests that star formation is highly correlated with H_2 and not the total gas mass or CO mass. The fraction of H_2 is found to be a strong function of metallicity, implying that galaxies with low metallicity will have larges masses of H_2 where there is little CO. Star formation rates per unit CO mass should be quite large in low

[6]The capital is used here to distinguish from the Leroy et al. (2013a) clumping factor.

metallicity galaxies if this is the case.

Carbon chemistry is an important parameter in the ISM, beyond its relation to inferring H_2. The large level spacings in H_2 make it a poor coolant, so the majority of cooling in molecular clouds is through the low rotational levels of CO (Krumholz, 2014). Where CO cannot exist in the peripheries of molecular clouds, the 158 μm line in neutral C is the main coolant line (Hollenbach and Tielens, 1999). This cooling plays the vital role in setting the temperature of molecular clouds. Krumholz (2014) shows that CO line cooling balances the dissociation from cosmic rays, which at their expected value in the Milky Way, gives an equilibrium temperature of ~ 15 K.

H I and CO Velocity Dispersion — Important clues on the relation of the molecular and atomic media can be discerned from their relative velocity dispersions. Caldu-Primo et al. (2013) used the HERACLES data to find the velocity dispersion of CO(2-1) adopting a similar stacking technique used in Schruba et al. (2011). They perform the same analysis on H I data from THINGS, using single Gaussian models instead of the two Gaussian model favoured by Ianjamasimanana et al. (2012)and Ianjamasimanana et al. (2015). They find that the line widths are comparable, with averages of 11.9 ± 3.1 km s^{-1} for H I and 12.0 ± 3.9 km s^{-1} for CO(2-1). These results do not change if alternate stacking techniques are used; the ratio remains at $\sigma_{\rm HI}/\sigma_{\rm CO} = 1.0 \pm 0.2$. The CO line width is twice the value expected for a cold molecular disk dominated by GMCs. The authors interpret their results to imply that a thick diffuse molecular disk component must exist. The notion of this "diffuse" component is molecular gas in small unresolved clumps not associated with GMCs. Evidence of a diffuse component is quite strong in the spiral galaxy M51. Pety et al. (2013) finds that $\sim 50\%$ of the emission arises from scales larger than 1.3 kpc in the PAWS survey data (Schinnerer et al., 2013). They suggest that a thick diffuse disk of molecular gas is sub-structured into a network of unresolved filaments

that have an expected filling fraction of $\sim 0.1\%$. Such a network would be associated with an accompanying filamentary network of cold H I (§1.3).

An expanded study by Mogotsi et al. (2016) reviewed and expanded the results of Caldu-Primo et al. (2013) by examining individual profiles instead of ones stacked over large regions. This study rejected all profiles that did not have a peak line flux greater than four times the rms noise. They find smaller CO line widths on average, $\sigma_{\rm CO} = 7.3 \pm 1.2$ km s^{-1}, which implies a ratio between the line widths of $\sigma_{\rm HI}/\sigma_{\rm CO} = 1.4 \pm 0.2$. This ratio between the line widths decreases when repeating the Caldu-Primo et al. (2013) stacking analysis with the increased noise thresholds, implying that higher velocity dispersions occur in low-brightness CO profiles and are not noise effects. This is taken as further evidence for the existence of a diffuse molecular disk component.

The idea of a diffuse molecular component has also been suggested by Combes et al. (2012). Using the H I and CO(2-1) data from Gratier et al. (2010) for M33, the power spectrum of the zeroth moments is found to show a break, indicating the transition from the large-scale two dimensional motion in the galaxy (i.e., rotation and radial flows) to small-scale three dimensional motion (i.e., turbulent motion and expanding shells). This break occurs at ~ 110 pc for both H I and CO(2-1).

Finally, it should be noted that several studies find good agreement between H I and CO(2-1) velocities. Leroy et al. (2009) find a 1σ scatter of 6 km s^{-1}, while Schruba et al. (2011) and Mogotsi et al. (2016) find a scatter of 7.0 km s^{-1} and 3.3 ± 1.2 km s^{-1}, respectively. The latter two surveys use a larger set of the HERACLES and THINGS data. Thus the assumption that H I velocities, or an H I derived rotation model, can be used with CO(2-1) without introducing significant additional uncertainties.

Localized Converging Flows in Wind-Driven Shells — WDS are predicted to drive converging flows on ∼100–150 pc scales, which can accumulate

into molecular clouds (Dobbs et al., 2014, and references therein). This process is expected to be the dominant formation mechanism on these scales. The maximum mass of a cloud that can be formed through this process is $\sim 10^4$ $M_\odot$, based on the mean density and the velocity correlation length[7] in the local environment. The rapid accumulation of the CNM requires an additional instability requirement, typically attributed to thermal instabilities. Hartmann et al. (2001) show that these requirements are satisfied above a column density of $\sim 10^{21}$ cm^{-2}.

Examples of this process occurring have been argued for two Milky Way supershells (Dawson et al., 2011) and in the LMC (Dawson et al., 2013). Dawson et al. (2011) finds one of the three molecular clouds associated with the GSH 287+04-17 supershell is consistent with having formed from triggering within the compressed shell. This cloud is hosting on-going massive star formation, which is driving an expanding H II region and creating a blister on the shell wall (§1.3). This is a possible example of second generation triggered star formation, as predicted by the Elmegreen et al. (2002) self-regulating model. The other two molecular clouds detected are rather small ($\sim 10^2$–10^3 $M_\odot$) and are possibly remnants from the natal molecular cloud, whose internal star formation created the shell. Dawson et al. (2011) suggest that these are remnants that were located at the tips of finger-like extensions of H I based on the dense heads and sharp intensity gradients on the leading edge.

Dawson et al. (2013) investigates the influence of multiple supershells in the LMC. They find that regions within in the shells tends to have a higher molecular fraction than those outside. They estimate that 12–25% of the molecular gas in the shells was formed due to compression in the shell. However these supershells do not have a significant aggregate impact on the total molecular fraction over the entire LMC. The authors suggest that this cannot be the dominant formation mechanism of molecular clouds in the LMC.

[7]Dobbs et al. (2014) sets the maximum length to the scale height of the galactic disk.

1.5 Why M33?

M33 is an early-type spiral galaxy at a distance of 0.84 Mpc (Freedman et al., 2001). Out of the three major members of the local group, it is the smallest with a mass around 10% of M31 and the Milky Way (Putman et al., 2009). However, it is the nearest face-on galaxy to the Milky Way, with a moderate inclination of $\sim 55°$. Thus we can observe the entire disk with minimal projection effects. Its spiral structure is *flocculent*, meaning it lacks the grand-design spiral arms that dominate other spiral galaxies, like M51, M101, and arguably the Milky Way. The two optical spiral arms dominate the mid H I disk (2–4 kpc radius), while the outer H I disk (out to $\sim$ 8 kpc) may have fragments of spiral arms based on their appearance (Verley et al., 2009). Several articles have found the disk to be warped, particularly beyond $\sim$ 8 kpc (e.g., Corbelli et al., 2014). The warp was likely produced from tidal interactions with close passings with M31 (Putman et al., 2009). The average metallicity of the galaxy is ~ 0.5 $Z_\odot$, however the metallicity has a radial gradient that varies by about ~ 0.8 dex (Rosolowsky and Simon, 2008). The metallicity affects the shielding of the gas, and hence the formation of H_2.

M33 is the ideal target for this study due to its spiral arm structure, proximity and orientation. M31 is nearer, but is more inclined, adding more uncertainty from projection effects. Its ISM is also highly disturbed, based on the location of *rings* of star-formation (Lewis et al., 2015). The Magellanic Clouds are also nearer, and their H I properties have been studied in detail (e.g., Braun et al., 2009). However, they lack the large-scale spiral structure that has been shown to dominate processes in the ISM when present (Colombo et al., 2014a). Finally, observations within the Milky Way provide by far the finest spatial and spectral detail, but have several disadvantages for large scale surveys. The first is the uncertainty in distance, which is alleviated by studying an external galaxy whose distance is well constrained. The second is the projection effects

from being situated within the disk. Questions about the spiral structure of the Milky Way's spiral structure remain topics of active research. Thus M33 represents an ideal combination of proximity and viewing parameters to study the large-scale galactic structure and its effect and influence on small scales.

1.6 Outline

The remainder of this thesis is outlined as follows: §2 describes the observational setup, reduction, and imaging of the VLA observations. §3 presents an analysis of the basic properties of the new H I data cube, compares those properties to the values presented in the literature, and introduces an new masking technique for analyzing the H I filamentary structure. §4 provides a comparison to the molecular medium with observations of CO(2-1) emission line from the 30-m IRAM telescope (Gratier et al., 2010; Druard et al., 2014). The spatial and kinematic relation to the H I is examined through multiple techniques. Finally, §5 concludes the findings presented in this thesis and provides a brief outline of my continuing work.

Chapter 2

Observations and Data Reduction

Our collaboration was awarded 52 hours of observations with the VLA in New Mexico[1] to observe 13 pointings to cover the disk of M33 (Project code:14B-088, PI: E. W. Rosolowsky). These observations are a portion of a large observational campaign to use the enhanced capabilities of the VLA to re-visit observations of the Local Group galaxies (namely M31 and M33). These observations achieve two significant improvements from past observations: (1) the upgraded VLA correlators allow for simultaneous observations of continuum and spectral lines, and (2) the enhanced bandwidth makes the continuum sensitivity much higher, allowing observations of very diffuse emission to be completed. These enhancements have led to far larger data volumes (for a single project) than have been encountered before, until only recently. In this Chapter, I describe the observational setup, target selection, data reduction, and imaging of these observations and the difficulties they entailed. I also provide a brief description of the workings of a radio interferometer.

[1]The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc

2.1 Interferometry & Aperture Synthesis

The use of interferometry is driven by the need for high resolution observations. When using a single-dish radio telescope, the angular resolution is $\sim \lambda/D$, where λ is the wavelength and D is the diameter of the dish. For the H I line, resolution on arcsecond scales would require a dish diameter several kilometres in diameter, which is unfeasible to construct (the largest single-dish radio telescope is around 500 m in diameter). In contrast, an interferometer's angular resolution is set by the maximum distance between two dishes, known as the baseline. The expression for angular resolution is analogous to that for a single-dish with the longest baseline b replacing the diameter, $\sim \lambda/b$. The diameter of the dishes in the interferometer now set the angular size of the field-of-view. This alleviates the resolution issue at long wavelengths — if you want better resolution, simply move the dishes further apart. However what the interferometer gains in resolution it loses in sensitivity. The actual collecting area of the interferometer is much smaller than the total collecting area of a single dish with a diameter of b, limiting its sensitivity. Furthermore, an interferometer only samples emission on angular scales between $\lambda/b_{\rm min}$ and $\lambda/b_{\rm max}$, where $b_{\rm min}$ and $b_{\rm max}$ are the shortest and longest baselines between the dishes. There is a lower limit of $b_{\rm min} \sim D$, since the dishes cannot be physically closer to each other.

Any interferometer can be represented as a combination of $N(N-1)/2$ pairs of dishes. This means the derivation of the response for a two-element interferometer can simply be generalized, which is convenient for understanding how they work. With this in mind, I provide a basic overview of the two-element interferometer, following Fomalont and Wright (1974) and Thompson (1999). The overall schematic is shown in Figure 2.1. The interferometer works by correlating the signals from the two antennas. The voltage measured at each antennae is proportional to the electric field E generated at the feed of the

antenna, but occurs at different times due to the spatial offset b between the dishes:

$$V_1 \propto E\cos(\omega t) \tag{2.1}$$

$$V_2 \propto E\cos[\omega(t-\tau)] \tag{2.2}$$

The time-delay between arrival is $\tau = b\cos(\theta)/c$ and $\omega = 2\pi/\lambda$, where I assume the incoming radiation is monochromatic for simplicity. The interference term, $\omega\tau$, describes the difference of the path length.

The voltage patterns are passed through amplifiers before being combined within the correlator. A correlator — consisting of a multiplier, low-pass filter, and an integrator — produces an time-averaged output $\langle V_1 V_2 \rangle$. It is the combination of multiplication and averaging that creates the correlation, and removes uncorrelated noise terms arising from the amplifiers. This is opposed to the case of an *adding* interferometer, where time averaging will retain absolute noise terms, decreasing the sensitivity to signal (Romney, 1999).

The time-averaging must satisfy $\Delta t \gg 1/2\omega$ to remove the unwanted high-frequency term from the multiplier. The low-pass filter provides the final check, allowing only the low-frequency term into the output voltage. The resulting output $R(t)$ is the aforementioned interference term

$$R(t) \propto S\cos\left(\frac{2\pi b}{\lambda}\cos(\theta)\right) \tag{2.3}$$

where $S = |E^2|$ is the squared modulus of the electric field or the power. This equation shows how the interferometer's response depends on the power of the source S and the spacing of the fringes in the interference pattern; this is the basic equation of interferometry. This form of the relation represents the response to a point-source. The response for an extended source can be shown

to follow (Fomalont and Wright, 1974)

$$R(t) = \int I(\boldsymbol{\sigma}) \cos\left[2\pi \mathbf{b} \cdot (\mathbf{s} + \boldsymbol{\sigma})\right] \mathrm{d}\boldsymbol{\sigma}. \tag{2.4}$$

where $\mathbf{b}$ is the vector baseline, and $\mathbf{s}$ is the phase centre per λ (i.e., the point defined to have zero phase). This treats a brightness distribution $I(\boldsymbol{\sigma})$ as a collection of point sources, such that the power is given by $S = \int I(\boldsymbol{\sigma})\mathrm{d}\boldsymbol{\sigma}$. The coordinate $\boldsymbol{\sigma}$ is defined with respect to the phase center $\boldsymbol{s}$ (the point of complete constructive interference in the response pattern). Equation 2.4 can be simplified by using $\mathbf{b}\cdot(\mathbf{s}+\boldsymbol{\sigma}) \approx \mathbf{b}\cdot\mathbf{s}+\mathbf{B}\cdot\boldsymbol{\sigma}$ since the maximum angular size of $I(\boldsymbol{\sigma})$ is set by the response pattern of the antenna and is usually less than a degree (Fomalont and Wright, 1974, 0.277° for the VLA at 1.4 GHz). $\mathbf{B}$ is the projected spacing of $\mathbf{b}$. The cosine term in Equation 2.4 is now the sum of two terms and can be expanded using the appropriate trigonometric identity. However it is more common, and convenient, to use the complex form

$$R(t) = V \exp\left(2\pi i \mathbf{B} \cdot \mathbf{s}\right). \tag{2.5}$$

The exponential term is simply the response to a point source, as shown in Equation 2.3. The visibility function V represents the interference of the brightness distribution integrated over the whole sky

$$V = \int I(\sigma) \exp(2\pi i \mathbf{b} \cdot \boldsymbol{\sigma}) \mathrm{d}\boldsymbol{\sigma}. \tag{2.6}$$

The desired output of the interferometer is the brightness distribution (i.e. the image). Equation 2.6 shows that V is the Fourier transform of $I(\boldsymbol{\sigma})$. Adopting a Cartesian coordinates system, where $\boldsymbol{\sigma} = (x, y)$, the map produced by an interferometer is shown to be

$$I(x, y) = \int\int V(u, v) \exp\left[-2\pi i \left(ux + vy\right)\right] \mathrm{d}u\mathrm{d}v \tag{2.7}$$

An interferometer produces a map of the emission within its primary beam by taking samples in frequency space — the uv-plane — which may then be Fourier transformed to form the spatial map. Unfortunately, the sampling in the uv-plane is sparse, negating the ability to simply use a Fourier transform to recover the map. Instead, this requires the visibilities to first be gridded using a convolution function, or a set of convolution functions, that will allow a Fourier transform to be used. Creating this convolution function requires setting the resolution of the grid, and hence the resolution of the image that will be created. This is the role of the *synthesized beam*.

Why is the sampling of the uv-plane sparse? To understand, I will expand the toy model of two antennas to the general case and how this will effect the sampling. The above derivation is the most simple case for sampling in the uv-plane — one baseline between antennae yields two uv samples at a time. Since the visibilities are complex and the intensity distribution is Hermitian, the conjugate of each measurement is known, providing two samples at a time. To increase the rate of sampling, the number of antennae can be increased. For N antennas, there are $N(N-1)/2$ baselines and hence $N(N-1)$ samples taken per unit time. Additional sampling is also naturally provided by the rotation of the Earth: as the source moves across the sky, the projected baselines between the antennas change, thereby filling more of the uv-plane. An example of uv-plane sampling taken with the VLA is shown in Figure 2.3.

Even with the increased sampling rate provided by more antennas, the portion of the uv-plane that can be sampled remains limited. The maximum baseline sets the resolution limit — no information on smaller scales is provided, Similarly there is an inner hole, since the antennas physically limit how short the baselines can be. The true baseline can be no smaller than the diameter of the antennas and if the projected baselines are smaller than the diameter, a portion of one of the dishes is blocked by the other. This is known as *shadowing* and is discussed in §2.4. This inner hole sets the largest scale the interferometer

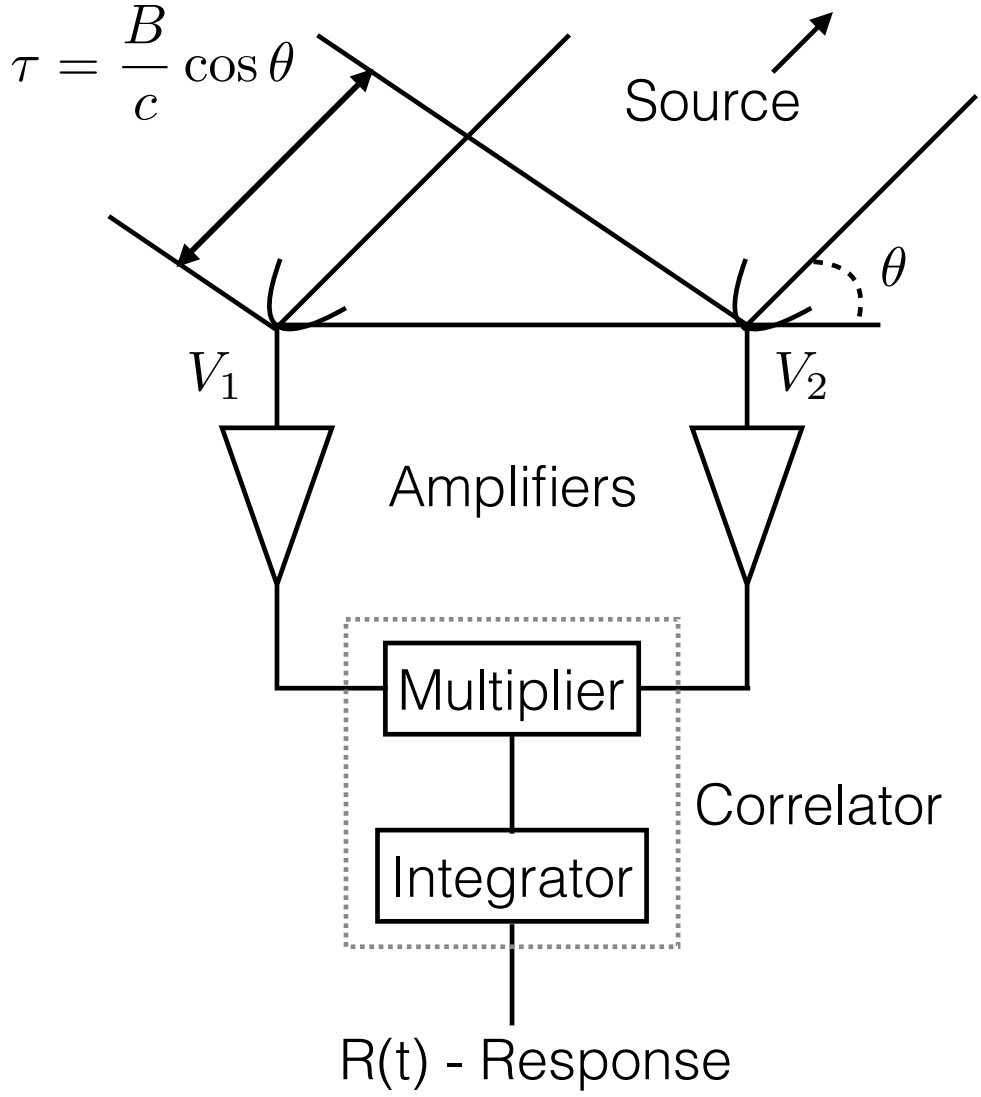


Figure 2.1: Schematic of a two-element radio interferometer. Based on Figure 10.3 from Fomalont and Wright (1974) and Figure 2.1 from Thompson (1999).

is sensitive to. The effect of limiting the available spatial scales in any image is analogous to the image analysis technique of *unsharp masking*, where an image is convolved by two (normally Gaussian) kernels, and the smaller convolved map is subtracted from the larger. This effectively only keeps structures whose size is in between the sizes of the two kernels.

These limitations coupled with the irregular sampling require the visibilities to first be placed on a grid, using a convolution method, to allow a fast-Fourier transform (FFT) to be performed. How the visibilities are weighted changes the shape of the synthesized beam, and thus the ultimate image. There are three typical choices for a weighting function: natural, uniform, or robust. Natural weighting uses the inverse noise variance in uv-cells (discrete sizes of the gridded data) that are occupied by samples, and zero elsewhere. This maximizes the sensitivity and tends to favour short baselines (since there are more samples there). This is the weighting chosen for the observations presented here. Uniform weighting assigns weights inversely proportional to the local density. This "fills" the uv-plane more thoroughly, enhancing the weights on longer baselines. The angular resolution is higher for uniform than natural weighting, but sensitivity is lost. Robust, or Briggs, weighting is a combination of the previous two that attempts to minimize the downsides of both other methods. The choice of weighting largely depends on the science goals of the observations. However this choice represents an interesting philosophical problem. Since the choice of weighting directly affects the final image, whatever image you obtain is not a unique solution for the set of visibilities measured. Because of this, some studies actually publish multiple images with different weighting choices (e.g., Andrews et al., 2009), where the increased angular resolution is enough to resolve more individual components.

With the visibilities gridded, an FFT can be performed to yield an image. However the image obtained is typically not a true representation of the sky. Instead it represents the true image convolved with the response of the antenna

— the *dirty beam*. This is referred to as the *dirty image*. The use of "dirty" here refers to imperfections in the interferometer, namely its side-lobes cause rippling in the response. An example of a dirty beam is shown in Figure 2.2. The VLA dirty beam has a distinct pattern resulting from the geometrical shape of the array. Each antenna also has a dirty beam, where the side-lobes here instead refer to the power pattern of the antenna. Side-lobes are angular extents where the antenna is sensitive to radiation away from the primary lobe. This allows stray radiation to be detected far away from the pointing centre. To recover the true image, a deconvolution algorithm is needed. These techniques aim to find a model compatible with the data in terms of an idealized beam shape (typically near-Gaussian), and reconstruct the true image in terms of the model.

The CLEAN algorithm is the canonical choice for deconvolution in radio astronomy. Other algorithm choices include maximum-entropy and compressive-sensing; deconvolution remains a topic of active research and we may see alternative methods become the preferred choice in the upcoming years. The CLEAN algorithm assumes that the dirty image is the response to a collection of point sources: each pixel in the image is an independent source. The algorithm iteratively subtracts the highest peak in the dirty image and adds the location and amplitude of the peak to a list of *clean components* that will comprise the model used to construct the clean image. This process is repeated until the maximum peak reaches a specified threshold, typically set to near the noise level in the data. The clean algorithm has numerous pitfalls, which I discuss in more detail in §2.5. The model image, containing the clean components, is convolved with an elliptical Gaussian, determined by fitting to the dirty beam. The residual map, from which the clean components were subtracted from, is then added to the convolved model image, creating the final cleaned image.

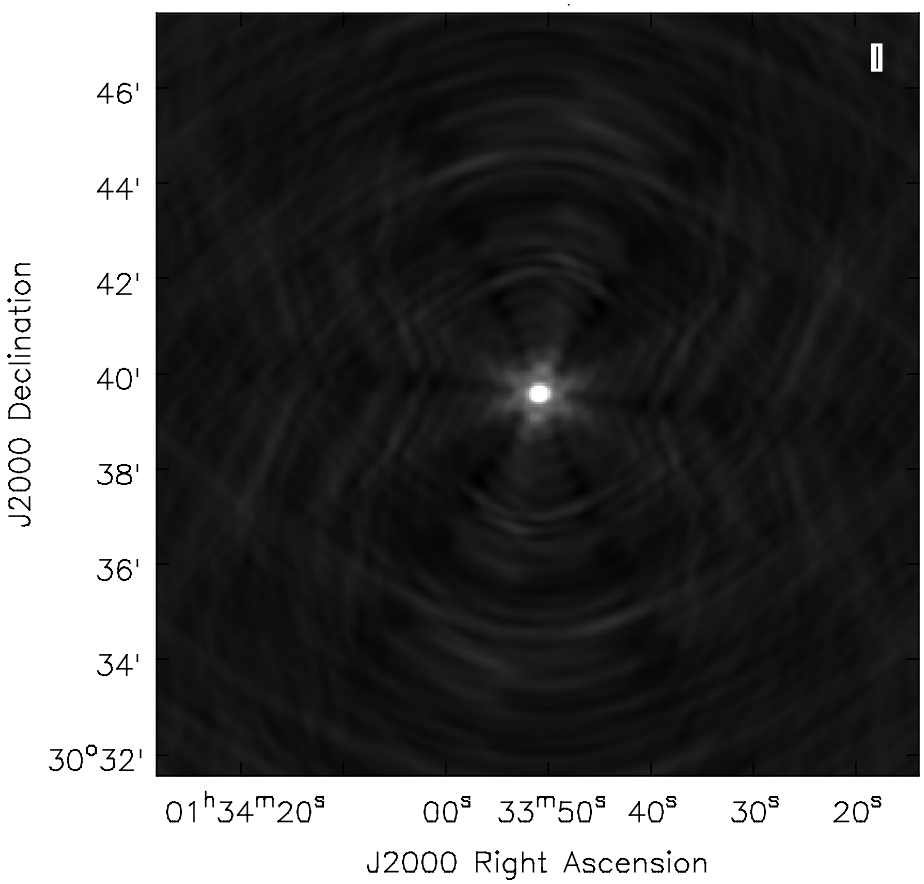


Figure 2.2: The VLA dirty beam from one of the H I channels in the observations presented here. Note the symmetry in the sidelobe response pattern; this shape is the Fourier transform of the Y-shape of the VLA antenna positions.

2.2 Project Goals & Target Selection

The overall goal of this observational project is to determine the dominant mechanisms of molecular cloud formation. Observations show that the Kennicutt-Schmidt Law is almost completely determined by the surface density of the molecular gas, while the atomic component appears to play no direct role in star formation (Leroy et al., 2013b, §1.4). This implies that the formation of molecular clouds is the bottleneck for star formation to occur (ignoring the actual sub-cloud scale processes). Conveniently, all of the proposed formation mechanisms can be traced through emission within L-band (1–2 GHz): kinematics and self-absorption are constrained using HI, and magnetic field strength and orientation are derived from the polarized continuum, which is dominated by synchrotron emission at these wavelengths.

2.3 Observations

Our observations were taken in 12 ∼4 hours tracks split equally before and after transit to maximize *uv* coverage (Table 2.1). M33 covers an angular scale ($\sim 2^\circ$) much larger than the primary beam of the VLA ($\sim 30'$ at 1.4 GHz). Thus we combine 13 pointings to create a mosaic of the entire galactic disk. To ensure uniform beam coverage over the extent of the mosaic, the 13 pointings were arranged into 3 overlapping hexagons (Figure 2.4) where the angular separation between each pointing is half of the primary beam to ensure Nyquist sampling[2]. Due to inputting one of the pointing centers incorrectly, our final observation is a single ~ 5 hour pointing on the missed centre. The time on source for this pointing centre very nearly matches the others, and thus our mosaic has uniform noise properties.

The spectral windows (SPWs) setup used in these observations includes 8

[2] `https://science.nrao.edu/facilities/vla/docs/manuals/obsguide/modes/mosaicking`

continuum SPWs and 11 spectral lines. The continuum SPWs span the entire radio L-band (1–2 GHz), each with a width of 128 MHz. The other sideband of the correlator was filled with SPWs for spectral lines (see Table 2.2): the H I line, 4 hydroxyl (OH) lines, and 6 hydrogen radio recombination lines (RRLs). Since the primary science case is driven by the H I observations, this sets the sensitivity requirements. The H I SPW was set to have channel widths of 0.2 km/s with a bandwidth covering 1690 km/s. This gives an ideal sensitivity of 3.72 K within each channel, assuming an expected 18″ beam. The expected sensitivity of the continuum observations is about 50 μJy, assuming that 40% of the 1 GHz bandwidth is lost to radio-frequency inteference (RFI). Both the H I and continuum observational setup provides the most sensitive observations of M33 in this band to date. The spectral resolution of the H I line is ∼6 times better than the archival AT0206 observations, whose spectral resolution is $\sim 1.2\ \mathrm{km\,s^{-1}}$.

Each observation was setup in a standard configuration, beginning with observations of standard calibrators. The calibration process is more thoroughly described in §2.4. One of the VLA primary calibrators 3C48 is conveniently located close to the position of M33. 3C48 is a quasar that is unresolved in the C-configuration of the VLA (synthesized resolution of $\sim 15 - 20''$). At 1.4 GHz, 3C48 has a flux of ∼16 Jy, making it one of the brightest L-band radio sources in the sky and ideal for minimizing overhead and calibration time. We use 3C48 as our flux calibrator, amplitude calibrator, and phase calibrator (see §2.4). The calibrator is observed at the beginning of the track before the science targets are observed. The 13 pointings in our mosaic are then observed in a loop, where we include short scans of the calibrator after each iteration. The additional calibrator scans ensure proper phase and amplitude solutions while calibrating. Each scan of the pointings in the mosaic lasts about 3.5 min, so one loop of all of the pointings takes about 45 min, about the maximum recommended time between calibrator scans. Five of these loops are executed

Table 2.1: All observations of M33 used in my thesis, including the archival data from the AT0206 H I project (PI: Thilker & Braun). Columns: Project code; VLA Configuration; Date observations; UTC start time; UTC end time; total duration of the observation.

Project	Conf.	Date (yy-mm-dd)	Start	End	Duration
AT0206	C	97-10-19	4:51:40	10:47:40	5:55
AT0206	B	98-07-14	9:16:39	17:11:40	7:55
AT0206	B	98-07-16	9:06:00	17:03:20	7:57
AT0206	B	98-07-20	8:50:20	16:47:59	7:57
AT0206	B	98-07-21	8:48:00	16:43:59	7:56
AT0206	B	98-08-15	7:39:19	15:35:40	7:56
AT0206	B	98-08-16	7:38:00	15:31:39	7:54
14B-088	C	14-10-22	2:08:43	6:24:56	4:16
14B-088	C	14-11-11	0:30:08	4:46:21	4:16
14B-088	C	14-11-14	5:38:28	9:54:56	4:16
14B-088	C	14-11-15	0:34:30	4:50:38	4:16
14B-088	C	14-11-15	5:50:03	10:06:28	4:16
14B-088	C	14-11-16	0:35:13	4:51:26	4:16
14B-088	C	14-11-16	5:21:41	9:38:09	4:16
14B-088	C	14-11-17	0:29:01	4:45:12	4:16
14B-088	C	14-11-17	5:43:10	9:59:32	4:16
14B-088	C	14-11-22	0:04:50	4:23:03	4:16
14B-088	C	14-11-25	4:54:19	9:10:45	4:16
14B-088	C	14-12-11	22:08:18	3:00:18	4:52

in each observation. The observation ends with a scan of 3C138, a standard polarization calibrator since our continuum observations include all polarization components. Our final observation of the missed pointing centre has the same overall setup, but only the missed pointing centre is observed rather than looping through the entire set.

2.4 Calibration

The raw data from the interferometer is not significantly more processed than the voltages measured by each antennae (i.e., Equation 2.1). The steps outlined in §2.1 need to be solved for to obtain the set of visibilities with which a usable

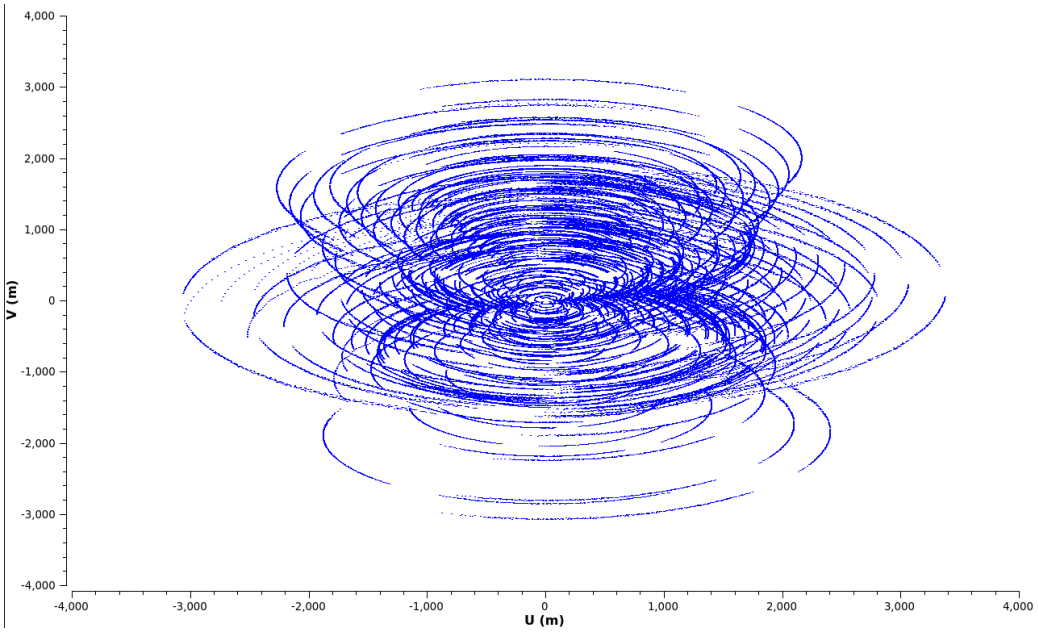


Figure 2.3: Tracks in the *uv*-plane for H I in the 14B-088 observations.

Table 2.2: The SPW setup for spectral lines used in the observations.

Line	Rest Freq. (GHz)	Bandwidth (km/s)	Channel Width (km/s)	Channels
HI	1.42	1690	0.206	8192
OH(1612)	1.612	743	1.45	512
OH(1665)	1.665	720	1.41	512
OH(1667)	1.667	719	1.40	512
OH(1720)	1.720	697	1.36	512
H(172)α	1.28	646	1.26	512
H(152)α	1.85	659	1.29	512
H(153)α	1.82	726	1.42	512
H(166)α	1.42	811	1.58	512
H(158)α	1.65	841	1.64	512
H(164)α	1.48	936	1.83	512

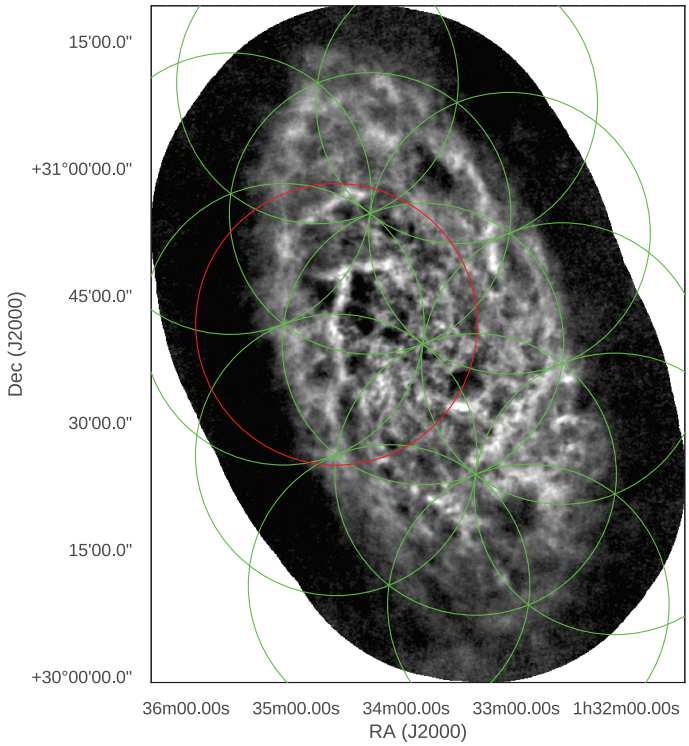


Figure 2.4: The pointing centres used for the new observations overlaid on the zeroth moment map of the new H I data. The radii of the circles is 0.277°, corresponding to the FWHM primary beam size of the VLA antennas at 1.4 GHz. The incorrect pointing centre is halfway down on the left side shown in red. The final observation was used to complete the mosaic at the pointing directly south of the incorrect pointing.

image can be made. However, there are many "real-world" problems that make this process difficult. Calibration is required to remove or minimize a number of effects: telescope imperfections, limitations of the electronics used in the amplifier and correlator, hardware failures, scheduling/observational setup issues, non-ideal atmospheric conditions, and radio frequency interference (RFI), to name a few. The outline I present here draws from two excellent resources that I credit in vastly helping my understanding: Fomalont and Perley (1999) and lecture notes by George Moellenbrock given at the 2015 ALMA Workshop in Penticton, B.C.

The VLA has a data-reduction pipeline that is run on all observations before being delivered to the scientist[3]. Though certainly not perfect and robust in every situation, the pipeline automates many repetitive calibration steps and is typically a good "first step" in preparing the data.

The pipeline follows a standard reduction path used for calibrating interferometric data. The following steps are an outline of this process. In some steps I have noted when the operation is inappropriate for spectral-line observations.

1. **Hanning Smoothing** — Low frequency radio bands (< 2 GHz) are marred by RFI. These can either appear as strong, narrow features in frequency, or "blips" at certain times during an observation. Typically RFI is significantly stronger than astronomical sources and this leads to spectral ringing, also known as the Gibbs phenomenon, within the data. Hanning smoothing can minimize theses effects by applying a triangular smoothing element, resulting in a loss of spectral resolution by a factor of 2. This is inappropriate to use for narrow spectral lines. However, the correlator sampling returns a sinc function for the spectral resolution element. The ripples from the sinc function may also be minimized through Hanning smoothing, as long as the reduction in spectral resolution leaves

[3] `https://science.nrao.edu/facilities/vla/data-processing/pipeline`

the data usable.

2. **Deterministic & Online Flags** — Online flags include hardware or antennae issues typically recorded by the telescope operator. The term "flagging" is used here to mean portions of the data that will be ignored during all further steps. Deterministic flagging removes typical correlator issues, notably channels near the end of the sub-bands. "Shadowed" data are also flagged at this stage. This refers to a period of time during the observation where the projected baseline between two antennae was less than the diameter between one of the dishes. These data are uncertain due to the blocking of one of the antennae and the diffraction pattern that this can cause (Fomalont and Perley, 1999).

3. **Flux Calibration** — The raw data do not have an absolute flux scale a priori. The conversion into a physical unit is achieved by using a model determined for the flux calibrator observed during the observation. There are only 4–5 primary calibrators that the VLA provides flux models for (the number available depends on the frequency of the observations); in our observations, we use 3C48. The model for the flux calibrator is applied to the scan of the object. This sets the flux scaling used for the rest of the observation.

4. **Prior Calibrations** — These include correcting for antennae positions, pointing errors, and atmospheric effects. The antennae corrections are vital to ensure the correct baselines are used; incorrect baselines will lead to solutions without phase coherence. Atmospheric effects, namely the opacity, is provided by the VLA for typical atmospheric conditions at the telescope site.

5. **Initial Delay and Baseband Calibration** — The delay and baseband calibrations remove effects from the correlator. Delay corrections are

needed to account for deviations in the correlator model. Correlating the data between the antennae requires injecting a time delay between the signals, due to the physical offset between the antennae (Equation 2.1). If the delay time is incorrect, the calculated phase difference will be incorrect leading to incoherence in the calibrated data. Baseband corrections reflect the variations in the instrumental response across a sub-band. These variations are typically larger than the actual signal in the data.

6. **Flagging Bad SPWs** — The pipeline determines if any of the calibration solutions have failed for a particular SPW. This can reflect a system issue (bad deformatter) or severe RFI effects (as expected in L-band). In these cases, the entire SPW is flagged and assumed to be uncalibratable through an automated process.

7. **Calibrator RFI Removal** — The RFLAG algorithm uses a sliding standard deviation window to detect and flag narrow spikes in time and frequency. These features, at least for a calibrator, are almost guaranteed to be RFI. The algorithm is used on the calibrator used for the delay and baseband calibration.

8. **Second Delay and Baseband Calibration** — With RFI hopefully minimized in the calibrator data, Step 5 is repeated to refine the previous solutions.

9. **Solving for Gain Solutions** — The gain refers to a complex quantity tracking the amplitude and phase changes during an observation. To retain coherence across the observation, variations in the gain must be accounted for. This drives the need for the repeated calibrator scans described in §2.3. The gain at each scan is used to create a smooth, interpolated model of the gain variations across the observation. The amplitude

typically varies slowly, but the phase may change rather quickly. Thus it is necessary to repeat scans of a calibrator to provide a baseline for changes in the phase. These phase changes depend on the length of the baselines: the longer the baseline, the more rapid the potential phase variations.

10. **Bootstrap Flux Density to the Gain Calibrator** — The gain calibrator is not typically the same source as the flux calibrator (though it is for our observations). This means that the flux of the gain calibrator is unknown. The model for the flux density calibrator can be "bootstrapped" to the gain calibrator to set the absolute flux density.

11. **Re-derive Calibration Solutions** — With the flux of the gain calibrator set, the delay, bandpass, and gain calibrations are re-calculated such that they now have an associated absolute flux.

12. **Apply Calibrations** — The calibration solutions from the previous step are applied to the science data, and any additional calibrators required. Though I have not included the polarization calibration in this overview, this would be the appropriate stage to derive the polarization dependent calibration solutions.

13. **RFI Removal** — The RFLAG algorithm is re-run on all of the data this time. If the spectral lines in the observation are expected to be narrow, RFLAG may remove the actual signal! Thus, this stage should typically be disabled for spectral lines.

14. **SPW Statistical Weights** — These reflect the RMS noise level within a SPW and are applied as Gaussian weights to the data (i.e., $\propto 1/\sigma^2$). The RMS standard deviation is given by

$$\sigma_{ij} = \sqrt{\frac{T^i_{\rm sys} T^j_{\rm sys}}{2\Delta\nu\Delta t}} \tag{2.8}$$

where i, j denote two antennae, $T^i_{\rm sys}$ is the system temperature of the $i^{\rm th}$ antenna, $\Delta\nu$ is the spectral bandwidth, and Δt is the time-on-source.

15. **Test Image the Calibrators** - The quality of the calibration can be partially examined by imaging the calibrator scans. If the data are well calibrated, the resulting image should be a clean point source, assuming the calibrators are point sources.

At the time of these observations, the VLA was using the 1.3.1 pipeline version, built on the CASA software package[4]. This version was primarily intended for "Standard Stokes I" SPWs, i.e., a 128 MHz SPW without polarization components. Modifying some aspects of the pipeline can make it more suitable for spectral line data, notably removing the Hanning smoothing step and the automatic radio-frequency interference (RFI) removal. Due to the mix of the continuum and line SPWs in our observational setup — a so-called "mixed" correlator setup — I found it necessary to create a modified version of the pipeline capable of automatically splitting the continuum and line SPWs. After this splitting step, the components are each run on a separate pipeline: the standard pipeline for the continuum portion and the modified pipeline for the line data.

The key to obtaining a good calibration solution for spectral lines with the pipeline is to use a strong calibrator. Since 3C48 is used as the calibrator, this is not an issue for these observations. In most cases, the pipeline returns well-calibrated data, but fails to remove all of the pathological portions. This behaviour is somewhat desirable to avoid flagging out good data and impacting the quality of the calibrated data. I found that by iteratively using the pipeline, segmented with by-hand flagging of the data in-between runs, produced an excellent final product. For the spectral lines, the extent of the by-hand flagging was removing bad integration periods (evident as outliers in their amplitude or

[4]CASA version 4.2.2

phase). Though time-consuming, the process was straightforward.

By-hand flagging of the continuum portion was significantly more substantial. This is due to the severe RFI across L-band[5]. This immediately makes 40% of the continuum data unusable. The pipeline recognizes regions of severe RFI and can completely flag SPWs altogether. Two of the eight SPWs were always flagged after the initial pipeline run. The by-hand editing was required to remove clusters of RFI spikes, while attempting to salvage the relatively unaffected channels they surrounded. Upon final calibration of the data, $\sim 55\%$ of the channels had been removed minorly impacting our expected sensitivity levels which were based on an assumed removal fraction of 40%.

The process of calibrating this data was slow-going and particularly hampered by the sheer volume of data. Once calibrated, each observation has a size around 500 Gb. Working with data of this size requires computational power beyond what most desktops are capable of. I performed the majority of the data reduction on the NRAO cluster, which is dedicated for handling of large data from the JVLA and Atacam Large Millimetre Array (ALMA). Since portions of the reduction process require user-interaction, local clusters available through Westgrid can not be effectively used. However, the large amount of computational power from cloud services (such as the University of Alberta's Cirrus Cloud or Amazon Web Services) are an ideal alternative for local data processing. Use of cloud resources for data processing is further explained in §2.5.

Archival Observations — In addition to processing these new observations, I also processed archival H I observations taken by the VLA before its upgrade. These observations are also listed in Table 2.1. The purpose of this exercise was two-fold: to introduce me to the calibration and imaging process of radio data

[5]The RFI spectrum at the VLA site is shown here: `https://science.nrao.edu/facilities/vla/observing/RFI/crosscorr/lband/index`

with a significantly smaller dataset, and to re-image the data using modern imaging techniques that were not available when the observations were taken. Since these data include observations with the VLA in a more extended configuration (the B configuration has longer baselines than C configuration), the archival data has roughly three times the angular resolution that the 14B-088 observations have (see §2.5).

The reduction of these data proceeded as described above. I utilized, and contributed to, a custom CASA pipeline for archival H I VLA observations developed by Adam Leroy[6].

2.5 Imaging

This section extends the discussion of imaging and deconvolution in §2.1. I also introduce a technique for speeding-up the imaging of massive spectral-line datasets. I have broken this section into the individual steps required to reproduce the imaging results; there is no particular order since each component must be used simultaneously.

Multi-Scale Clean — The basic implementation of the clean algorithm is scale-free, since it assumes complete independence between pixels (Cornwell et al., 1999). However, this representation is unrealistic for extended sources and clean is often unable to completely account for extended emission — non-random features will remain in the residual map down to the noise level. This can be alleviated, to an extent, by allowing clean to use a range of clean component sizes. This variant of the algorithm is called multi-scale clean (Rich et al., 2008). I used this technique with scales of 1, 2, 5, 10, and 20 times the beam size for imaging both the new and archival M33 observations. The typical single pixel delta function scale was also used. By visual inspection, the addition

[6] `https://github.com/akleroy/hi_reduction_scripts`

of multi-scale clean components reduces the amount of non-random residuals at the noise limit. This is key for recovering the maximum amount of diffuse emission, which will dominate the structure of H I in M33. The drawback to its approach is the computational cost. Additional convolution models must be created at each scale, and in my experience with using CASA, this can take enormous amounts of time compared to the actual clean process.

Clean Models from Single-Dish Data — In §2.1, the effect of missing short-spaces in the *uv*-plane was discussed. These artifacts can be removed by adding a source of short-spacing data, observations with a single-dish telescope. Observations from a single-dish telescope can be thought of as a near-uniform sampling in the *uv*-plane up to the dish diameter, where minor deviations result from the primary power pattern of the antenna. Thus, the Fourier transform of the low-resolution single-dish data can be added into the *uv*-plane with the interferometer data. The inverse Fourier transform of the added data will yield a map with the short-spacing portions, and the high-resolution provided by the interferometer, provided there is overlap in the *uv*-plane between the dish diameter and the shortest interferometer baseline. This method is known as "feathering." Alternatively, the single-dish data can be used as the initial clean model for deconvolution. This is the approach used for imaging in this thesis as it helps to stabilize the clean algorithm (see below). The provided single-dish data is converted into a clean model image, which stabilizes and improves the clean process since it does not have to begin adding components from zero.

The single-dish data I use is presented in Putman et al. (2009). These observations were taken with the Arecibo telescope as part of the GALFA survey (Peek et al., 2011). The Arecibo beam size is 3.8'×3.3' and the published data cube has a spectral resolution of 736 m/s. Since the shortest baseline of the VLA in C-configuration is 0.035 km[7], the largest angular scale covered in the

[7] `https://science.nrao.edu/facilities/vla/docs/manuals/oss/performance/`

VLA data is ~ 20' at 1.4 GHz. Thus, there is enough spatial overlap between the datasets for them to be merged correctly, according to the recommendations found by Kurono et al. (2009). CASA is unable to automatically perform the preparation steps to use the Arecibo data as a model. Two pre-processing steps are needed. First, the Arecibo data must be up-sampled to match the grid that the VLA data will be imaged at. To find the appropriate grid, I first imaged a single channel of the VLA data without deconvolution to ensure that: (1) the grid does not cut-off any regions with signal, and (2) there are no obvious aliasing effects from not providing enough pixels within the beam size[8]. To regrid the data I used Adam Ginsburg's FITS_tools package[9].

Regridding is typically a straightforward process. However the VLA grids of my data require a significant amount of memory and it is unfeasible to run such a process in a single step. The size of the VLA grids I produced were $1200\times2560\times2560$ pixels for the new observations and $230\times4096\times4096$ pixels for the archival data, where the first dimension is the number of spectral channels and the last two are the spatial dimensions. Since the data is being up-sampled in each dimension, the regridding can be done in portions. First, I regridded the Arecibo in the spectral dimension to match the VLA channels. The second regrid on the spatial dimensions can then be performed in two dimensions. The memory issue can then be alleviated by streaming each spatially regridded channel into a pre-made empty FITS file. This streaming approach is very useful in general for writing FITS files too large to fit into memory. The final regridded cube, converted into units of Janskys per beam, can now be used with CASA's CLEAN task.

Figure 2.5 demonstrates the effect of using single-dish data as the starting model for clean. When no model is given, the VLA-only spectrum shows signs

`resolution`

[8]It is recommended to choose the grid-spacing 3-4 times smaller than the beam size to provide convolution support for the beam element.

[9]`https://github.com/keflavich/FITS_tools`

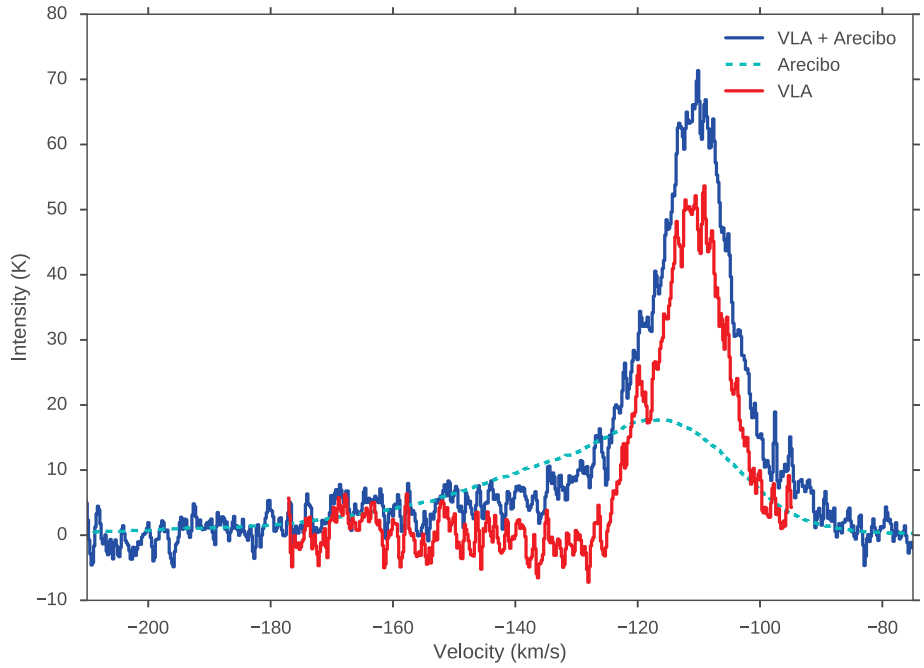


Figure 2.5: A spectrum from the new data is shown with and without the use of the Arecibo data as a starting model. When no model is given, the resulting spectrum shows negative bowls from a lack of small-spacing information, and the extended tail in the Arecibo data is completely missed. The difference in the peak brightness is due to the different beam sizes between the VLA and Arecibo data.

of negative bowls, which occur due to the lack of small-spacings in the data. The extended, low brightness tail in the Arecibo data is not detected in the VLA data alone, and may partially be the result of beam smearing.

Clean Masking — The clean algorithm is well known to be unstable. This is at least in-part due to the clean algorithm not having a deterministic solution (Cornwell et al., 1999). While the use of single-dish clean models can improve stability, a further substantial improvement can be made by limiting where the clean algorithm can look for clean components to subtract. These are known as clean masks, windows or boxes. This keeps clean from wasting computational time on regions where there is no signal or the telescope coverage is low (e.g.,

away from the pointing centre). On the other hand, one must be sure that no real signal falls outside of the clean mask, as it will not be included in the model image. The use of clean masks is encouraged in CASA through "interactive cleaning", where the user chooses regions to mask after some number of clean iterations. This approach yields good results and increases the stability of clean. However it is unfeasible for large data sets, particularly for spectral-lines where there may be thousands of spectral channels, each requiring a slightly different mask. I again used information from the Arecibo H I observations to create clean masks for each channel. The data in the up-sampled data cube were cut above 3σ to create an initial mask. A binary closing operation[10] was then performed with a structuring element of the Arecibo beam size, eliminating spurious regions much smaller than the beam size. This ensured that the clean mask would fully encompass the regions with signal. I also applied an additional masking step, limiting the mask to where the beam coverage in the mosaic is above 0.5. This keeps the clean algorithm from cleaning the strongly noise-dominated edges of the map. If this last cut was not applied, I found that clean rapidly became unstable. This mask was also applied to the initial Arecibo clean model to ensure the model image would have no components outside of the clean mask[11]. Just like the single-dish model, the clean mask must match the grid shape to be used in the CASA clean task. I created the mask from the original Arecibo data, then regridded it onto the VLA grid following the same procedure used for the model.

Rapid Spectral-Line Imaging: Divide and Conquer — The volume of the data is the limiting factor throughout the reduction and imaging process. After splitting out the relevant velocity range in the H I data, the combined visibility data across all 12 observations is about 150 GB, even after I have actu-

[10]This is defined as a dilation of a binary mask, followed by an erosion with the same structuring element (Shih, 2009).

[11]This was likely a bug in CASA 4.4 and I believe it has been fixed in newer releases.

ally *removed* all flagged data in the assumption that this is the final calibrated data. The data footprint is significantly reduced (from $\sim$ 500 GB for HI), but it still too large to operate on in memory. CASA is designed to handle these situations by reading the data in chunks. This leads to massive amounts of I/O operations, where the limiting factor in the speed becomes how quickly data can be read/written to the storage medium. For large operations, like imaging 1200 spectral channels at once, this is extremely time-consuming. A second factor in the time required to image large amounts of data are the convolution models. Excluding the multi-scale addition, each channel in the data requires a different convolution model, since the resolution is frequency dependent.

Given these drawbacks, my approach to imaging and cleaning large amounts of data is to first split them by channel. Each channel in the data is imaged separately, and combined back into a cube in post-processing. This alleviates much of the issue with the two problems above: I/O operations are reduced because there is significantly less data input at once, and only the multi-scale convolution models need to be created before the cleaning process proceeds. This idea is not particularly novel, and I doubt it is new (though I have not found a publication mentioning such an approach), but it has proven extremely useful with these new observations. When working in a cluster environment, each channel image can be submitted as its own job, in theory allowing each channel to be simultaneously imaged. To use the clean mask and model created from the Arecibo data, they must also be split into individual channels.

As an example of the speed-up gained, I found that imaging a single channel using the full data set as the input was **8 times slower** than when it was split out. I do not know how much longer imaging the entire cube at once is versus the approach explained here as I have not been able to have the latter process complete. When trying to find this out with the archival H I data, the process had been running for about 5 weeks before the cluster was taken offline for maintenance.

There are two issues with using this per-channel approach. First, the data must be in the correct frame and spectral resolution expected for the final image. The clean command in CASA is capable of regridding visibility data to specified parameters internally. Second, the time to split off each channel from the entire data is non-negligible. Using a lustre filesystem on the Jasper cluster, I found that CASA took 15 minutes to split off one channel[12]. For the 1200 channels in the data, this means it took 2 weeks just to create the individual channel datasets! Afterwards, the imaging and deconvolution would complete in about an hour for each channel. While this approach is significantly quicker, it still requires a significant commitment of time and computational resources.

Finally, I think it is important to note that the CASA development team is working on incorporating MPI processing into CASA. A fully-functioning MPI-based clean process would likely render the approach described here fairly obsolete, provided one has enough resources to run on many cores at once. The MPI approach will split portions of the data by channel (in the case of spectral-lines), as is recommended here, but will do so in a far more automated way. However for astronomers' who do not have access to large clusters, the divide-and-conquer approach described here can be conducted for relatively low-cost using cloud resources.

Developing well-tested alternatives for large-data volumes will be essential in the coming years for radio astronomers. The (upgraded) VLA and ALMA have begun to usher in the age of commonly large-data in radio astronomy, and future projects, such as the Square Kilometre Array (SKA), will place radio astronomy firmly in the realm of "big data".

The final H I data cube was imaged using natural weighting, which gives a beam size of $19'' \times 17''$. The pixel dimensions are 1178 channels of 2560 by 2650 pixels, respectively. Assuming a distance to M33 of 840 kpc (Gil de Paz

[12]Using the split command. The mstransform task was available in CASA 4.4, but I found that it was 25% slower.

et al., 2007), this projects to a physical size of $\sim 76 \times 68$ pc.

Comparing to Arecibo — It is necessary to compare the flux contained in the combined VLA and Arecibo cube to that of the original Arecibo cube. The single-dish data should contain *all* of the integrated flux since it covers all of the short-spacing information. If the combined VLA and Arecibo data is convolved back to the resolution of the Arecibo data, it should match the Arecibo data and have the same integrated flux, ideally. I performed this test for each channel in the respective datasets and found that the combined cube contained 1.45 ± 0.04 times the flux in the Arecibo cube, based on averaging the ratio across the channels. The offset in the total spectra of these datasets is shown in Figure 2.6, with the corrected version shown in Figure 2.7. This appears to be a bug within CASA. The combined cube can only contain more flux than the single-dish data if, through dilution of the large beam, some of the emission is smoothed outside the bounds of the footprint of the interferometric data. However this should not be the case for my data, and if it was, it seems highly unlikely this could account for an increase of 50% in the total emission. Furthermore, I call this a bug since the flux scaling applied by CASA in the clean task is not set by the user. The assumption is made that the single-dish data has a better absolute flux calibration, and the interferometric data should be scaled based on that. This near constant offset across all of the channels suggests that the scaling applied internally by CASA is incorrect by the same factor.

I found a similar result when checking the Arecibo data against the combined archival VLA and Arecibo cube, but the scaling factor was closer to 1.2. Since the synthesized beams are different between the new and archival observations, this suggests the bug may lie in the conversion from Jy per beam to Jy per pixel, which is the appropriate units needed to rescale the flux based on the single-dish data.

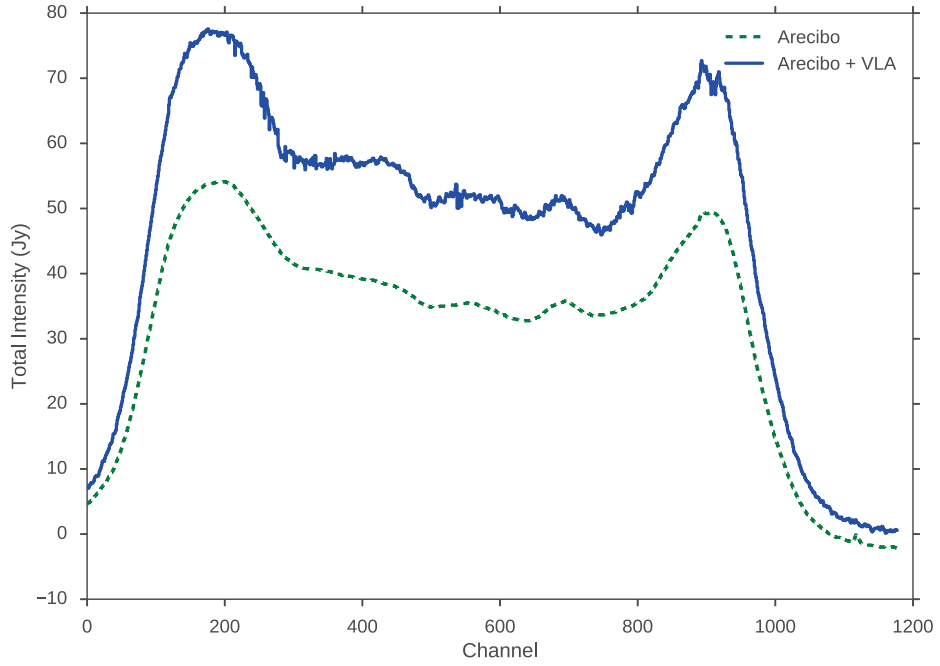


Figure 2.6: The total spectra, summed spatially within each channel, are shown for the Arecibo and combined VLA and Arecibo cubes.

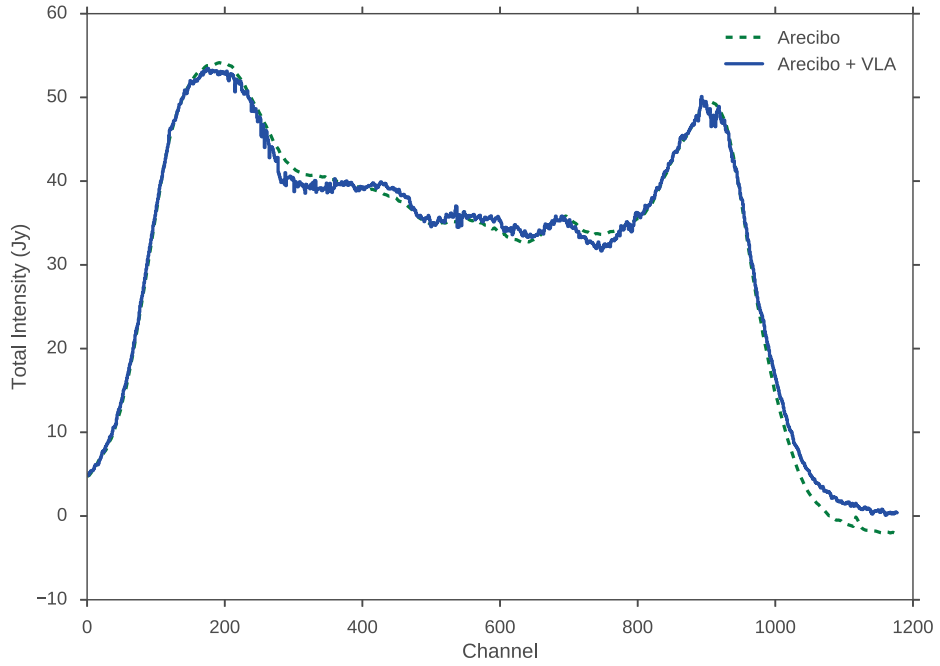


Figure 2.7: The total corrected spectra with the 1.45 offset factor applied.

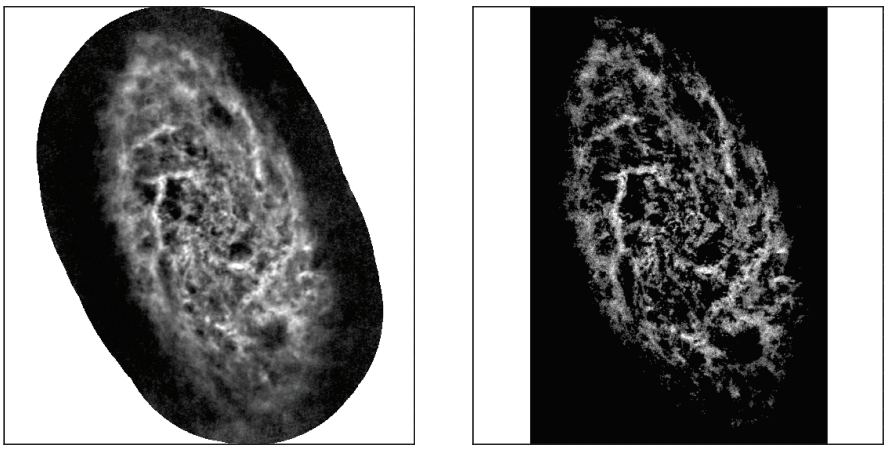


Figure 2.8: The zeroth moment maps are shown for the archival and new observations. Due to the higher noise levels in the archival data, the data has been masked more heavily causing multiple holes to appear in the data. These points correspond to line-of-sights where all of the data has been masked.

Archival versus New Observations — To highlight the differences between the improvements in the data quality, I have plotted the zeroth moment images (sum over spectral channels) in Figure 2.8. The differences in the zeroth moment maps are due to a higher noise level in the archival data. The archival observations have less integration time in a more extended VLA configuration. For these reasons, the archival observations are unable to pick up as much of the diffuse component as the new observations. Thus, there are fewer "holes" of missing data in the new zeroth moment map.

A representative spectrum from both of the data cubes is shown in Figure 2.9 towards a fairly bright H I source, the giant HII region NGC 604 in the Northern spiral arm. This highlights the real advantage of these new observations: the significant improvement in spectral resolution reveals complex velocity structures that are not apparent in the archival data. Further spectra are presented and discussed in Chapter 3.

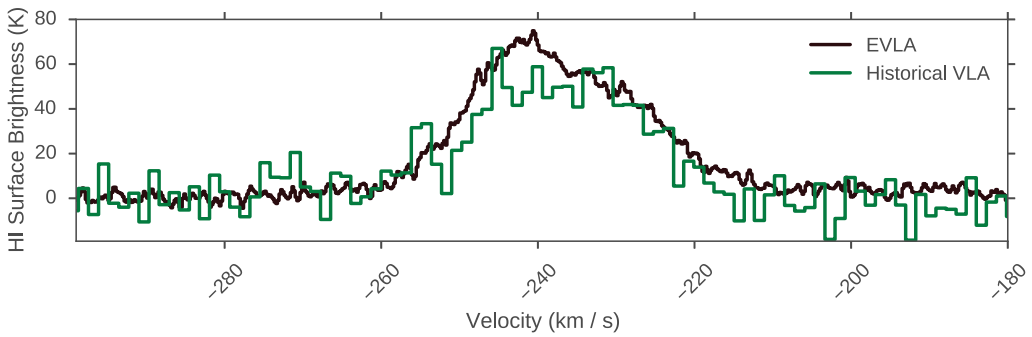


Figure 2.9: The same H I spectrum is shown in the archival and new data cube. This line-of-sight is near the giant HII region NGC 604.

2.6 Data Overview

Of the observations described in the previous sections, the remainder of this thesis focuses only on the H I line data. The continuum observations suffer from dynamic range effects, where bright background objects within the field causing severe imaging artifacts. This makes imaging the diffuse continuum component difficult. Recent improvements in wide-field and wide-band imaging may provide significant improvement (Rau et al., 2016). Furthermore, I have avoided discussing the polarization components of the continuum observations throughout this chapter. The VLA pipeline (§2.4) does not yet support polarization calibration because, particularly for diffuse emission, it is highly non-trivial. The polarization represents a fraction to the total continuum emission (the Stokes I component), and thus the imaging difficulties will need to be solved before attempting work on the polarization component. This work will be completed as part of my PhD thesis. Given the sensitivity limit of ~ 2 K (assuming an $\sim 18''$ beam at 1.5 GHz in a 1.2 km s^{-1} channel), it is unlikely that our data of the OH and hydrogen RRLs spectral lines contain any significant signal. Preliminary imaging of these lines around the giant HII region NGC 604 does not show any discernible emission. As the most active star-forming region in M33, the abundance of ionized gas outflowing from NGC 604 makes it the most likely region to detect signal in these spectral lines. Future work on

this portion of the observations includes line averaging (between the OH and hydrogen RRLs, respectively) to further lower the noise level and combining with additional JVLA observations that will be taken in the fall of this year.

Chapter 3

The Atomic ISM in M33

The properties of H I emission can be used to find several key properties of a galaxy and the nature of its ISM. The analysis within this chapter uses the typical approximations made for extragalactic H I observations, namely approximating each spectrum across the disk as a Gaussian. I show that while these traditional approaches continue to be useful, the high sensitivity and spectral resolution provide additional information that deserves a more complex and thorough spectral modeling approach.

3.1 Previous Work

As the second nearest spiral galaxy to the Milky Way, M33 has been a target of considerable interest for studying ISM properties since the inception of radio interferometers. Much of the early H I interferometry work on M33 is presented in Wright (1974), where he discusses several properties of the map taken with the Cambridge Half-Mile Telescope (Wright et al., 1972). Their data has a synthesized beam of $1.5' \times 3'$ and a spectral resolution of 39 km s^{-1}. Comparing to the observations presented in this thesis, it highlights the extent of the improvements of radio interferometers in the past 40–50 years. Wright (1974) finds an H I radial profile with an approximately constant hydrogen column

density out to a radius of about 8.3 kpc[1], where it then drops off rapidly. They also publish the first M33 rotation curve using HI. Using a Brandt rotation curve (Brandt, 1960), they find a maximum rotational velocity of 100 $\mathrm{km\,s^{-1}}$ and an index of 1.0. Assuming Keplerian rotation and a homogeneous disk, this implies a disk mass of 1.7×10^{10} $\mathrm{M_\odot}$. Wright (1974) finds regions with significant residual velocities after removing circular motion using the rotation curve. He uses these deviating regions to constrain streaming motions within the spiral arms (which are not resolved) to 5 $\mathrm{km\,s^{-1}}$. Finally, he finds that the H I velocities scale linearly with the velocities implied from optical studies of HII regions in M33, with a scatter of about 20 $\mathrm{km\,s^{-1}}$.

Deul and van der Hulst (1987) presents H I observations of M33 taken with the Westerbork Synthesis Radio Telescope. Their map has a spatial resolution of $12'' \times 24''$ and a spectral resolution of 8.3 $\mathrm{km\,s^{-1}}$. Importantly, they include zero-spacing information by combining their map with observations taken with the Effelsburg 100-m telescope. This is the first H I map of M33 that was capable of resolving the intricate H I structures throughout the disk. In Deul and den Hartog (1990), the authors present the first catalog of bubbles in M33.

The next substantially improved observations of H I in M33 were conducted by the VLA in 1997 and 1998 as part of the AT0206 project (PI: Thilker; these are the archival observations reduced in tandem with the our new observations in §2.3). Initial reductions were presented by Thilker et al. (2002), but the completed reduction and imaging was not presented in a publication until Gratier et al. (2010). Further improvements of these observations are presented by Braun (2012) and Corbelli et al. (2014), who added zero-spacing data using observations with the Green Bank Telescope (GBT). The brief analysis presented by Thilker et al. (2002) emphasizes the the presence of narrow line features towards bright regions, resulting from the superposition of unresolved

[1]A distance of 630 kpc was used originally used, giving a radius of 6.5 kpc. I have adjusted this value to the distance of 840 kpc used here.

CNM H I clouds, and a broader line features in more diffuse regions, as expected from the WNM. Using a region out to a galactic radius of 8.5 kpc, Gratier et al. (2010) find a total H I mass of 1.4×10^9 $M_\odot$, largely agreeing with the mass found by Putman et al. (2009), 1.5×10^9 $M_\odot$, from their Arecibo observations. Given the lack of zero-spacing data in the map used by Gratier et al. (2010), these value are remarkably close. Corbelli et al. (2014) also find a total H I mass of 1.5×10^9 $M_\odot$.

3.2 Rotation Curve

The smooth ubiquity of H I emission across galactic disks provides an excellent tracer for the rotation curve. I perform a simultaneous fit to the first moment map of my data using the DISKFIT routine (Spekkens and Sellwood, 2007; Sellwood and Sánchez, 2010). The first moment map is created by calculating the centroid for each spectrum in the cube

$$\bar{v} = \frac{\sum_i v_i T_i}{\sum_i T_i} \tag{3.1}$$

where i is the channel, v_i is the velocity of channel i and T_i is the associated intensity. The resulting surface is dominated by the rotation curve of the galaxy. While many other studies use a tilted-ring analysis, where annuli in the velocity surface are independently fit, Sellwood and Sánchez (2010) show that DISKFIT provides a superior fit for a set of the THINGS galaxies analyzed by Trachternach et al. (2008). I restrict the rotation modeling presented here to the $m = 1$ term of a harmonic decomposition — DISKFIT can also fit higher-order terms, however previous results show that there are no large non-circular excursions within M33's disk up to $\sim$ 10 kpc (Putman et al., 2009; Corbelli et al., 2014). The model for each term is given by

$$V_{\rm circ}(\theta) = V_r \cos(\theta) \sin(i) \tag{3.2}$$

Table 3.1: Galactic properties of M33 from DISKFIT. All physical values are in agreement with previous work. The 1σ errors are found using 200 bootstrap iterations (Sellwood and Sánchez, 2010).

	Value	Error	Unit
Position Angle	201.7	0.5	deg
Inclination	54.6	2.7	deg
Systematic Velocity	-180.6	0.8	$\mathrm{km\,s^{-1}}$
Right Ascension	23.120	0.003	deg
Declination	30.366	0.005	deg

where θ is the position angle relative to the galactic centre, i is the inclination of the disk, and V_r is the rotational velocity at the radius r. The resulting disk properties of M33 are shown in Table 3.1. For 38092 degrees of freedom, the minimum χ^2_{red} of this fit is 6.8. This would seem to point to a bad fit, however the χ^2 calculation in DISKFIT relies on a user-supplied typical ISM dispersion, Δ_{ISM}, to account for expected scatter in the velocity surface (Spekkens and Sellwood, 2007). As discussed in §1.3, estimates of the "central velocity" are uncertain in themselves and increasing Δ_{ISM} only acts to hide this uncertainty by artificially lowering χ^2_{red}. The residual map in Figure 3.1 does not show large-scale residuals from remaining rotational velocities not accounted for in the model, and so I consider this to be a valid rotation model.

Calculating the correct uncertainties for rotation models is non-trivial. Errors based on the covariance matrix tend to severely underestimate the true uncertainties (Trachternach et al., 2008). DISKFIT addresses this issue by deriving errors on the fit parameters using a regularized bootstrap technique described by Sellwood and Sánchez (2010), which takes correlation between pixels into account. I use 200 iterations bootstrap iterations to calculate the errors in Table 3.1 These iterations require computational time of several hours (DISKFIT is not parallelized); a check with only 100 iterations showed no significant differences. The inclination, position angle, and systematic velocity (Table 3.1) agree with the respective values found by Corbelli et al. (2014).

While DISKFIT produces a model surface, the independent rotation veloc-

Table 3.2: Fit values of the rotation curve in Figure 3.2 using Equation 3.3. Errors are the 1-σ value derived from the covariance matrix of the least squares fit.

	Value	Error	Unit
n	0.52	0.03	–
$V_{\rm max}$	109.4	1.4	$\mathrm{km\,s^{-1}}$
$R_{\rm max}$	13.2	1.3	kpc

ity in each of the rings leaves artifacts in the residual velocity field. To remove these and create a smooth version, I fit a Brandt rotation curve (Brandt, 1960) following Meidt et al. (2008). This has an analytical form of

$$V_{\rm rot}(R) = \frac{V_{\rm max}\,(R/R_{\rm max})}{(1/3 + 2/3\,(R/R_{\rm max})^{n})^{3/2n}} \tag{3.3}$$

where $V_{\rm max}$ is the maximum rotation velocity, $R_{\rm max}$ is the radius at which $V_{\rm max}$ occurs, and n determines how rapidly the curve becomes Keplerian ($V \propto R^{-1/2}$). The fit parameters are shown in Table 3.2, and the fit is plotted in Figure 3.2. I note that the parameter values from the fit are not used to derive any results, only to produce a smooth approximation of the rotation curve.

Several rotation curves of M33 are available in the literature, the most recent of these being the thorough analysis of Corbelli et al. (2014). Figure 3.3 shows a comparison to the rotation curve I find to that from Corbelli et al. (2014). The curve of Corbelli et al. (2014) is systematically above the one I find over a large extent of M33's disk. Due to the different datasets and approaches used, there are multiple reasons for this disagreement. Corbelli et al. (2014) constructs the velocity surface using the *peak velocities* - the velocity at the maximum of the spectrum - instead of the line centroids, as I have done here. Using the peak velocities will make the velocity surface more susceptible to noise, since only the peak position is taken into account rather than the whole spectrum as is done for the centroid. Corbelli et al. (2014) calculate a rotation curve based on a centroid H I surface (see their Figure 6), but ultimately adopt the peak

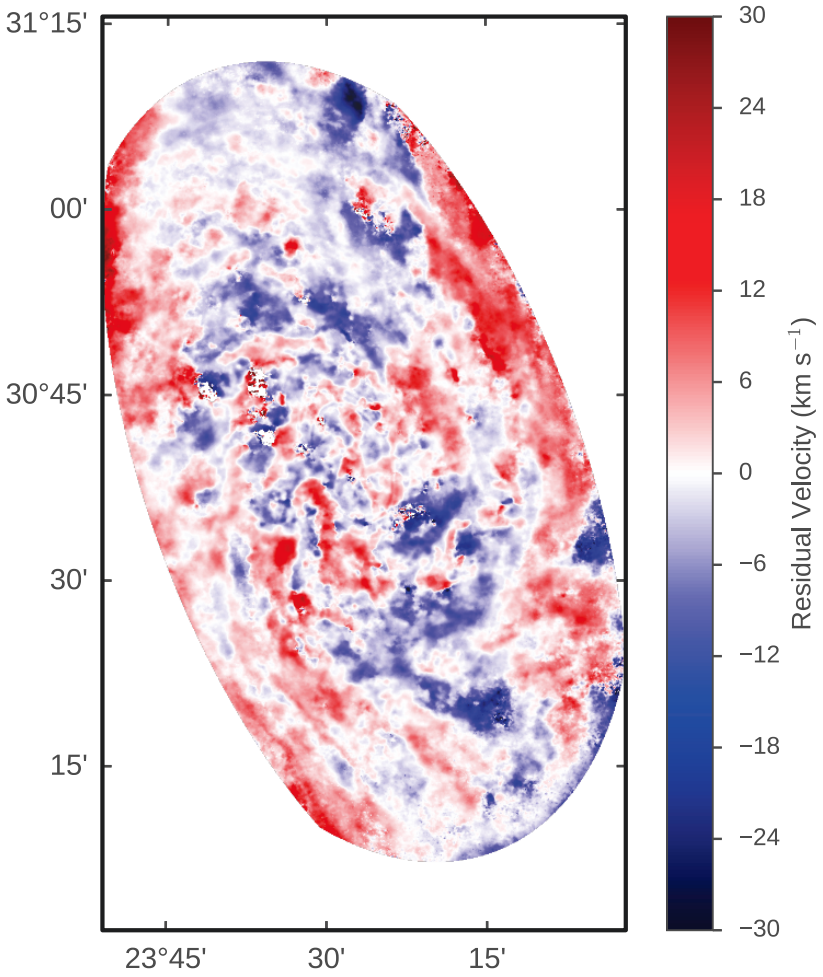


Figure 3.1: The velocity residual surface. This is the difference between the centroid surface and the rotation model from DISKFIT. The Northern plume visible in the last panel of Figure 3.4 clearly shows up as a large velocity deviation. The map here has been masked by hand to only include regions with emission, since DISKFIT uses the entire surface while fitting.

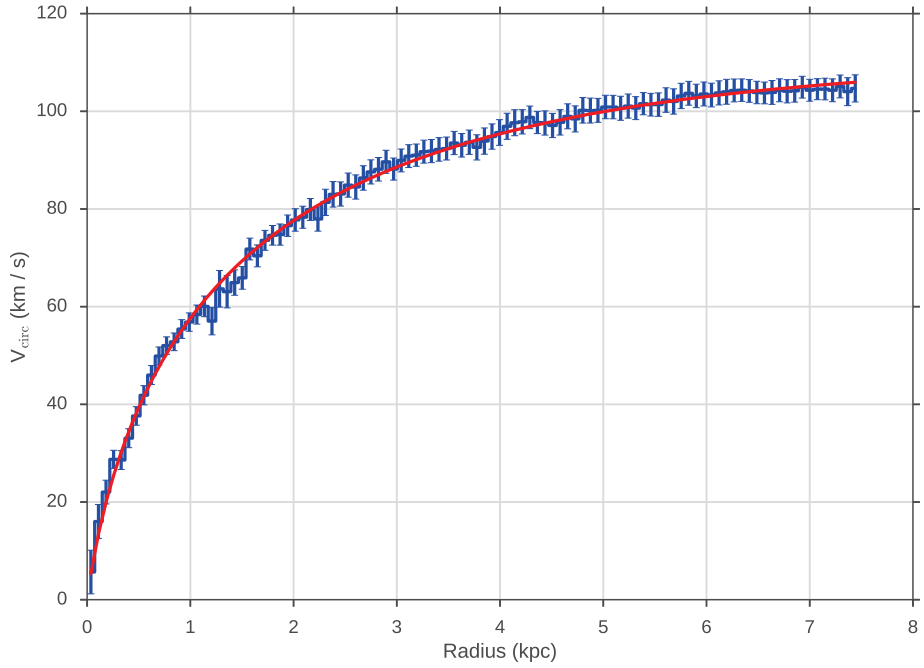


Figure 3.2: The rotation curve for the centroid surface of the H I cube using DISKFIT (blue) and the best fit modeled with Equation 3.3 (red). The error on each point is found using the bootstrap method of Sellwood and Sánchez (2010).

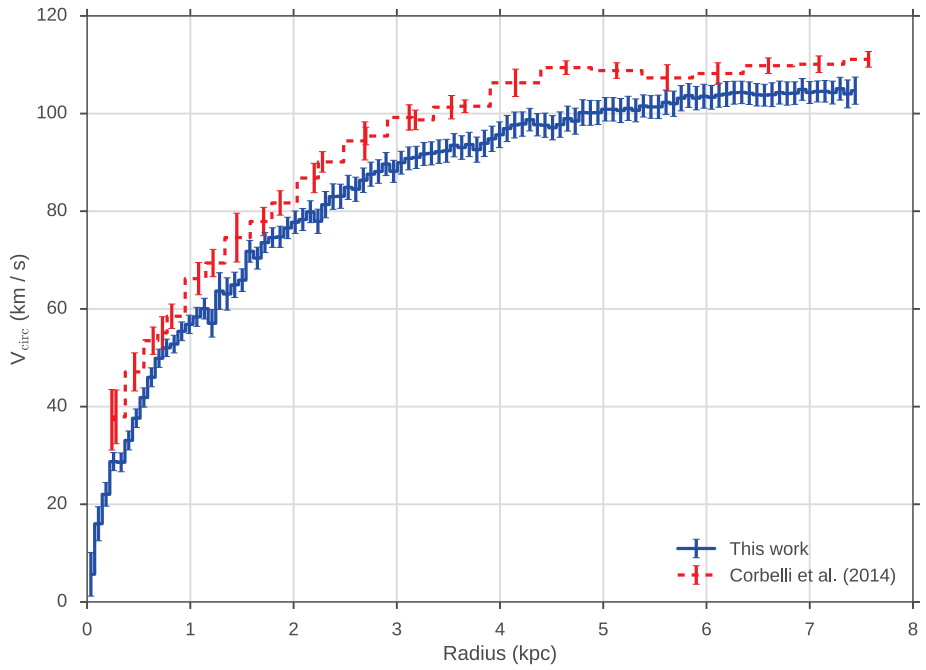


Figure 3.3: A comparison of the rotation found by DISKFIT with my H I data (blue solid) and the rotation curve published by Corbelli et al. (2014) (red dashed) using a smoothed version of the archival H I data discussed in §2.4.

velocities for their rotation curve out to 9 kpc. Their rotational velocities using the centroid values are indeed lower than those found using the peaks. I believe the author's decision in choosing this rotation curve stems from its agreement with the rotation curve found from the peak velocities of CO.

Using the fit values in Table 3.2 and Equation 3.2, I generate a smooth rotation surface for M33. I then calculate the difference between the model and the systematic velocity of the galaxy and shift each spectrum within the full cube using that difference. The result is a datacube with the circular motions removed. This, along with the residual centroid surface, emphasize the regions with significant non-circular motions, such as the spiral arms, regions of active star formation, and signs of disk warping in the outer edges. The residual centroid map is shown in Figure 3.1 and a series of spectral channels in the rotation-subtracted datacube are shown in Figure 3.4. These panels show the progression of multiple features within the disk. Proceeding from positive to negative velocities, there appears to be a progression in intensity from the outer edges of the disk inward (i.e. the disk begins to appear "skinnier"). This is potentially due to the warp in M33's disk. Corbelli et al. (2014) finds that the position angle of the disk deviates significantly from the value in the inner disk at a radii beyond $\sim$ 8 kpc (see their Figure 7). Since I assume a single overall value for the position angle and inclination of the disk, this change is not accounted for in the rotation model in Figure 3.2; this is also near the radius where the noise begins to dominate in my data. A clear visualization of the disk warp is shown in Figure 8 of Putman et al. (2009).

Figure 3.4: Channel maps of the rotation subtracted H I cube are shown. Each panel is the integration of 10 channels, or a velocity of 2 km s^{-1}. The annotation in each window is the velocity range in km s^{-1}, centered on the systematic velocity of -180.6 km s^{-1} (Table 3.1). Each map is shown with an arcsinh stretch to highlight the diffuse features.

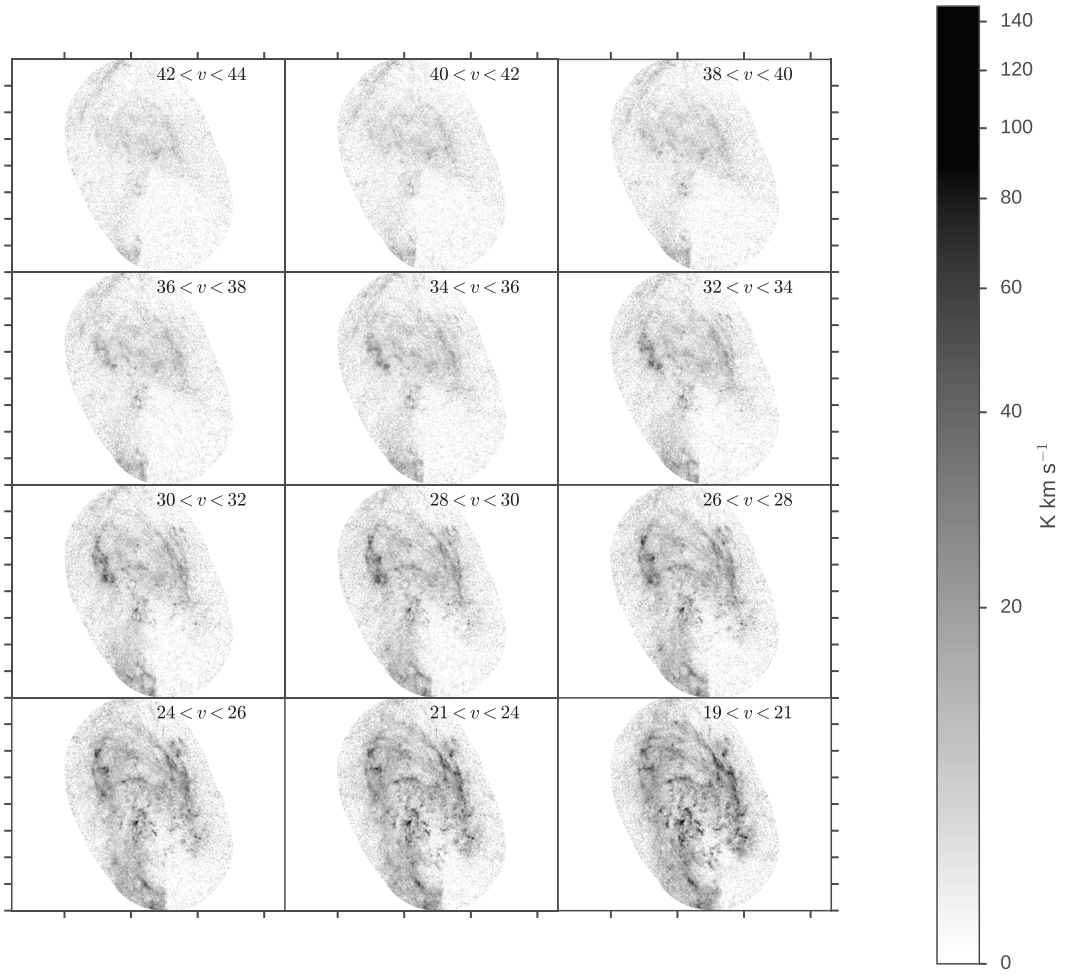


Figure 3.4: *cont.*

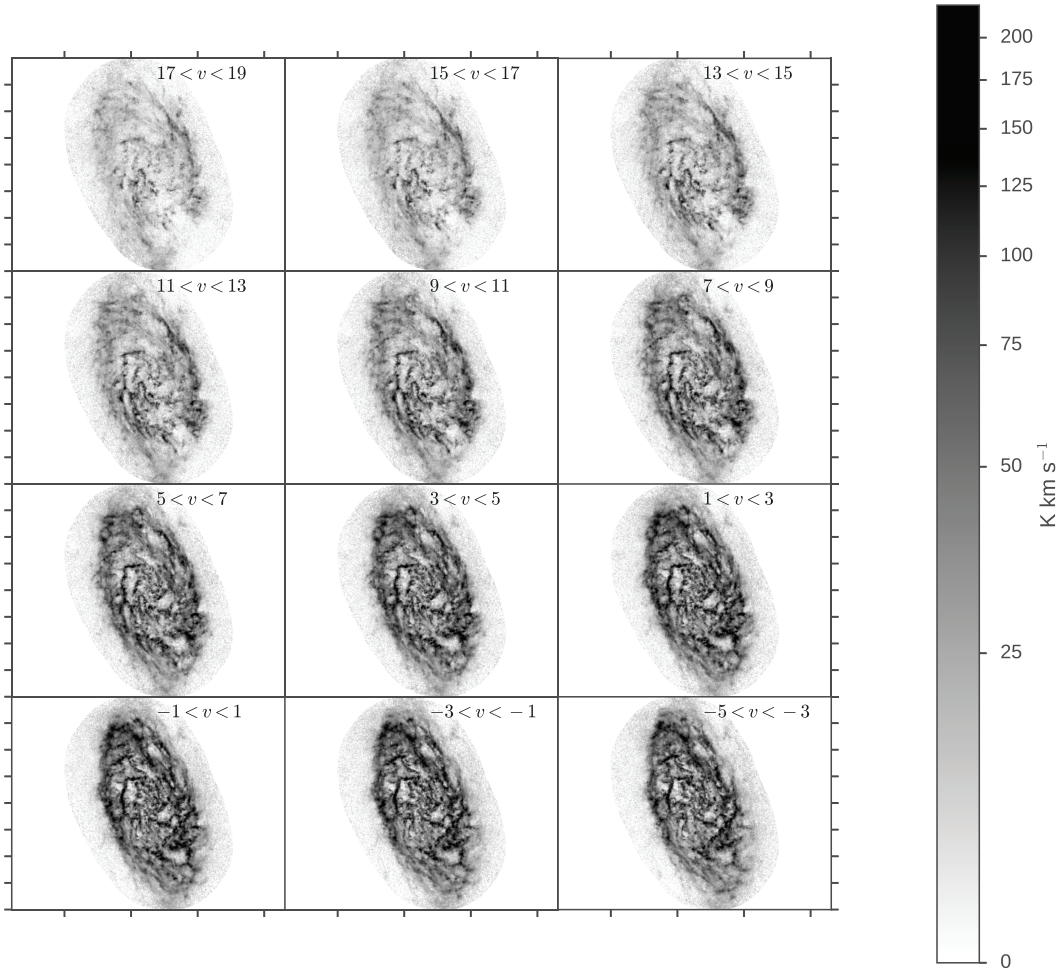


Figure 3.4: *cont.*

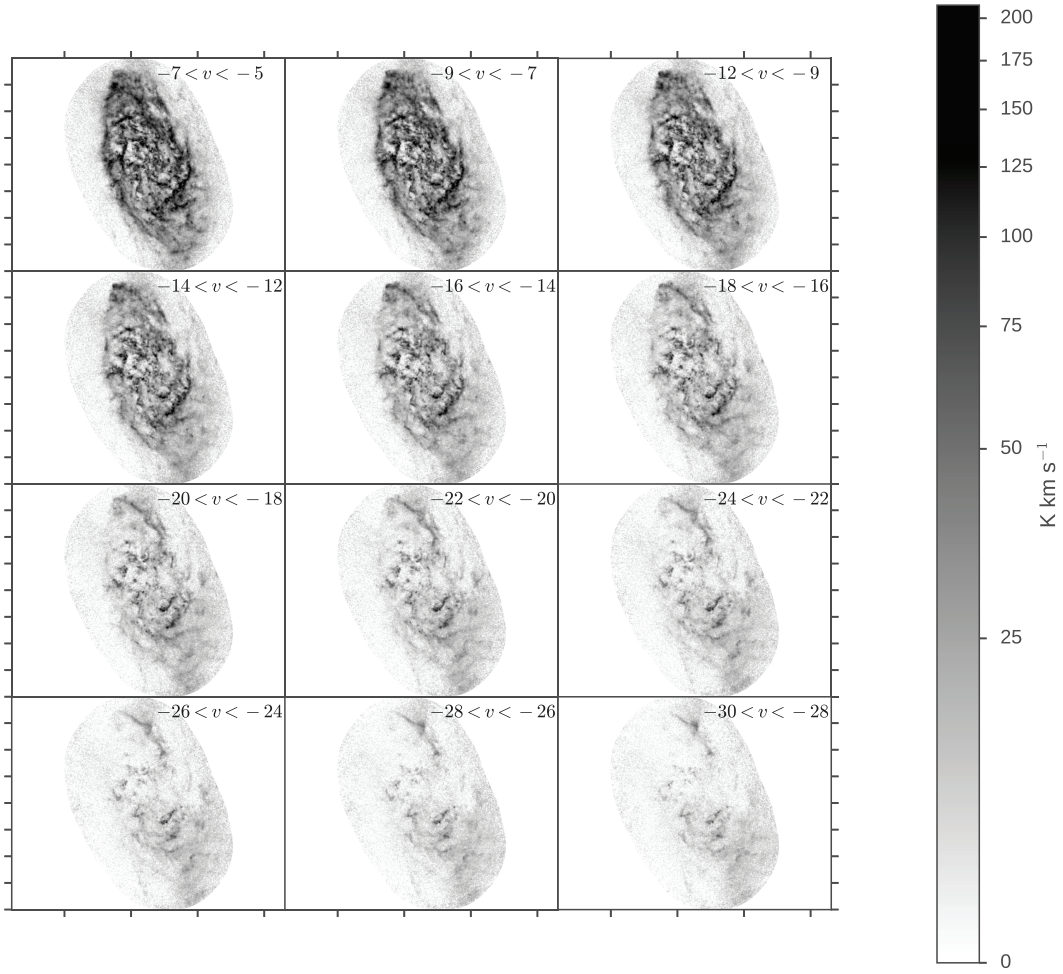
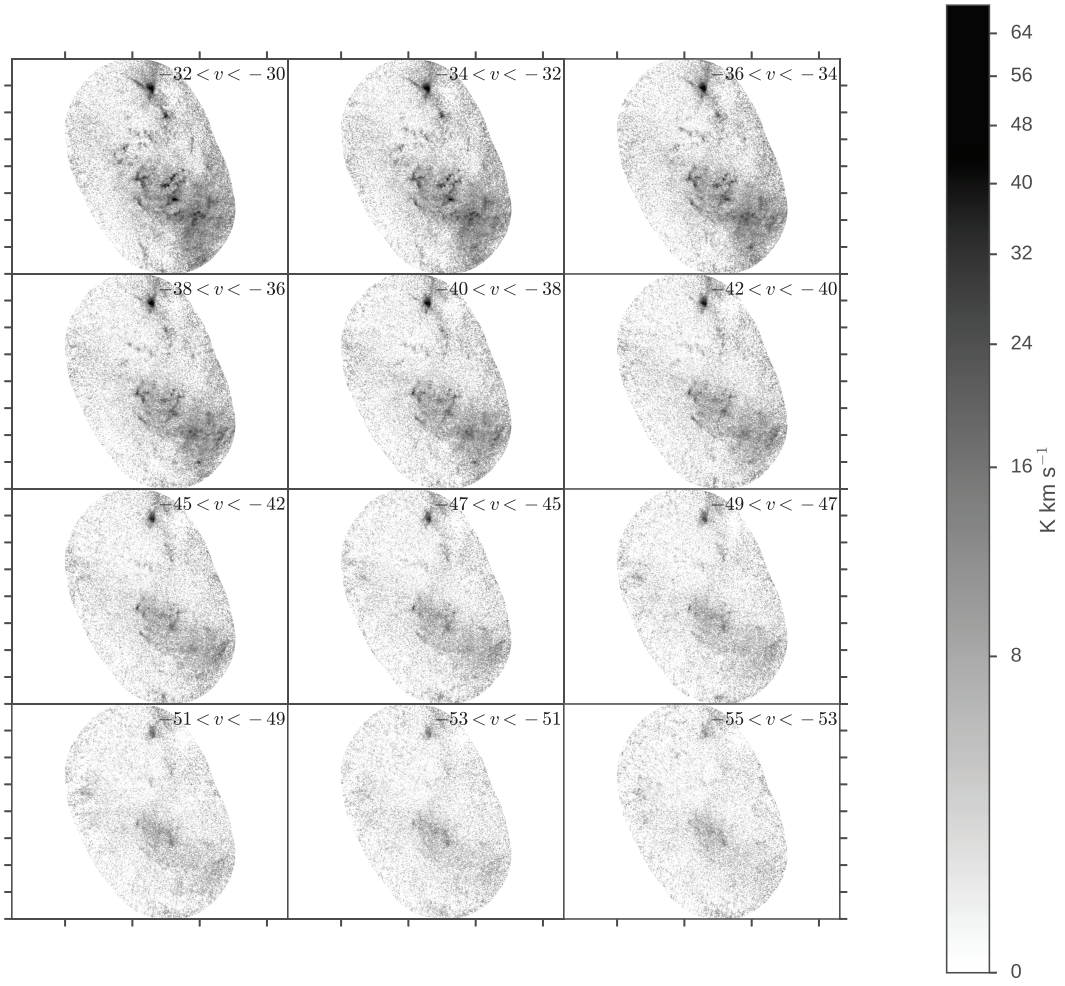


Figure 3.4: *cont.*



An Adequate Representation of Rotation? — Calculating rotational models based on a centroid surface is relatively easy and computationally efficient, but does it correctly capture enough of the velocity variations within the entire datacube? As I argue in §3.5, these new observations show significant line structure, and a single Gaussian model (for which the centroid is an approximation) is not capable of properly describing the spectral variations. An improvement can be made by extending the tilted-ring analysis to use the entire datacube, instead of a velocity surface. This is the approach used in the 3DBarolo algorithm by Di Teodoro and Fraternali (2015). I attempted to use this algorithm, but was unable to get the algorithm to run due to the size of the datacube. When running on the smaller, archival datacube, I found the computational time to be unfeasible, often taking several days to calculate scaled-down versions (using large ring widths) of the rotation curve. This was without adding error finding in, which slowed the algorithm down by about a factor of 60! While certainly a step in the right direction, this 3D modeling approach needs to be modified for handling large datasets and extended to run on multiple cores, ideally in a cluster environment.

3.3 Global Line Properties

The H I datacube contains approximately two million spectra across M33. By combining these spectra, global properties of the neutral ISM in M33 can be examined. In Figure 3.5, I present the total spectrum of M33 with its rotation curve, which shows the traditional dual peak morphology. The edges of the spectrum are not symmetrical; the approaching side of the galaxy appears slightly cut-off. Below an LSRK velocity of around -75 km s^{-1}, significant Galactic H I emission becomes an issue. I chose this cut-off by examining individual channel maps to ensure all emission contained within the cube originates from M33 and not stray Galactic emission. Regions with significant Galactic

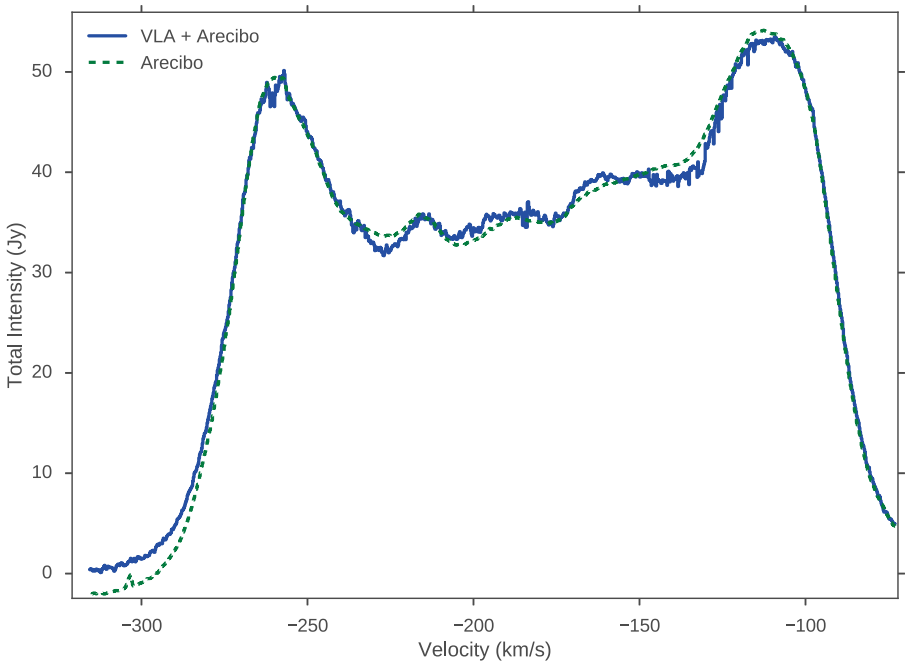


Figure 3.5: The total spectrum of M33 with the rotation curve is shown for the combined VLA and Arecibo data, and the Arecibo data alone.

H I emission are also not properly flux calibrated, since the flux density calibrator shows signs of absorption at those velocities.

By integrating over the total spectrum in Figure 3.5, I find the total H I mass in the disk to be 1.42×10^9 $M_\odot$. This values agrees with the masses found by Putman et al. (2009), Gratier et al. (2010) and Corbelli et al. (2014), who all find values of $1.4 - 1.5 \times 10^9$ $M_\odot$.

Using the rotation subtracted datacube, a total spectrum for the galaxy containing only the velocity residuals can be found. The rotation-subtracted spectrum for M33 is shown in Figure 3.6. I model the spectrum with two Gaussians with their means tied together. The parameters from the fit are shown in Table 4.1. A two-Gaussian model is used due to the expectation that the neutral atomic medium contains two distinct phases — the CNM and WNM — whose thermal line widths should be significantly different. The preference of a two-Gaussian model over a single Lorentzian profile was investigated by

Ianjamasimanana et al. (2012) using the THINGS galaxies (Walter et al., 2008). This preference was determined by comparing the relative χ^2 values of the fits. They find widths of the velocity profiles ranging from 3.4–8.6 $\mathrm{km\,s^{-1}}$ and 10.1–24.3 $\mathrm{km\,s^{-1}}$ for the narrow and wide components, respectively. The errors are estimated to be around 20% mostly due to the finite channel widths. The widths found here for M33 are both larger than these ranges, despite the vast increase in spectral resolution (0.2 $\mathrm{km\,s^{-1}}$ here versus 2.6 or 5.2 $\mathrm{km\,s^{-1}}$ in THINGS). Some artificial line broadening is expected due to uncertainties in the centroid velocities (§1.2).

In this model, the CNM is likely represented by the narrow component, though this modeling is crude for the reasons discussed below. However, Ianjamasimanana et al. (2012) provide evidence for this by finding correlations between the ratio of the amplitudes of the components and the ratio between the surface density of H_2 (from CO) and HI. This suggests a connection between the narrow component and the molecular medium. The ratio of the amplitudes here is ~ 2 (though see below for uncertainties), greater than any in the Ianjamasimanana et al. (2012) samples by a factor of 0.5. The relative widths of the components, however, are consistent with their sample; the widths are different by a factor of ~ 0.36. Integrating over the spectrum shows that the components contribute almost equally to the total intensity. The wide component contributes slightly more to the total intensity with a fraction of 0.56.

Identifying the wide component with the WNM is tenuous at best. The thermal line width of 10^4 K gas is ~ 8 $\mathrm{km\,s^{-1}}$ (§1.3), a factor of three smaller than the width of this wide component. Thus the large line width likely also represents a combination of other effects, such as turbulent motion, outflows/inflows, and radial flows along the spiral arms. The line profiles shown in §3.5 also demonstrate that there may be significant uncertainties in the line centroid when multiple strong line components are present. Those uncertainties will

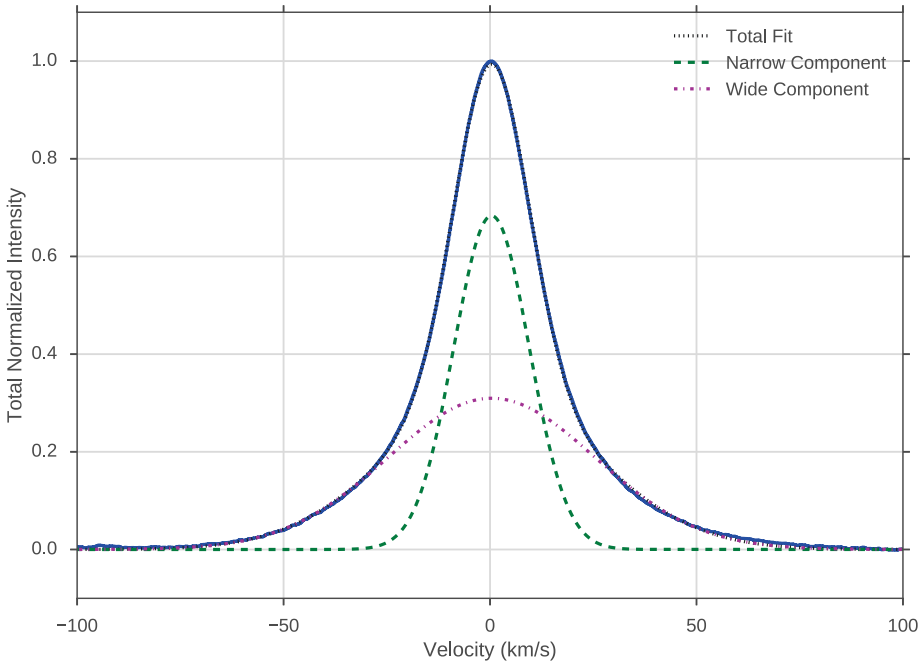


Figure 3.6: The total spectrum of M33 with the rotation curve subtracted. A two-Gaussian model is fit to the spectrum and overplotted in the figure. The intensity has been normalized to unity to demonstrate the relative components in the model.

contribute more to this wide component than the narrow component.

Finally, it should be noted that modeling the CNM in this manner is necessarily crude since: (1) our data lacks the resolution to fully resolve CNM clouds, expected to be found on $\sim$ 20 pc scales (Braun et al., 2009); and (2) simply performing a sum over all spectra implicitly assumes that the H I emission is optically thin (Equation 1.8), which is only a good assumption for the WNM component. Because of point (1), the components of the spectrum due to the CNM is a superposition of cold unresolved H I clouds, which may artificially increase the width. An analysis similar to that performed by Braun (2012) is required to address the optical depth issues in (2), though this too suffers from resolution effects without the inclusion of higher-resolution data.

Table 3.3: Fit parameters for the two-Gaussian model shown in Figure 3.6. Errors are the 1-σ value derived from the covariance matrix of the least squares fit. The means are tied together in the fit. The additional error 0.20 $\mathrm{km\,s^{-1}}$ in velocity is the channel width in the data, accounting for rounding errors when the spectra are shifted in the rotation subtraction.

	Value	Error	Unit
Mean	0.38	0.001 + 0.20	$\mathrm{km\,s^{-1}}$
Narrow Amplitude	0.683	0.002	Normalized Intensity
Wide Amplitude	0.310	0.002	Normalized Intensity
Narrow Standard Deviation	9.12	0.02 + 0.20	$\mathrm{km\,s^{-1}}$
Wide Standard Deviation	25.16	0.06 + 0.20	$\mathrm{km\,s^{-1}}$

3.4 Radial Profiles

Similar to the ring analysis used to find the rotation curve (§3.2), the integrated intensity and line width maps can be averaged over concentric rings to find average radial profiles across the disk.

Surface Density — The traditional surface density profile is weighted by the area:

$$\langle \Sigma \rangle^{\mathrm{A}} = \frac{\sum_{\mathrm{A}} \Sigma \delta A}{\sum_{\mathrm{A}} \delta A}, \tag{3.4}$$

where A is the area of the ring, δA is a pixel area (though the true independent area element is the beam; see below) and Σ is the mass surface density. The integrated intensity, or the zeroth moment, is simply the sum over the spectral dimension, multiplied by the velocity channel width,

$$I = \sum_i T_i \delta v_i, \tag{3.5}$$

Assuming that the emission is optically thin, the column density is independent of the spin temperature of H I (§1.2). It is then a simple extension to convert column density into the surface density of the gas; the conversion factor is 0.019 $\mathrm{M_\odot\,pc^{-2}/K\,km\,s^{-1}}$. The surface density profile using rings with a width of 100 pc for H I is shown in Figure 3.7 using the combined VLA and Arecibo cube

and the Arecibo cube alone. The errors are the standard deviations about the mean, scaled by a factor of $\sqrt{N_{\rm beam}/N_{\rm ring}}$, the number of pixels in the beam over the number of pixels in the ring, to reflect the approximate number of independent measurements averaged over (Druard et al., 2014). There is good agreement between the two datasets, as expected once the correction factor has been applied (§2.4). The noticeable peaks near 2 and 4 kpc correspond approximately to the radius where the spiral arms are found. In Figure 3.8, I compare the profile I find versus the one found by Corbelli et al. (2014). There are some differences between the two profiles, which may due to the differences in how each was computed. Corbelli et al. (2014) computes their profile directly from their tilted ring analysis which yields their rotation curve and thus has some dependence on the model. Furthermore, their rotation curve modeling used a version of the archival VLA data smoothed to a resolution of 30 arcseconds. Given that the Arecibo data compared to in Figure 3.7 shows less variation then the Corbelli et al. (2014) profile, the difference in resolution is not likely to be the significant cause of the differences.

To investigate these fluctuations in the surface density further, I create separate profiles of the North and South halves, defined by the position angle of the disk (Table 3.1). The significant drops in the Northern half about the peak are certainly due to material cleared by the Northern spiral arm. The Southern spiral arm is less pronounced than the Northern arm, so there is not as significant of a drop seen. The Southern half has, on average, a greater surface density than the North; this point is visited in more depth in §4.4. Due to the small rings within 1 kpc, the variations between the North and South are not significant within that region.

Studies of star-formation on large kpc scales are typically unable to resolve individual individual H I clouds (Leroy et al., 2008) and must resort to statistical descriptions of the intensity to attempt to understand what is occurring on smaller scales. One of these methods is to calculate a *clumping factor*. As

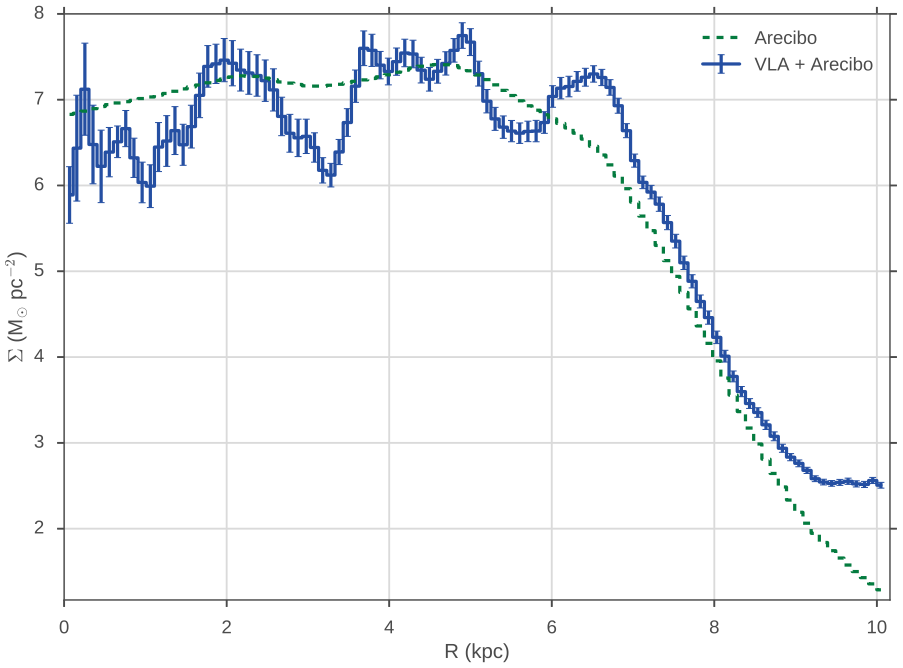


Figure 3.7: The surface density profiles of the combined VLA and Arecibo data (blue solid) and Arecibo (green dashed). The combined data has had the correction factor applied.

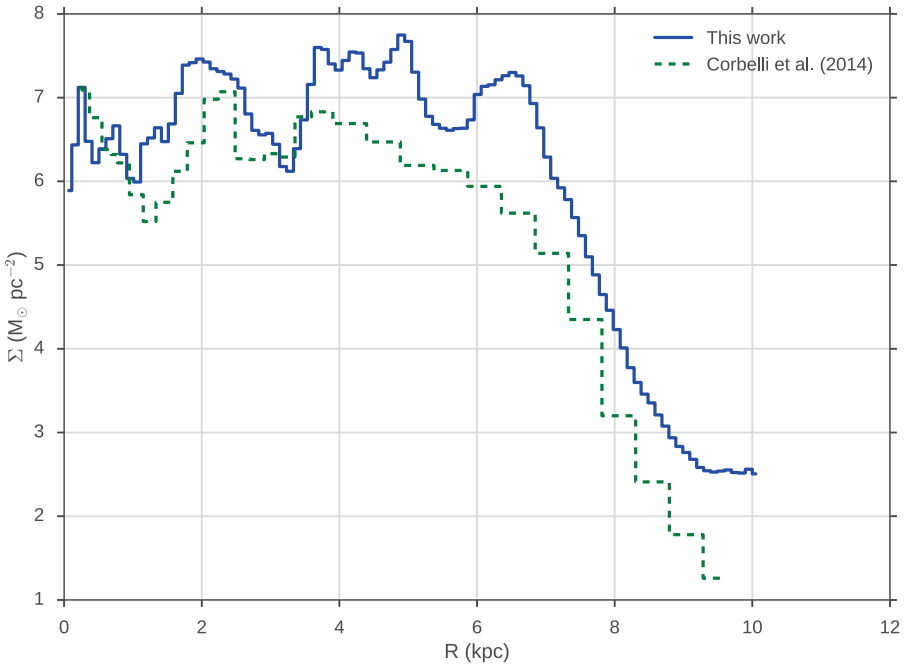


Figure 3.8: The surface density profiles of M33 from this work (blue solid) and Corbelli et al. (2014) (green dashed) are compared. Errors for the profile I find are shown in Figure 3.7; no errors are provided by Corbelli et al. (2014).

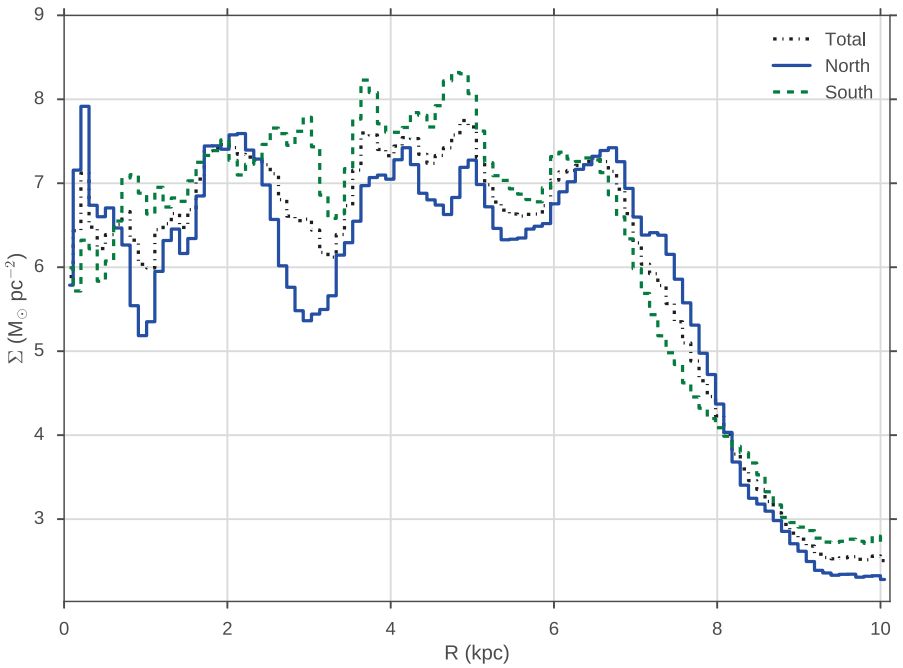


Figure 3.9: The surface densities of the North (blue solid) and South (green dashed) are plotted separately. The profile for the entire disk from Figure 3.7 is plotted in black (dash-dot). Errorbars are not shown in order to highlight the differences between the halves. The errors are approximately a factor of $\sqrt{2}$ larger than those in Figure 3.7.

introduce by Leroy et al. (2013a), the clumping factor relies on finding the mass-weighted surface density profile,

$$\langle\Sigma\rangle^{\mathrm{M}} = \frac{\sum_{\mathrm{A}} \Sigma^2 \delta A}{\sum_{\mathrm{A}} \Sigma \delta A}. \tag{3.6}$$

Leroy et al. (2013a) then defines the clumping factors as the ratio of the surface densities,

$$c = \frac{\langle\Sigma\rangle^{\mathrm{M}}}{\langle\Sigma\rangle^{\mathrm{A}}}, \tag{3.7}$$

or equivalently the ratio of the squared integrated quantity divided by the integrated quantity squared. This is called a clumping factor since $\langle\Sigma\rangle^{\mathrm{M}} > \langle\Sigma\rangle^{\mathrm{A}}$ when averaging over small, dense regions spread throughout a much larger diffuse medium. A smooth medium will have $\langle\Sigma\rangle^{\mathrm{M}} \approx \langle\Sigma\rangle^{\mathrm{A}}$. I calculate a clumping factor for the 100 pc wide rings used in Figure 3.7 and find that the H I emission throughout the disk is smooth, with clumping factors around 1.11 ± 0.03. These results are shown in Figure 3.10. This concurs with the results of Leroy et al. (2013a), who find that the H I clumping must be low based on smoothing a large set of data to a minimum of 250-500 pc scales. On those scales, they find an overall H I clumping factor of $1.34^{+0.32}_{-0.17}$. The average clumping factor across M33 on 100 pc scales is slightly below their 1σ error range, but the difference is not large and this may be due to not first smoothing my data to large size scales.

Velocity Dispersion — A second radial profile of interest is that of the average line widths, which reveals trends in the local velocity dispersion of the neutral gas. The dispersion profile shown in Figure 3.11 is computed from the line width, which is the square root of the second moment,

$$\sigma_v^2 = \frac{\sum_i T_i \left(v_i - \bar{v}\right)^2}{I}, \tag{3.8}$$

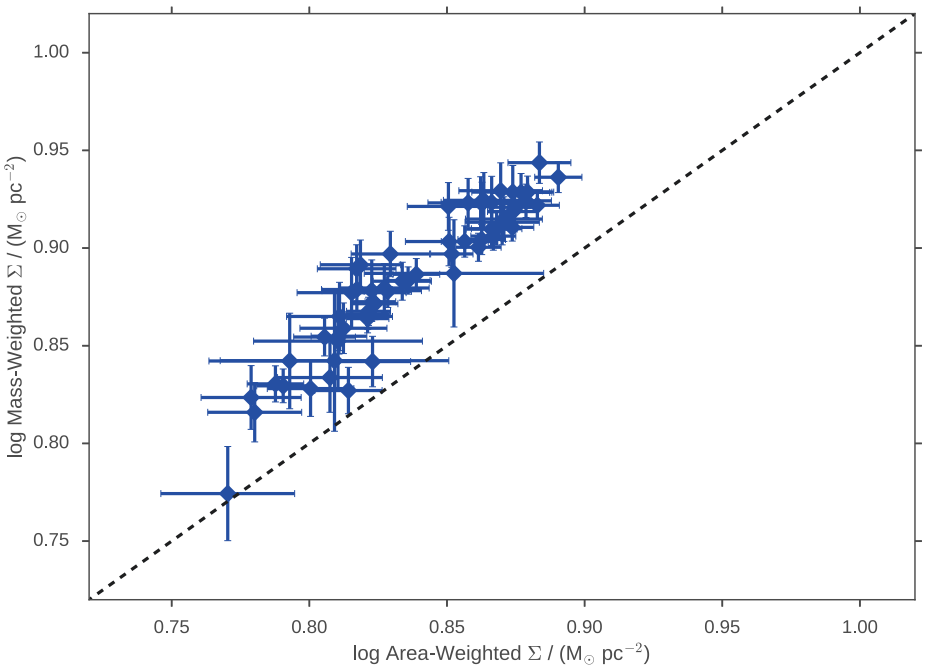


Figure 3.10: The area-weighted surface densities (Equation 3.4) are plotted against the mass-weighted surface densities (Equation 3.6). The ratio is the clumping factor (Equation 3.7), which is around 1.1 across the disk. The dashed line indicates where $\langle\Sigma\rangle^{\mathrm{M}} = \langle\Sigma\rangle^{\mathrm{A}}$ (i.e., where $C = 1$).

where I is the zeroth moment (Equation 3.5) and $\bar{v}$ is the centroid (Equation 3.1). This is an approximation of the standard deviation of a Gaussian. Figure 3.11 shows that the velocity dispersion does not vary significantly across the entirety of the disk; the dispersion is consistent with 10 $\mathrm{km\,s^{-1}}$ over much of the disk, though there is perhaps a weak decline with increasing radius. A decline in the dispersion with increasing radius is found by Ianjamasimanana et al. (2015) across a sample of 22 galaxies.

There are two significant drops in velocity dispersion near 1 and 4.2 kpc. The central region of M33 is dominated by the stellar component (Corbelli et al., 2014) and the H I morphology is arranged into a series of wind-driven shells. This sudden drop may signify a change in the dominant large-scale structure mechanism, as the mid-disk region becomes dominated by the two spiral arms. The outer disk shows signs of a warp Corbelli et al. (2014) resulting from past interactions with M31 (Putman et al., 2009). Indications of different ISM structure throughout the disk have been noted before; Rosolowsky et al. (2007) find that the mass spectrum of GMCs changes within the central (< 2 kpc), mid ($2 - 4$ kpc), and outer regions of the disk. The highest mass clouds are all located in the mid-disk region, particularly in the Northern spiral arm.

If this is the case, a drop should be nearly independent of θ throughout the ring. In Figure 3.13, I show separate dispersion profiles for the North and South halves. Indeed, these drops are evident in both profiles, though both are more pronounced in the Northern profile where the spiral arm is more distinct. A more thorough modeling of the rotation curve is needed to further investigate this claim. M33's flocculent structure suggests that this transition to spiral-arm dominated regions will not be very strong and may indicate why no such deviations are evident in the rotation curve. This point is discussed in further detail in §4.4.

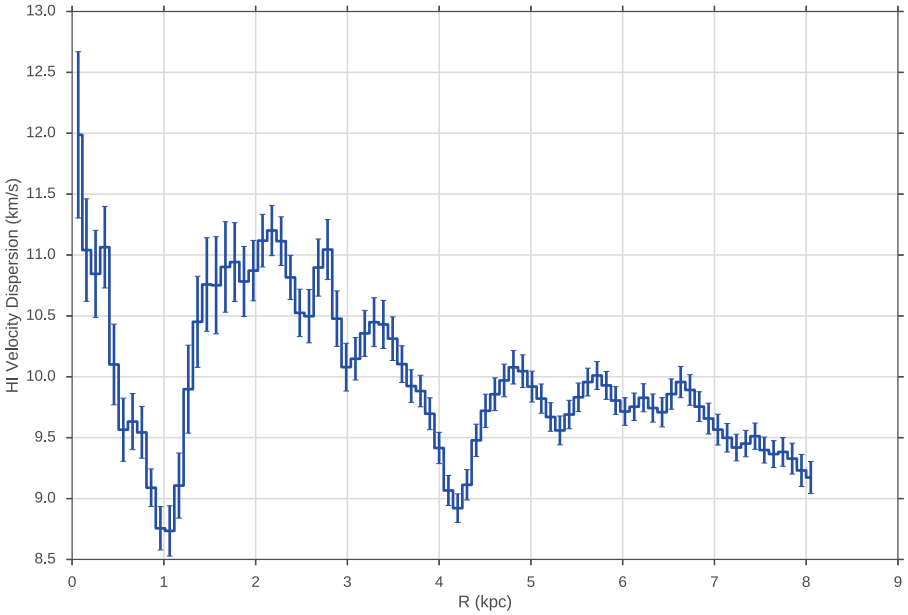


Figure 3.11: The velocity dispersion profile across the disk is shown. The errorbars are found as described in Figure 3.7

.

3.5 Higher Order Moments

Typical analyses of extragalactic H I observations are limited to computing the second moments from the spectra in the datacube. Even finding usable second moments can be difficult, however, and properly masking the data becomes essential for finding useful results (e.g., Walter et al., 2008; Mogotsi et al., 2016). The combination of high sensitivity and spectral resolution in the new data presented here alleviate some of these difficulties. Masking the data is still necessary, but most spectra through the disk have a sufficient number of non-masked elements to calculate an accurate line width value without the need for localized stacking. The line width properties are presented and discussed in §3.4.

The quality of the data allows for accurate representations of higher order moments to be calculated. The third and fourth moments - skewness and

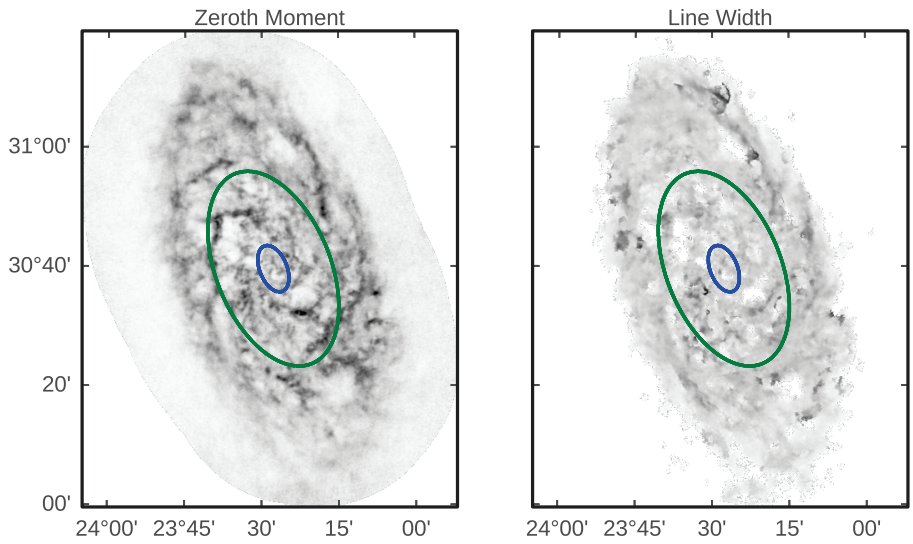


Figure 3.12: The zeroth moment map is shown with the location of the velocity dispersion drops in Figure 3.11 indicated. The inner, blue contour is at a galactic radius 1 kpc, and the outer green contour is at 4.2 kpc.

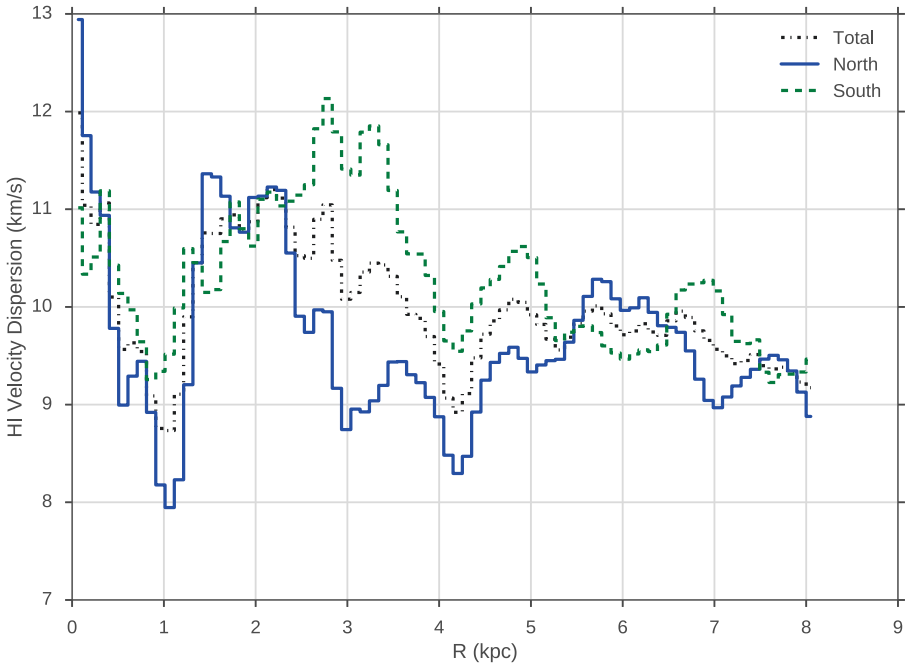


Figure 3.13: The velocity dispersion profiles in the North (blue solid) and South (green dashed) halves across the disk is shown. The profile over the whole disk from Figure 3.11 is also shown (black dash-dot).

kurtosis - are shown in Figure 3.14. Skewness is sensitive to the strength and direction of the tail in a distribution, and has the form,

$$S = \frac{\sum_i T_i \left(v_i - \bar{v}\right)^3}{I} \frac{1}{\sigma^3}, \tag{3.9}$$

where the first term is the third moment of the distribution. The kurtosis describes how significant the tails of the distribution are

$$K = \frac{\sum_i T_i \left(v_i - \bar{v}\right)^4}{I} \frac{1}{\sigma^4}. \tag{3.10}$$

Due to the fourth power, the kurtosis is quite sensitive to the shape of the mask. Some regions with beam-element shapes can be seen in Figure 3.14, however other areas continue to show interesting variations within regions of similar intensity, where the shape of the spectral mask will have less of an influence. Both of these quantities are dimensionless through the normalization of the line width with the appropriate power.

Both maps show several interesting features and it is clear that there is complex spectral structure throughout the disk. I discuss three of the interesting regions identified in the skewness and kurtosis maps: the spiral shock along the southern arm, the complex variations near the giant HII region NGC 604, and the curious Northern "plume" that is evident in the last panel of Figure 3.4. A common feature in the spectra presented for these three regions is the amplitude between two dominant components changing with respect to each other.

Figure 3.15 shows the location of the shock front moving through the Southern spiral arm. The shock is most evident in the skewness map, where there is a long continuous feature where the skewness flips signs. This appears as an elongated region in the line width and as an abrupt change in the kurtosis. Identifying the location of the shock front and its relative position from collections of molecular gas is important to constrain H_2 formation resulting from

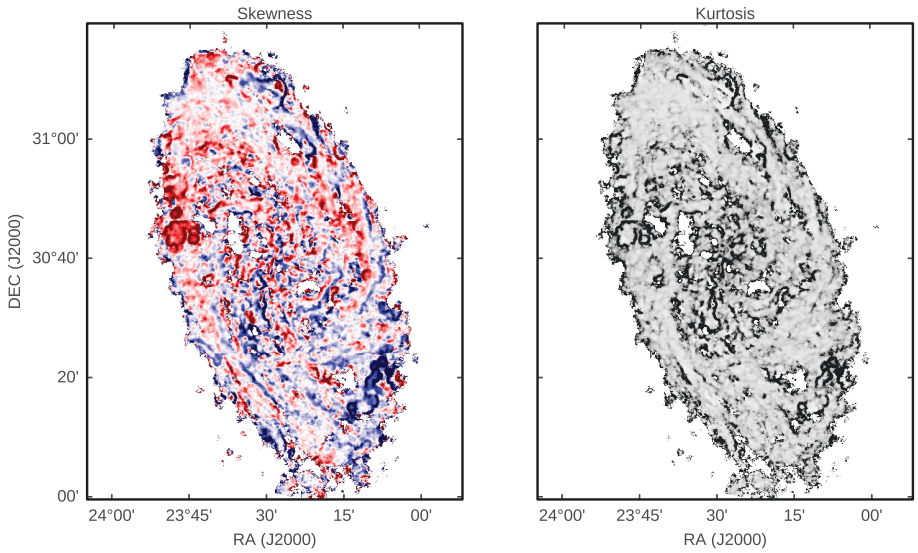


Figure 3.14: The skewness and kurtosis maps. The skewness is shown on scales between -3 and 3 as this highlights the variations across the disk. Similarly, the kurtosis map is scaled to be between 0 and 20. In the skewness map, the blue and red represent the negative and positive skew, respectively. The kurtosis is a positive quantity. Black regions have a higher kurtosis.

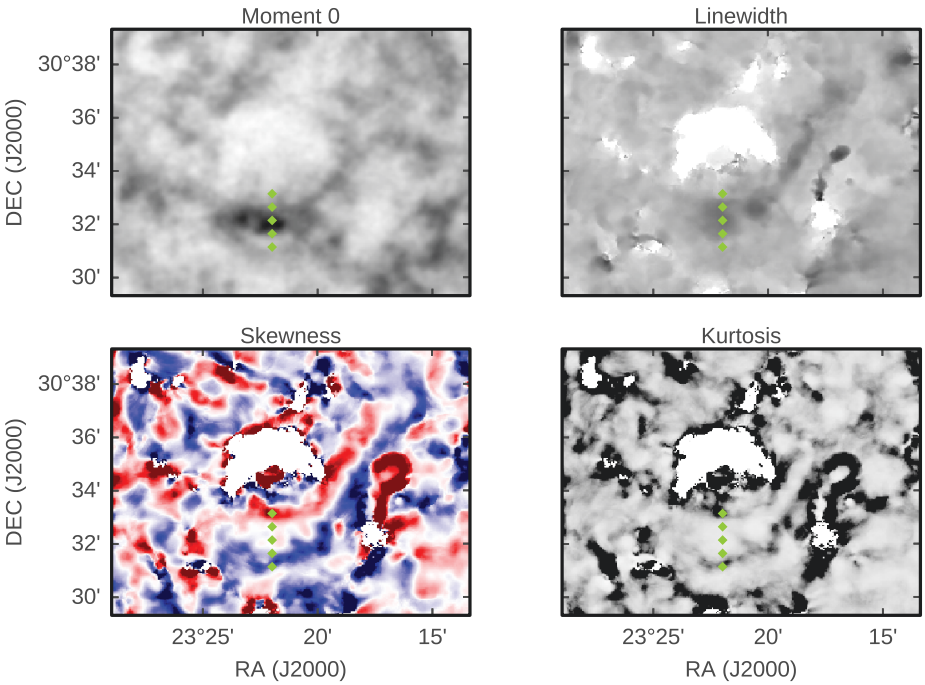


Figure 3.15: The zeroth moment, line width, skewness, and kurtosis maps around the southern spiral arm. There are broader line widths and a switch in the sign of the skewness along the arm due to the spiral shock. The kurtosis is enhanced only on one side of the arm. The points in each of the four frames correspond the the positions of the spectra in Figure 3.16.

compression of the passing shock front (Dobbs, 2015). This example shows the potential of using higher-order moments to identify and study the dynamical effects within spiral arms.

The most active star forming region in M33 is the giant HII region NGC 604. Embedded in the Northern spiral arm, this region has been the subject of multiple papers on sequential star formation (e.g., Muraoka et al., 2012) — a self-regulated propagation of star formation within wind-driven shells created by embedded massive stars (Elmegreen et al., 2002, §1.4). This is similar, though on smaller scales, to the model introduced by Palous et al. (1990). Figure 3.17 shows the interesting morphology of the higher moments around this region. Rapid small scale fluctuations in skewness are observed; fluctuations of this magnitude and frequency on small scales appear to be rare throughout

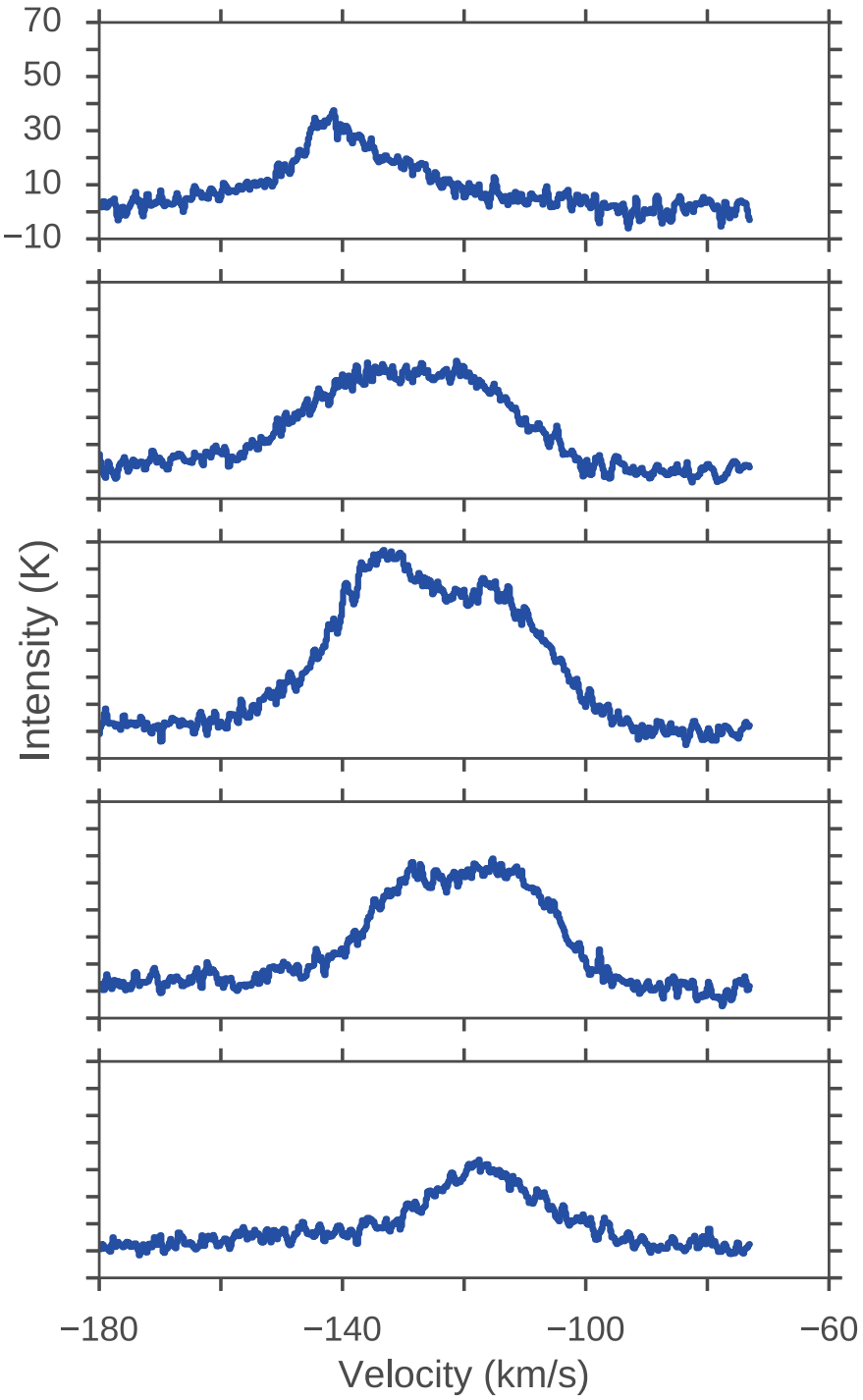


Figure 3.16: Spectra from the dots in Figure 3.15, proceeding from North to South. The y-axis is common across all of the spectra.

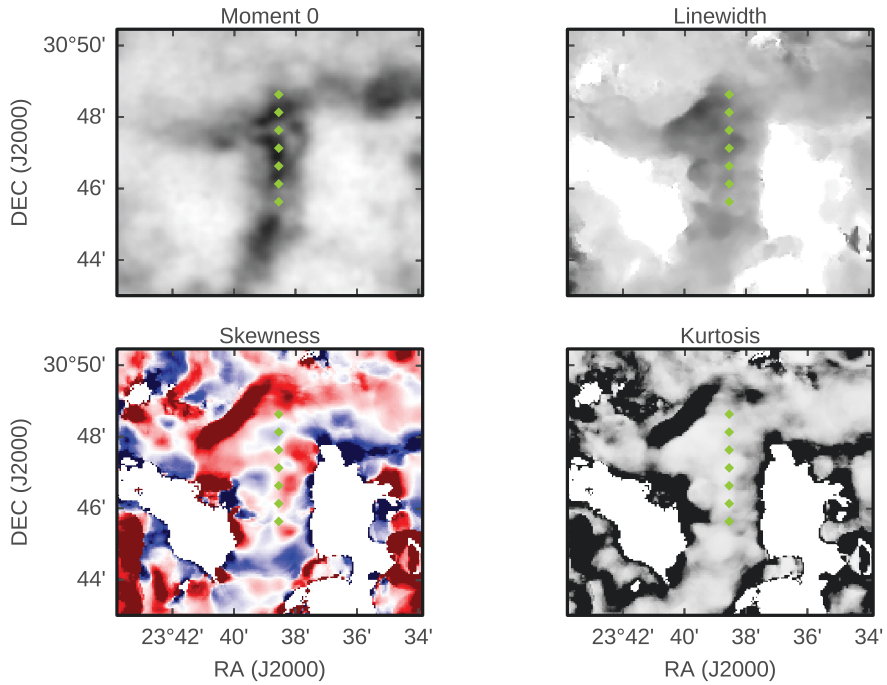


Figure 3.17: The zeroth moment, line width, skewness, and kurtosis maps around NGC 604. The points in each of the four frames correspond the the positions of the spectra in Figure 3.18.

the disk. Unlike for the southern arm, there are no clear structures in the line width or kurtosis. The enhanced kurtosis on the edges of the spiral arm are the result of the lack of signal along those lines-of-sight. The lack of clear signatures in the line width suggests there is complex morphology along the line-of-sight and that these features are unresolved at this resolution. Example spectra through the middle of this region are shown in Figure 3.18. These show a wide component, some with two peaks, whose peak brightness changes by 20 $\mathrm{km\,s^{-1}}$ multiple times over small spatial separations. There are clearly interesting spatial and spectral variations throughout this region. Figure 3.17 also shows a bar-like feature to the North that shows high positive skewness and kurtosis. This feature occurs where the Northern arm "kinks".

Several ring-like structures appear in the outer disk in both maps. These are potentially signatures of impacts in the disk from infalling H I clouds, analo-

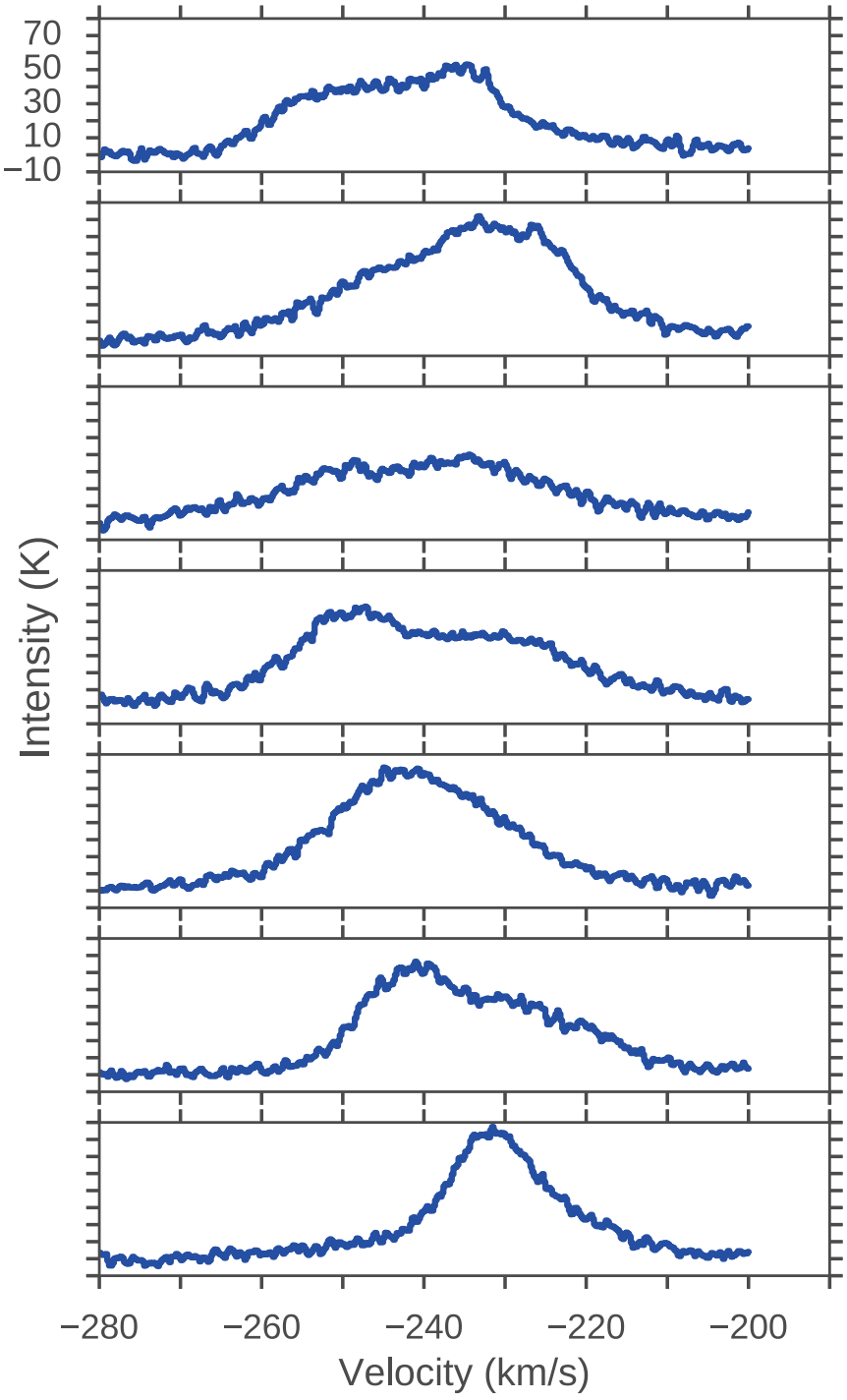


Figure 3.18: Spectra from the dots in Figure 3.17, proceeding from North to South. The y-axis is common across all of the spectra.

gous to the high-velocity clouds (HVCs) observed within the Milky Way (Heiles, 1984). These regions do not host significant star formation, making the opposite case of tracing H I outflows unlikely since there is no source to drive them. The impact of HVCs on a galactic disk is one of the proposed formation mechanisms for H I super-bubble in spiral galaxies (Heiles, 1984; Tenorio-Tagle and Bodenheimer, 1988). These features are not obviously connected to bubble structures in the disk, however, indicating that these features are relatively young or the material from the material HVC integrated along the line-of-sight hides the bubble morphology. In Figure 3.19, the zeroth moment, line width, skewness and kurtosis maps are shown for the Northern plume evident in the last panel of Figure 3.4. The spectra in Figure 3.20 correspond to the diamonds in the aforementioned plot to highlight the spectral changes giving rise to the interesting behaviour in the higher moments. The plume itself appears as a broader feature centered around -280 km s^{-1} and is not clearly distinguished from the galactic disk emission. There is, however, a bright narrow feature superimposed on the blue-side. The peak of this narrow feature rapidly changes velocity, becoming more spectrally separated from the plume nearer the interface of the enhanced kurtosis.

How Valid is the Gaussian Form Assumption? — The example spectra shown in this section have a common theme: they are not Gaussian and show multiple kinematic components. This suggests that the assumption that a single Gaussian shape, which underlies the definitions of the moments, is not valid in these cases. These spectra are shown specifically because their higher order moments show interesting variations, suggesting that some portion of the high skewness and kurtosis regions are a result of poor modeling. Incorporating the ability to decompose multiple kinematic features of varying shapes is necessary to fully explore the information in this dataset.

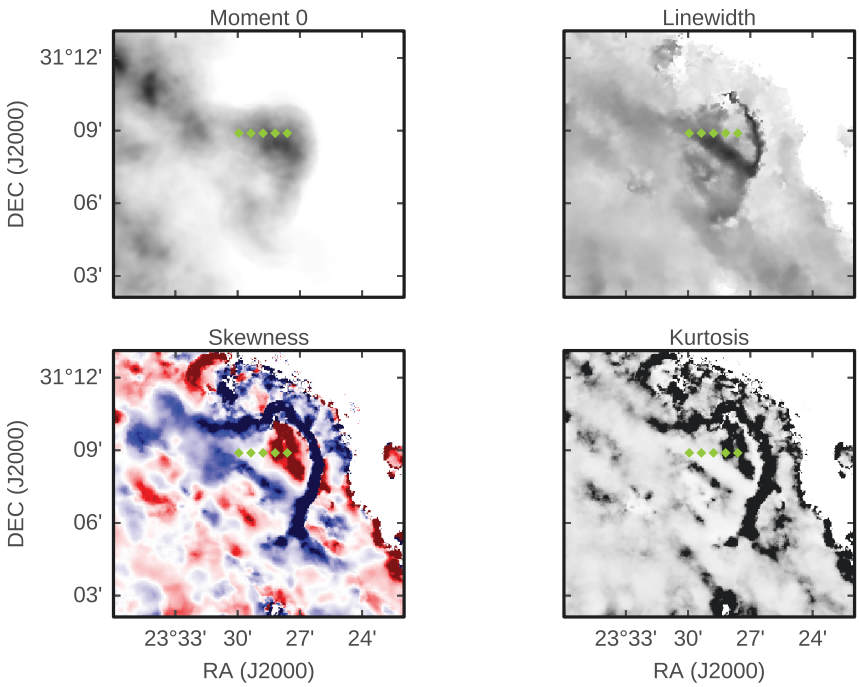


Figure 3.19: The zeroth moment, line width, skewness, and kurtosis maps around the Northern plume. The points in each of the four frames correspond the the positions of the spectra in Figure 3.20.

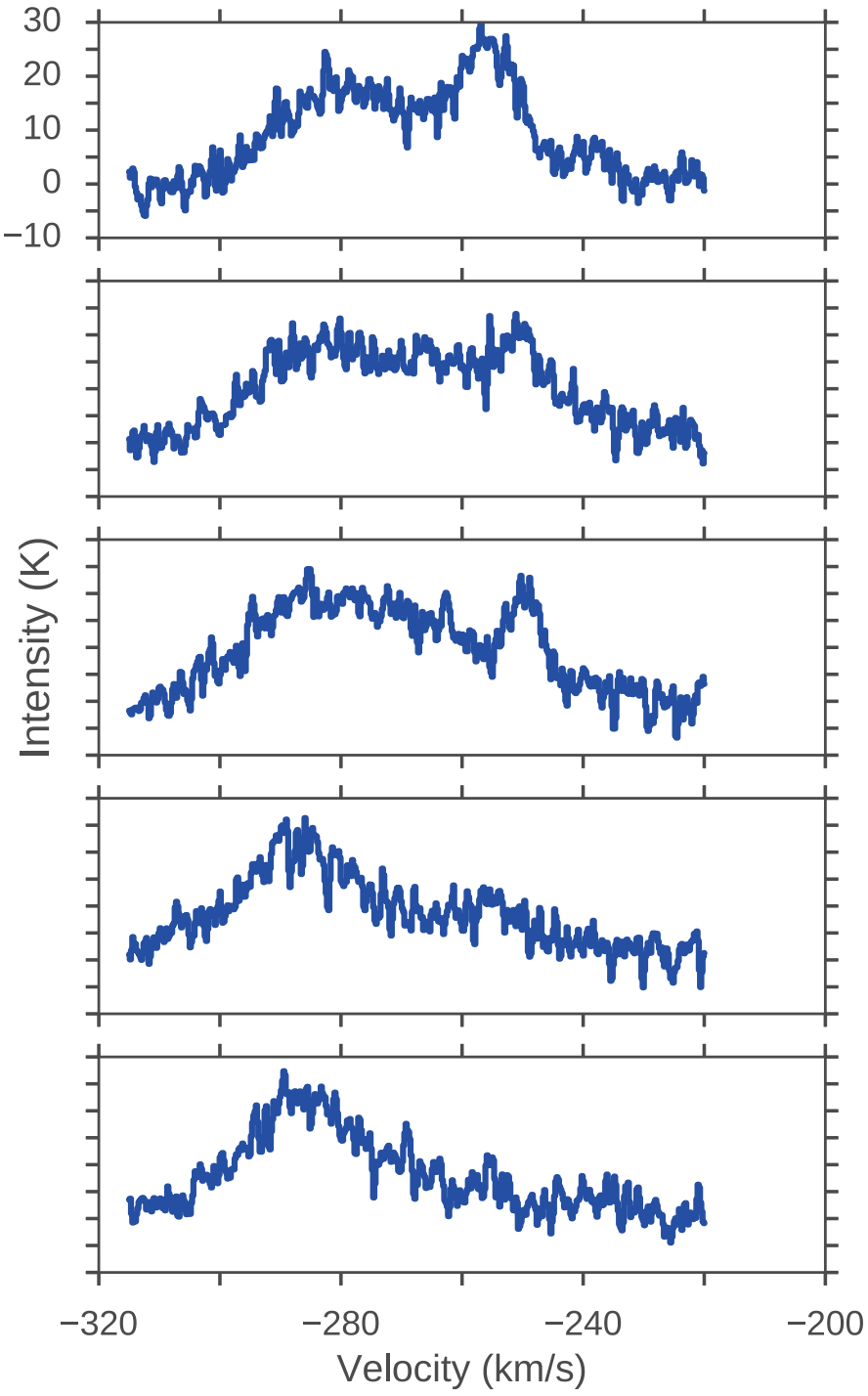


Figure 3.20: Spectra from the dots in Figure 3.19, proceeding from left to right. The y-axis is common across all of the spectra.

3.6 HI Morphology

The complex morphology of the H I emission makes simple descriptions difficult (§1.3). With a lack of spatial resolution, these descriptions are limited to averaged quantities, such as the clumping factor of Leroy et al. (2013a). Other options include by-eye analyses of bubbles and their associated shells (e.g., Bagetakos et al., 2011). The H I morphology in M33 has been described as having a filamentary structure (Deul and van der Hulst, 1987) that varies with the galactic radius as discussed below. Several studies have attempted to determine the relation of the H I morphology to the events that cause them. Apart from the large-scale structure, dominated by spiral arms and tidal disruption in some systems, the small-scale holes and other features likely occur from a combination of supernovae and winds from massive stars (or massive star clusters). These effects are described in more detail in §1.3 and §1.4.

Ignoring the exact nature of individual features, the adaptive thresholding technique can be used to produce a mask of localized bright emission. These features will primarily be comprised of the spiral arms and wind-driven shells surrounding holes, since they are locally bright. Adaptive thresholding calculates a threshold condition for each pixel within a map based on the values of the pixels in a patch around it. The threshold condition used defines a pixel as part of the mask if its value is above the mean of the patch after applying a Gaussian smoothing kernel to lessen the effect of noise. This thresholding alone is not sufficient to produce a usable mask. Regions of pure noise are detected as a large number of tiny fragmented regions. I account for this using a modified masking procedure from the one presented in the FilFinder algorithm (Koch and Rosolowsky, 2015). The FilFinder algorithm was developed to decompose the filamentary structure within nearby molecular clouds using two-dimensional integrated intensity maps. The algorithm was thoroughly tested on the publicly released data from the Herschel Gould Belt Survey. The

masking steps are shown as follows:

1. The adaptive thresholding is performed on the image with a patch size of ten times the beam width. This choice of patch size was found to detect small and large scale variations well based on visual inspection. Changing the patch size by a couple times the beam width in either direction does not change the mask significantly.

2. A global mask is created by rejecting all pixels below a certain brightness cut-off. A value of 3σ is found to work well, where σ is the standard deviation calculated by a sigma-clipping routine (Men'shchikov et al., 2012).

3. Pixels not within the global mask are rejected from the adaptive threshold mask.

4. The edges of the mask are smoothed using a median filter with a kernel size of half the beam width. This reduces the fragmentation of the skeletons of these regions, which are highly sensitive to small variations in the shape.

5. Regions whose peak brightness is below some threshold are removed. A threshold of 6σ is found to effectively remove spurious noise features.

Here I use FilFinder's ability to decompose skeleton structures to describe the width of large-scale features across M33. The results are shown in Figure 3.21. Nearly all of the bright structure is connected into one massive region within the mask, as can be seen in the skeleton structure in Figure 3.21. This is expected since the emission is widespread and H I has low clumping factors over a large range of scales (§3.4). There are large-scale differences in the skeleton structure within the inner, mid, and outer disk that can be connected with the descriptions of Deul and van der Hulst (1987). The inner disk skeleton is not

well connected to the rest of the structure and lacks a particularly complex morphology. This is likely due to a lack of large-scale structure in this region — when integrated over, the network of WDS at different velocities becomes "washed-out," appearing as an enhanced background instead of distinct bright structures. The mid-disk (2–4 kpc) is dominated by the spiral arms. Both are well traced over their entire extents by the skeleton. A clear shift in skeleton morphology is seen in the outer disk. The skeleton is very well connected with a number of loops and "weaved" regions. A few linear regions are scattered throughout this region; these are likely fragments of spiral arm features. This supports the claims of distinct regions within the disk made in §3.4 — the connectivity of higher-column density regions appears to evolve from the inner to outer disk. In Figure 3.22 I investigate if these changes in connectivity are reflected in the average spacing between the skeletons. A distance transform is performed on the skeleton mask with respect to the skeleton location (i.e., a skeleton pixel has a distance of 0). The distance values are then radially binned, as performed in §3.4. Several local dips and peaks are evident, ranging from distances of 120–220 pc. However, there are no large-scale variations between the inner, mid, and outer disk that reflect the change in connectivity described above.

Also using the distance transform from the skeleton, an intensity radial profile can be constructed. This allows for the average widths of the filamentary network to be characterized. The radial profile is shown in Figure 3.23. Fitting a Gaussian with a fixed mean and constant background to this profile gives a FWHM of 274.1 ± 8.8 pc for the filamentary structure of M33, where this value has been deconvolved with the beam size. I note that the radial profile is not corrected for the inclination of the disk, making this width maximally uncertain by a factor of ~ 2 for the portions of the profile along the semi-minor axis. Structures traced by the skeleton have an average peak brightness of 999.2 ± 4.8 K km s^{-1}, a factor about 2.5 times brighter than the low-brightness, large-scale

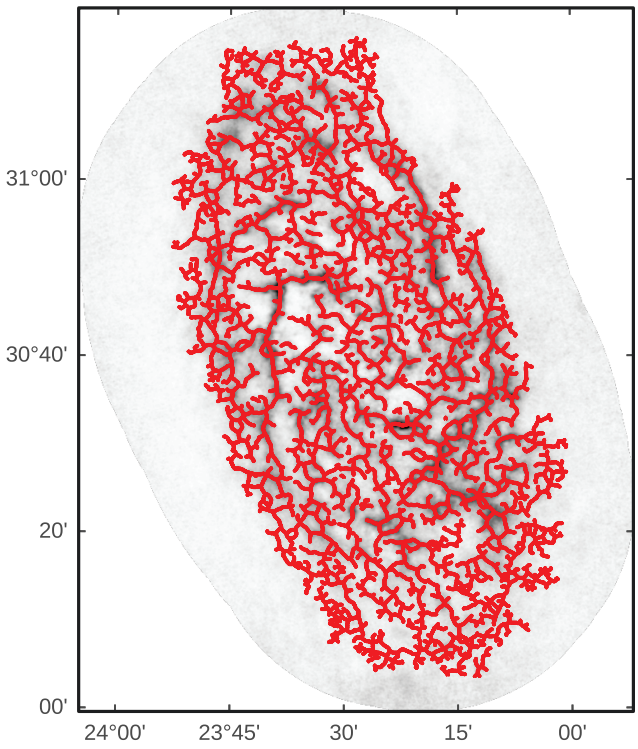


Figure 3.21: The skeleton network found by FilFinder. The skeleton structure traces the flocculent spiral arms in the outer disk and the main optical spiral arms in the mid-disk (2–4 kpc). The skeleton structure in the central region is noticeably less connected than those in the rest of the disk.

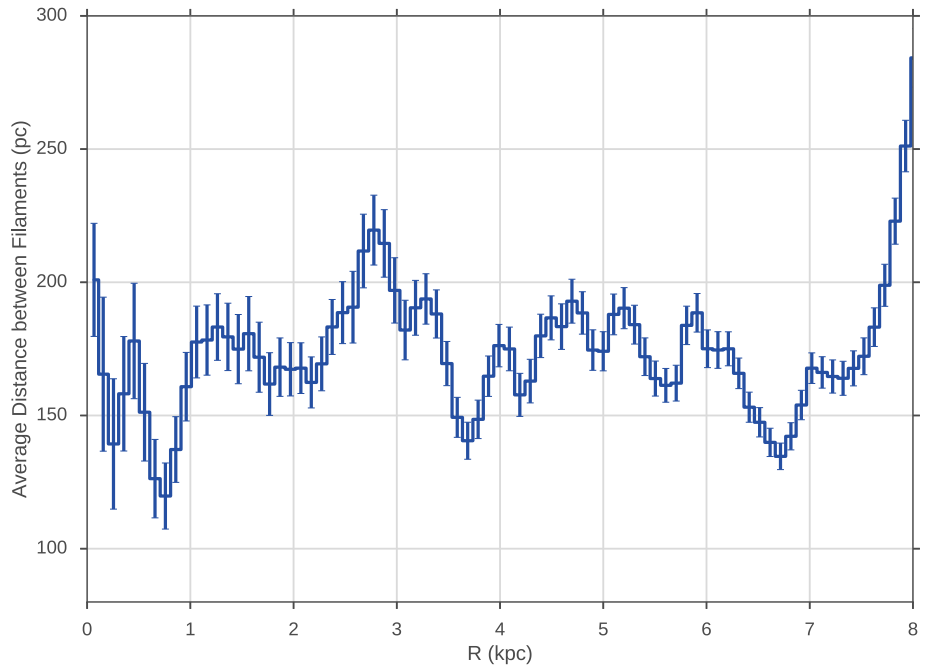


Figure 3.22: The radially averaged distances between filaments. There are no significant variations between the inner, mid, and outer disk regions. The skeletons end near 7 kpc, leading to the increase in distance near 8 kpc.

Table 3.4: Fit results of the radial profile in Figure 3.23. The model is Gaussian with a constant background with its mean fixed to zero. The errors are the 1σ values from the covariance matrix. The FWHM has been deconvolved with the beam size.

	Value	Error	Unit
Amplitude	999.2	4.8	K km s^{-1}
Background	389.1	16.9	K km s^{-1}
σ	121.0	3.6	pc
Decon. FWHM	274.1	8.8	pc

diffuse emission fit by the background. Taking the area above the background level and multiplying by the number of pixels in the skeleton, I find that the bright component account for 54% of the total intensity in the zeroth moment map. This is similar to the results found by Braun (1997), who developed an alternate spatial filtering technique to analyze these *high brightness networks* (HBNs) of filaments in 11 spiral galaxies. He finds that 60–90% of the total emission comes from HBNs within the optical radius R_{25}. This drops to around 50% when including the entire H I disk. The other half of the emission comes from a diffuse component in the inter-arm and outer disk regions, consistent with the large-scale background found here. By smoothing his data at different scales and examining how this diminishes the peak brightness, Braun argues that the HBNs have marginally resolved widths of about 150 pc. The direct width measurement here is nearly twice that value. This likely stems from the adaptive threshold mask being more inclusive than the masking performed by Braun (1997), who necessarily had to define restrictive masks due to the low signal-to-noise ratios in the data.

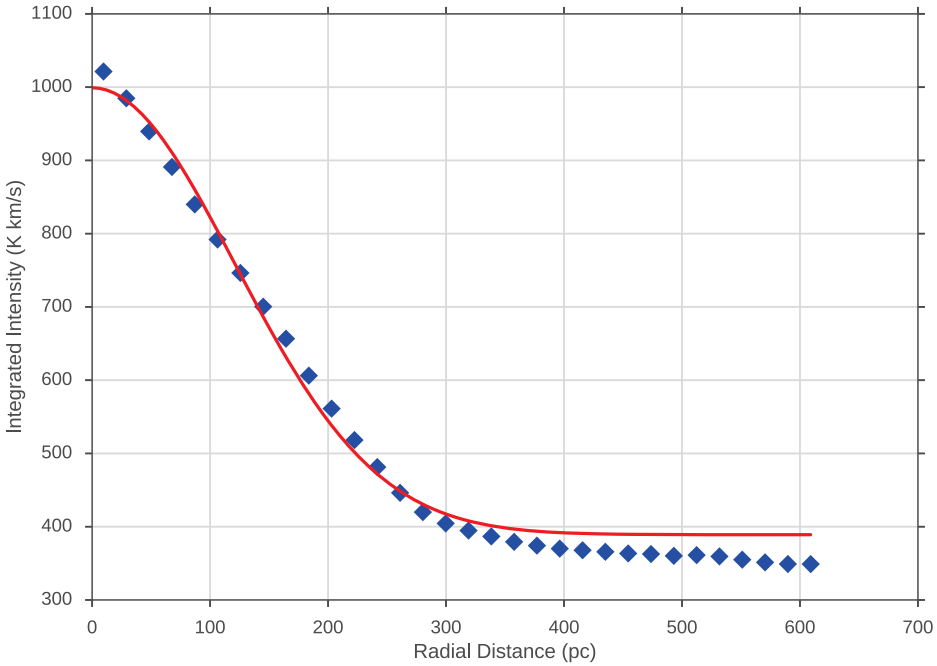


Figure 3.23: A Gaussian fit to the skeleton in Figure 3.21. Distances are produced by a distance transform of the skeleton network. The Gaussian model has a constant background and its mean fixed to zero.

Chapter 4

Linking the Atomic and Molecular ISM in M33

Star formation is inherently linked to the molecular medium, but is typically insensitive to the atomic medium (Leroy et al., 2008, §1.4). However the H_2-H I conversion requires sufficient shielding from FUV photons. This observational evidence implies that the bottleneck in the star formation process is the conversion of H_2 to H I. I investigate the link between these phases of the ISM by comparing the properties of the atomic medium (§3) to the molecular medium as traced by the CO(2-1) transition.

As discussed in §1.4, observing the conditions of GMCs is complicated by the lack of H_2 transitions at the low temperatures ($\sim$ 15 K) within molecular clouds. Using CO to trace H_2 introduces chemical dependencies when deriving properties of the molecular medium, since the ionization potential of CO is lower than H_2 and is not inhibited by the symmetry in H_2. Thus CO requires greater shielding from photodissociation. The outer regions of a molecular cloud will then lack CO, and therefore will not be detected. A conversion factor of $X_{CO} = 2 \times 10^{20}$ cm^{-2} / (K km s^{-1}) is typically assumed to convert the CO(1-0) intensity to H_2 column density. The conversion for the 2-1 rotation level uses X_{CO} multiplied by the ratio of the line emission between the two tran-

sitions. A thorough review of the assumptions implied by using this conversion is presented in Bolatto et al. (2013).

4.1 Previous Work

The proximity and moderate inclination has made M33 the target of several large-scale studies of the molecular medium. Single dish observations of the inner disk were taken by Wilson and Scoville (1989) tracing CO(1-0) with the NRAO 12-m telescope. The authors followed this study by observing 19 fields at high-resolution (7–8″ $\approx$ 30 pc) with the Owens Valley interferometer. They compare the properties of the GMCs in these fields against Milky Way clouds, and find a consistent size-line width relation. Wilson et al. (1997) extend this study with observations of seven GMCs in CO(2-1), CO(3-2), and ^{13}CO (2-1). Two of the clouds in their observations are related to the giant HII region NGC 604 (Wilson and Matthews, 1995) – one, NGC 604-2 is located within the HII regions, and the other lies just beyond the edge. They find differences in the physical properties of the clouds related to HII regions versus those that are not. This lead to the postulation that the changes arose from post-star formation effects in the environment, rather than the natal properties of the initial GMC.

Improved observations with the Berkeley Illinois Maryland Association (BIMA) interferometer were taken by Engargiola et al. (2003), allowing GMCs to be identified across the star-forming disk. Higher resolution observations with BIMA covering about a third of the GMCs (45 in total) were presented by Rosolowsky et al. (2003). They confirm the results of Wilson and Scoville (1990), finding consistent size-line width and mass-line width relationships with Milky Way clouds. They favour a cloud formation mechanism through the accumulation of atomic gas within < 500 pc, whereby significant magnetic braking is required to reach the angular momenta values calculated. Rosolowsky et al.

(2007) combines these data with single-dish observations by Heyer et al. (2004) with the Five College Radio Astronomy Observatory (FCRAO) 14-m telescope and the Nobeyama 45-m telescope. This provides coverage on all scales down to a resolution of 20″, lower than the 13″ afforded by the BIMA data, but with improved overall sensitivity. They present an updated catalog of 149 GMCs, complete to a mass limit of 1.3×10^5 $M_\odot$. Within 100 pc of GMCs, 90% of the flux is found to occur, suggesting that the location of low mass clouds is highly correlated with the position of GMCs. The mass distribution of clouds is found to vary radially within the disk, with all clouds above 8×10^5 $M_\odot$ located in the spiral arm dominated mid-disk (2-04 kpc). A possible asymmetry between the North and South portions of M33 is also suggested, again due to variation in the mass distribution of GMCs. I discuss this point in §4.4.

Bigiel et al. (2010) observed a small region of the outer disk of M33 at much higher resolution ($\sim 1.7'' \approx 7$ pc) with the Combined Array for Research in Millimeter-wave Astronomy (CARMA) interferometer. They identify 8 GMCs in this region, which are well into the atomic dominated region of M33's disk (§3.4). Compared to clouds at smaller radii, they find these clouds to be smaller signifying a steeper GMC mass distribution in the outer disk, or a incomplete sampling of spatial scales in the data. Alternatively, some of the H_2 may be "CO-dark" as the outer disk has a lower metallicity than the inner disk (Rosolowsky and Simon, 2008, §1.4).

4.2 CO(2-1) Data

I use new observations of the CO(2-1) line taken with the IRAM 30-m telescope to compare the properties of the atomic and molecular media in M33. A portion of this map was first presented by Gardan et al. (2007) and Gratier et al. (2010), and the completed observations are described in Druard et al. (2014). The data covers an area of $2400'' \times 3400''$, encompassing the disk completely to ~ 6.6

kpc. The spatial resolution is 12″ and the spectral resolution is 2.6 $\mathrm{km\,s^{-1}}$. The median noise level in the combined data is 20.33 mK, however severe striping and edge effects are apparent. The higher sensitivity and spatial resolution provide an improvement of the combined map from Rosolowsky et al. (2007), though the higher rotational level (2-1 versus 1-0) implies that this emission arises from slightly more excited gas.

A catalog of GMCs and their properties from the completed dataset has not yet been published. A preliminary catalog is presented in §4.5, though only for the purpose of confirming the co-location of H I and CO(2-1). A catalog of the Gratier et al. (2010) partial map is presented in Gratier et al. (2012). Using CPROPS (Rosolowsky and Leroy, 2006), 337 GMCs are found in the mapped region. They find the same radial variation in the GMC mass distribution identified in Rosolowsky et al. (2007), but with higher significance due to the increased sensitivity of the data. It should be noted that the preliminary catalog of the complete map from §4.5 has only 314 clouds. This suggests my requirements for a valid detection are more strict than Gratier et al. (2012).

Druard et al. (2014) finds a total molecular mass of 3.1×10^8 $\mathrm{M_\odot}$, assuming a conversion factor of $\mathrm{X_{CO}} = 4 \times 10^{20}$ $\mathrm{cm^{-2}}$ / (K $\mathrm{km\,s^{-1}}$). This conversion is twice the canonical value used for within the Milky Way. Rosolowsky et al. (2003) find that $\mathrm{X_{CO}} = 2 \times 10^{20}$ $\mathrm{cm^{-2}}$ / (K $\mathrm{km\,s^{-1}}$) across M33 independent of radius and metallicity. Bolatto et al. (2013) find that there is significant scatter in the data used for determining $\mathrm{X_{CO}}$ across nearby galaxies and recommend using $\mathrm{X_{CO}} = 2 \times 10^{20}$ $\mathrm{cm^{-2}}$ / (K $\mathrm{km\,s^{-1}}$). Using this $\mathrm{X_{CO}}$ value implies a total mass of 1.7×10^8 $\mathrm{M_\odot}$ for the molecular medium (traced by CO) with an uncertainty of 50% to reflect the scatter found by Bolatto et al. (2013).

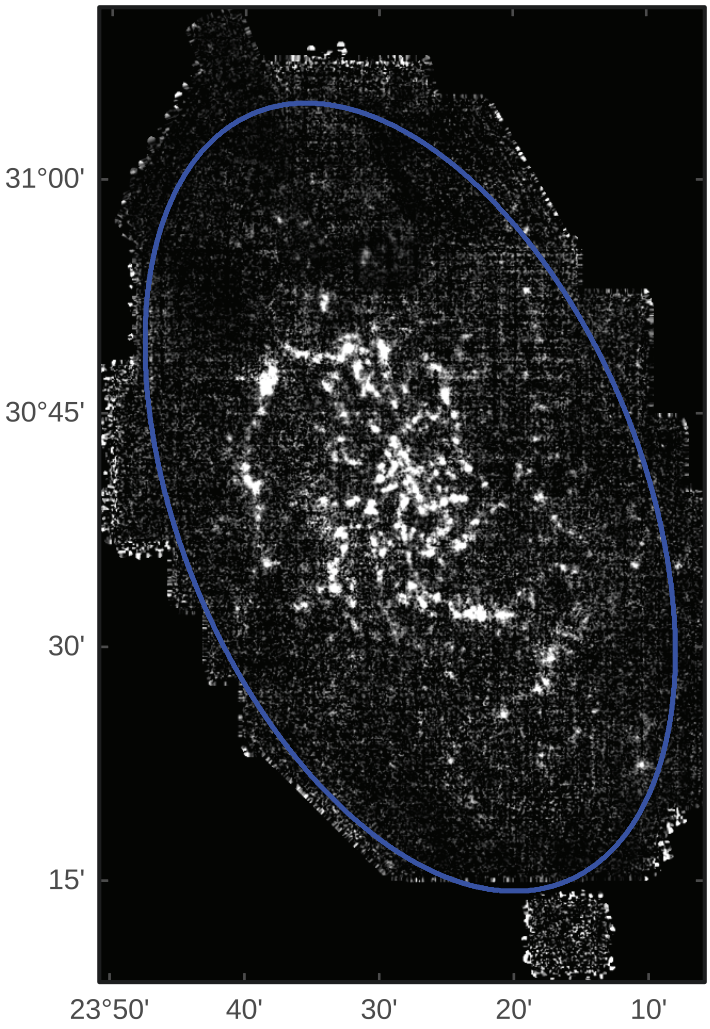


Figure 4.1: The zeroth moment map of the CO(2-1) observations. The blue contour indicates the radial cut applied to the cloud catalog in §4.5 to exclude edge effects. The emission is mostly localized to the inner disk and spiral arms. Striping effects are evident in most of the map.

4.3 Connections to the Atomic ISM

Similar analyses to those presented for H I in Chapter 3 may also be applied to the CO(2-1) data. Assuming the canonical X_{CO}, the CO(2-1) intensity can be converted to the surface density of H_2; this mass surface density factor used is 6.7 $M_\odot\,pc^{-2}/\,K\,km\,s^{-1}$. This conversion assumes a CO(2-1) to CO(1-0) line ratio of 0.65 (Leroy et al., 2013b).

Using the same radial profile technique presented in §3.4, I find an H_2 surface density profile that decreases rapidly with the galactic radius (Figure 4.2). When plotted with the H I profile from Figure 3.7, I find that the inner kpc is dominated by the molecular medium; the rest of the disk is increasingly dominated by the atomic component as the molecular portion drops quickly. The H_2 profile is consistent with the profile found by Druard et al. (2014) using the same data. A smaller bin width of 100 pc was chosen, however, to match the H I profiles shown in §3.4. This leads to a small truncation compared to the maximum radius of the profile from Druard et al. (2014) since the CO(2-1) emission is sparse beyond ~ 6 kpc. The large-scale features, namely the spiral arms, do not appear as clear peaks in the profile, though the decline is perhaps more steep past 4.2 kpc.

With the profiles of H I and H_2 calculated on the same scale, the total gas surface density across M33 is the sum of these two parts. This is plotted in Figure 4.3 with the stellar surface density profile found by Corbelli et al. (2014). Within 6 kpc of the center, M33 is dominated by the stellar component. The total gas surface density beyond the inner kpc is fairly constant, ranging between 6–8 $M_\odot\,pc^{-2}$. The peak surface density of the gas is ~ 20 $M_\odot\,pc^{-2}$ within the inner kpc where the molecular component dominates. Corbelli et al. (2014) finds that the atomic portion dominates the outer disk to around 7–8 kpc, beyond which its surface density drops steeply (Figure 3.7).

The total gas surface density and the relative H I to H_2 fractions can be

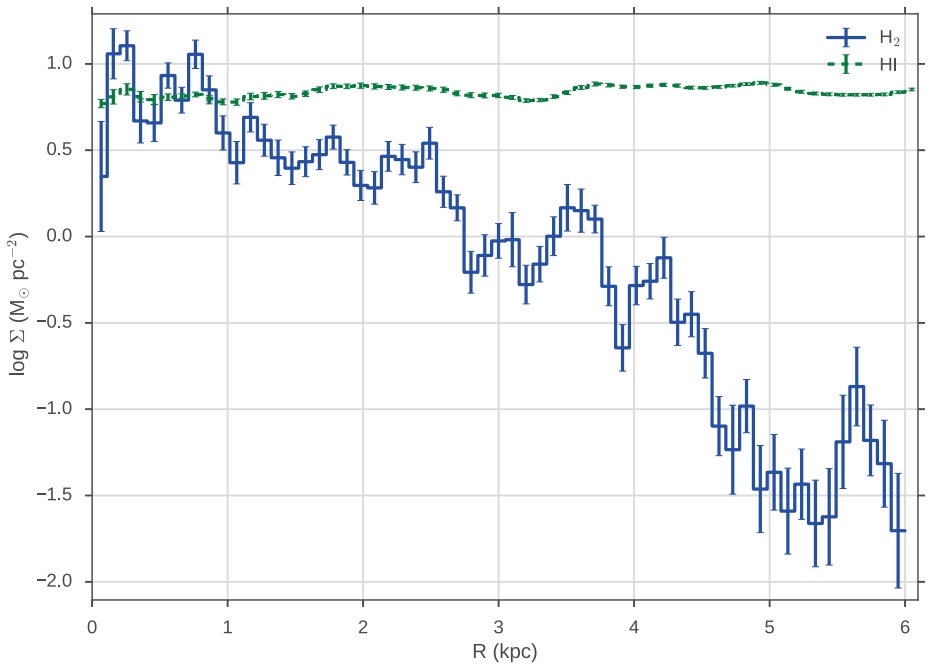


Figure 4.2: The radial profile of the surface density of H_2 (blue solid), with radial bins of 100 pc, is plotted with the H I profile (green dash) from Figure 3.7. Outside of the inner kpc, the neutral ISM is dominated by the atomic gas.

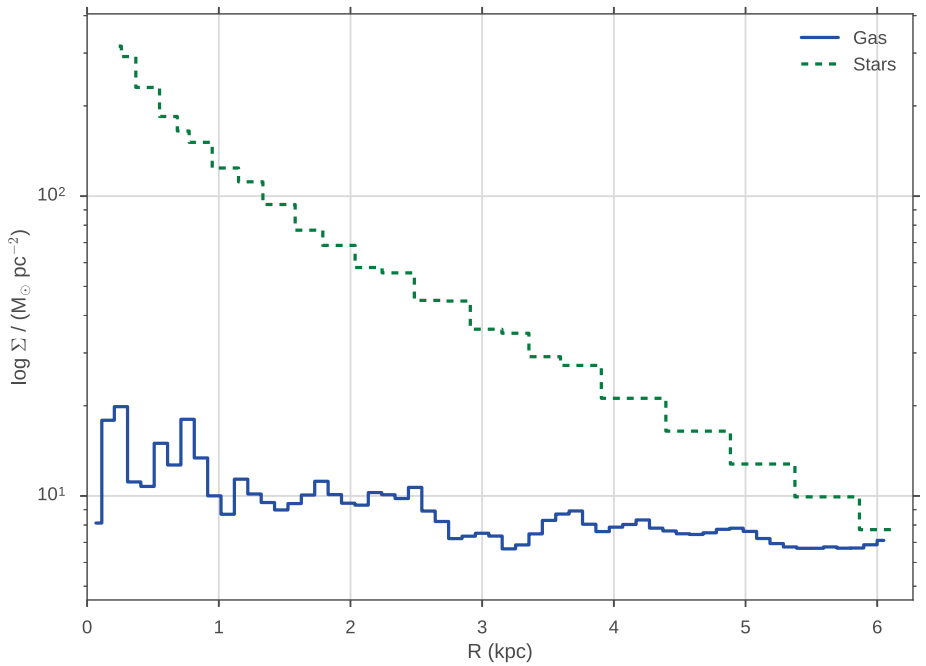


Figure 4.3: The radial profile of the total gas surface density (blue solid) is presented with the stellar surface density (green dash) found by Corbelli et al. (2014).

used to investigate the formation of H_2 and its dependence on the galactic environment. A model for the molecular fraction of the ISM was introduced by Krumholz et al. (2009b, see also Krumholz et al. (2008, 2009a)).

Using the total gas surface densities found for M33, I plot solutions to Equation 1.13, varying the clumping factor and metallicity, against the ratio of H_2 and H I in Figure 4.4, as recommended by Schruba et al. (2011). In terms of the molecular fraction, this ratio is $f_{H2}/(1-f_{H2})$ (see Equation 1.13). The data points show that the molecular component dominates when the overall surface density of the gas is greatest, consistent with the H I saturation reported in numerous studies (e.g., Bigiel et al., 2008; Leroy et al., 2013b). I find that the surface densities calculated in 100 pc radial bins require a clumping factor of 4–6 to match the highest surface density points. Given the assumptions in the model, the clumping factor should be near one. At lower surface densities, the data creates a steep tail that is not well-matched by any of the three models shown. Changing the average metallicity does not produce a better match to this tail either. Moreover, M33 has a radial abundance gradient (Rosolowsky and Simon, 2008) and the choice of a single metallicity is not necessarily optimal.

These results are somewhat surprising since Schruba et al. (2011) find the Krumholz et al. (2009b) model with a clumping factor between 1 and 2 match the molecular-to-atomic ratio from a sample of THINGS and HERACLES galaxies. The physical resolution in this sample is on kpc scales however, and a constant clumping factor may be a naive choice when these regions are resolved at 100 pc scales. Figure 13 in Schruba et al. (2011) shows that the total gas surface density versus galactic radius (normalized to the optical radius) shows a near constant decrease. M33's profile does not show this decrease; in fact the gas surface density is nearly constant until the H I surface density rapidly drops off near 7 kpc. This leads to the sharp drop-off seen in Figure 4.4; the atomic surface density is constant while the molecular component drops-off rapidly.

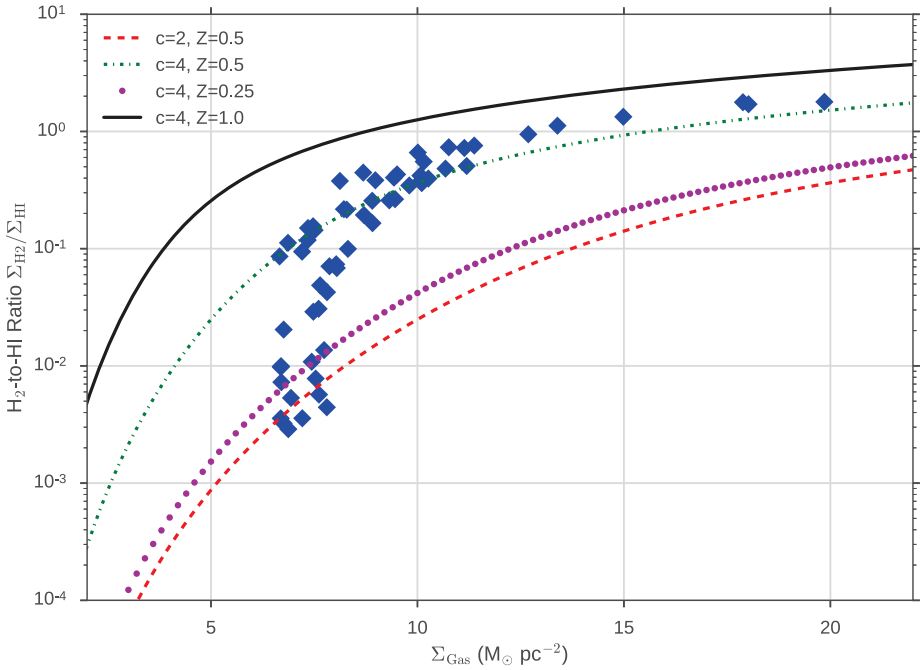


Figure 4.4: The ratio of the molecular gas to atomic gas surface densities is plotted against the total gas surface density (the sum of the atomic and molecular components). The lines indicate different choices of a clumping factor and metallicity in the Krumholz et al. (2009b) model for H_2 formation.

Furthermore, the points in Figure 4.4 are lower limits on the surface density. As suggested by the analytical model of Krumholz et al. (2011), there may be a significant amount of CO-dark H_2 in M33 — it is both H I dominated over most of the disk and has a lower metallicity, making the shielding necessary for CO formation weaker and giving the appearance of smaller GMCs (Bigiel et al., 2010, §1.4).

Global Line Properties — Using the H I rotation curve, I create a rotation-subtracted cube of the CO(2-1) emission as was done for the H I data (§3.2). H I emission typically provides a better measurement of the rotation within a galaxy since it is bright and widespread throughout the disk. Applying the model to the CO(2-1) data assumes that the molecular gas has similar line-of-sight velocities as the HI. However Corbelli et al. (2014) already investigated

Table 4.1: Fit results of the total rotation-subtracted line profile in Figure 4.5. The errors are the 1σ values from the covariance matrix. The additional error of 2.6 $\mathrm{km\,s^{-1}}$ in velocity comes from the finite channel width in the data.

	Value	Error	Unit
Mean	1.19	0.006 + 2.6	$\mathrm{km\,s^{-1}}$
Amplitude	0.970	0.001	Normalized Intensity
Standard Deviation	8.31	0.02 + 2.6	$\mathrm{km\,s^{-1}}$

this relation using the portion of this CO(2-1) data that was presented in Gratier et al. (2010). They find good agreement with the H I emission based on the peak CO velocities. The total spectrum across the regions with detected emission is shown in Figure 4.5. The spectrum is well-modeled by a single Gaussian with a width of 8.3 ± 2.6 $\mathrm{km\,s^{-1}}$, where I have assumed that the minimum error is the width of a spectral channel in the data due to rounding when subtracting the rotation curve. This width agrees with the 8.4 $\mathrm{km\,s^{-1}}$ found by Druard et al. (2014), who use a similar method for removing the galactic rotation.

There are clear differences in the total CO spectrum versus the total H I spectrum shown in Figure 3.6. The two spectra are plotted together in Figure 4.6. The CO spectrum clearly does not require a second wide Gaussian component. This is expected since: (1) CO(2-1) is more localized than H I emission and (2) the molecular ISM only has one cold component. The widths of the CO spectrum and the narrow component (9.1 $\mathrm{km\,s^{-1}}$) of the H I spectrum are comparable, and due to the finite channel widths, are consistent with each other. The CNM, as the probable envelope of H_2 and with a temperature of ~ 150 K, is expected to have a similar line width as the molecular component. Since the spectral resolution of the H I far exceeds that of the CO(2-1), the width of the CO(2-1) total spectrum may be overestimated due to the 2.6 $\mathrm{km\,s^{-1}}$ channels. However, the width obtained is consistent with the study by Mogotsi et al. (2016), who compare the H I and CO velocity dispersions of individual spectra in the galaxies observed by THINGS and HERACLES (§1.4).

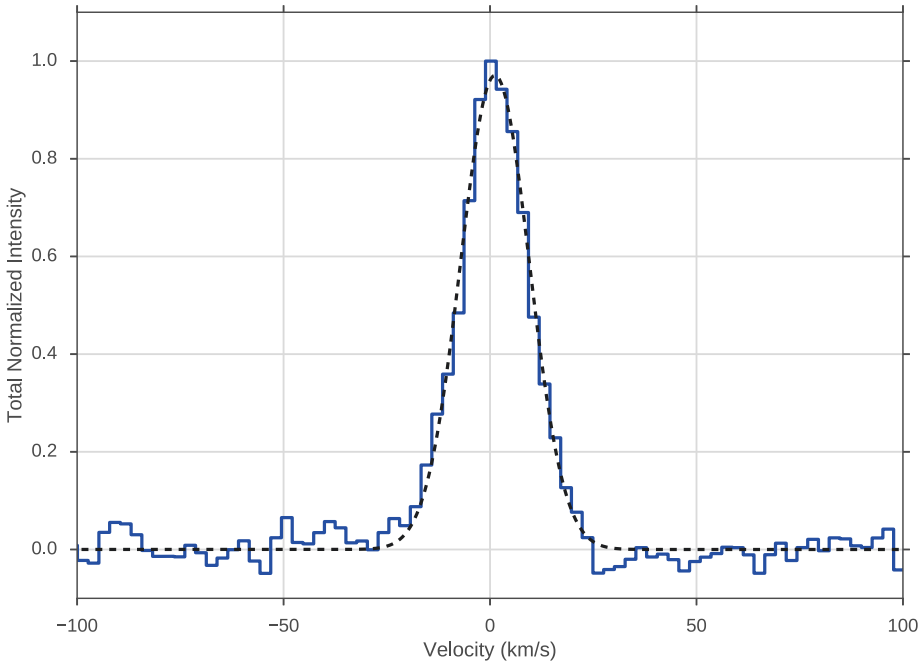


Figure 4.5: The total CO(2-1) spectrum is shown after subtracting the rotation curve derived from the H I data (§3.2).

4.4 North-South ISM Asymmetry in M33?

Is there an asymmetry of the ISM in the Northern and Southern halves of M33? This has been suggested by Rosolowsky et al. (2007) based on the asymmetrical GMC mass distributions found in each half. Furthermore, they find differences of 5%, 13%, and 5% in the total intensities in the halves for CO, HI, and Hα, respectively. I investigate this difference by creating separate CO(2-1) radial profiles for each half in Figure 4.7. A similar figure for H I is shown in Figure 3.9. Variations between the halves in the H I are apparent, however they favour the South half as having a higher average surface density. Since the typical error for each bin is ~ 0.5 M$_\odot$ pc^{-2}, most of the variation is not significant. It is interesting to note that the velocity dispersion profiles in Figure 3.13 show enhanced line widths in the South half from ∼2–5 kpc, the same radii where the H I surface density in the South is greater than in the North. In contrast, the

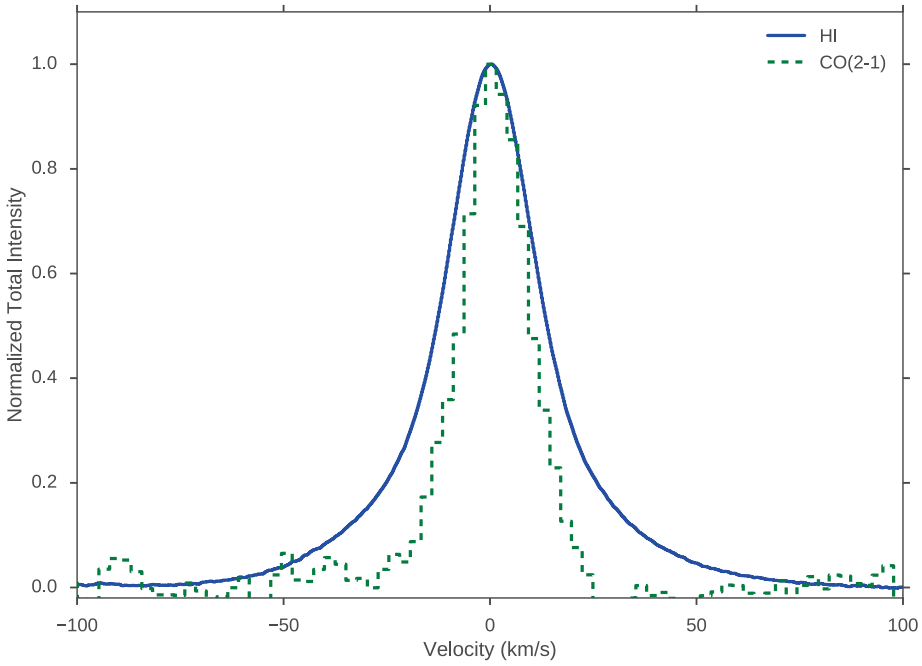


Figure 4.6: The total CO(2-1) and H I rotation corrected spectra are shown. The width of the CO(2-1) is slightly smaller than the narrow H I component (§3.3).

discrepancies in the H_2 profiles show several peaks at higher average surface densities in the North half. This suggests that the asymmetry in the GMC mass distribution found by Rosolowsky et al. (2007) is more related to the small asymmetries in the H_2 profiles, where most of this emission arises from GMCs. The H I profiles show no overall preference to higher surface densities in the Northern half. Putman et al. (2009) finds that the large-scale H I shape where the warp dominates the shape ($\sim$ 15 kpc) is nearly symmetrical. If there is a true asymmetry in the halves of M33, it does not seem to be reflected on the outer edges of the galaxy.

Based on these radial profiles, I find no substantial evidence for a link between the asymmetrical mass distributions found by Rosolowsky et al. (2007) and the global properties of the halves of the galaxy. While there may be an overall asymmetry based on the total intensities in a number of tracers, it does not appear to be directly related to differences on GMC scales. Instead, I suggest that the asymmetry in the mass distributions may be related to the strength of the Northern spiral arm and its effect on the mid-disk. In §3.4, I find that the dips in the velocity dispersion at 1 and 4.2 kpc are much more prominent in the northern profile. If these features are related to changes in the dominant mechanism governing the disk structure, it suggests that northern arm acts more strongly on the local environment in gathering and compressing the ISM. The locations of the most massive GMCs are within the northern arm, supporting this view. A more definitive result can be obtained by examining the dominant cloud formation mechanisms within the spiral arms (see §5.1).

4.5 Relation Between Molecular Clouds and H I Morphology

The previous sections in this chapter all largely consider the molecular and atomic media from a large-scale perspective. This is typical of extragalactic

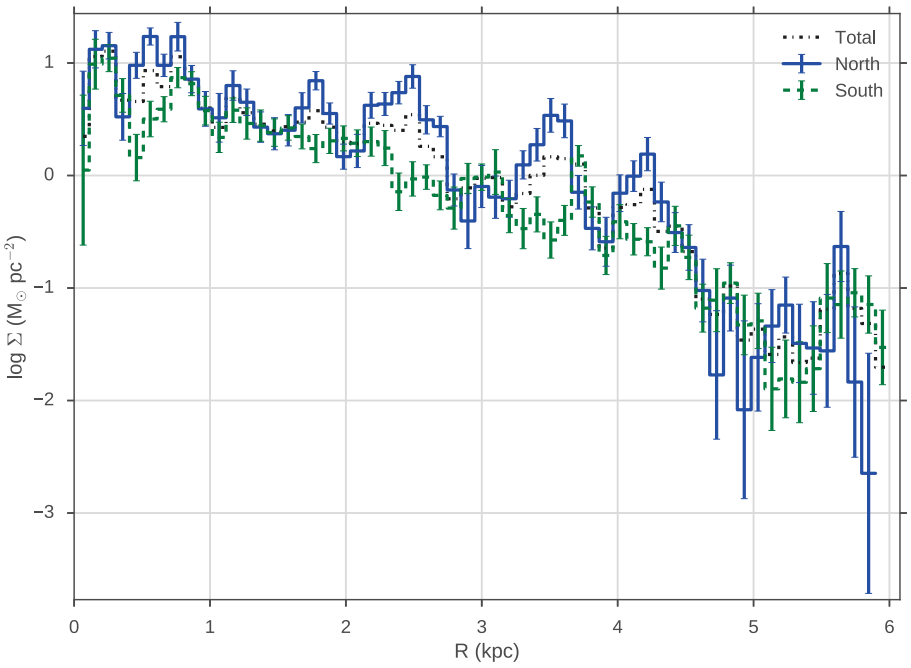


Figure 4.7: The H_2 radial profiles of surface density are shown for the North (blue solid) and South (green dash) halves separately. The total profile from Figure 4.2 is also shown (black dash-dot). The error bars are shown since the sparsity of CO emission beyond 4.5 kpc greatly increases the uncertainty in the bins.

observations: the kpc scales resolved in the data necessarily rely on the properties averaged over significant portions of a galaxy, or the entire galaxy itself. In this view, the Kennicutt-Schmidt law and its lack of sensitivity to the H I surface density imply that, once H_2 is present, star formation simply proceeds (§1.4). However, resolving < 100 pc scales across a galactic disk allows this to relation to be tested more explicitly. In this section, I investigate where the bulk of the CO(2-1) emission occurs relative to the H I filamentary network presented in §3.6.

A cursory check of the relative emission locations can be found by identifying the location of GMCs. If the bright CO, expected to be related to a GMC, is co-located with the bright HI, these regions should preferentially appear in areas such as the spiral arms and inner kpc, since the ISM is dominated by the molecular component in that region (Figure 4.2). A catalog of GMCs is constructed using the CPROPS algorithm (Rosolowsky and Leroy, 2006). The noisy edges of the mask lead to many spurious detections of clouds that are seemingly highly significant. To remove these I impose a radial cut-off of 6.6 kpc which, as shown in Figures 4.1 and 4.8, retains the regions of the map where real clouds can be located. I further apply a S/N threshold of 5, since the derived cloud properties become susceptible to noise below a S/N of ~ 4 (Colombo et al., 2014a). This gives a tentative final catalog of 314 GMCs sufficient to examine their locations relative to HI. The locations of the GMCs in Figure 4.8 show a clear correlation with bright H I in the central disk and the spiral arms, as expected. They do not, however, appear on every bright H I peak, suggesting that the properties of the local environment likely have an effect on the formation of molecular clouds. A catalog of GMCs using this new dataset have not been published, but will be presented in a future publication (Gratier et al., in prep). Furthermore, the remaining analyses in this section are too coarse to determine the H I-H_2 relation on the scales of individual clouds. Thus the GMC properties will not be explored within this thesis. A per object

Table 4.2: Fit results of the radial profile in Figure 4.9. The model is a Gaussian with a constant background and its mean fixed to zero. The errors are the 1σ values from the covariance matrix. The FWHM has been deconvolved with the beam size.

	Value	Error	Unit
Amplitude	0.281	0.005	K km s^{-1}
Background	0.018	0.005	K km s^{-1}
σ	83.5	3.0	pc
Decon. FWHM	190.7	7.3	pc

based analysis between GMCs and bubbles is discussed in §5.1.

Using the same H I filament network from §3.6, I create the equivalent radial profile using the CO(2-1) zeroth moment. The profile, as shown in Figure 4.9, is well fit by a Gaussian with mean zero and a constant background (Table 4.2). Unlike the H I profile, there is no evidence of a large-scale diffuse CO(2-1) background, which is consistent with viewing the zeroth moment in Figure 4.1. The average background intensity is slightly below the median noise level quoted by Druard et al. (2014). This can be further tested by integrating over the Gaussian profile and multiplying by the number of pixels in the skeleton. This Gaussian component actually accounts for more flux than is contained within the data, implying that the mean value is skewed high by a small number of very bright regions along the filamentary network. I find an average FWHM of 190.7 ± 7.3 pc, after deconvolving by the beam width. Thus the width of H I filaments are 1.4 times wider on average than those of the CO along the skeletons. The models of the radial profiles are shown in Figure 4.10. The clear CO peak at the center of the H I skeletons implies that most of the CO emission is spatially concurrent with the bright H I along the spine of the skeleton. However, many of the kinematic features, such as the shells around H I holes, are hidden after integrating over the spectral dimension unless they are particularly bright. Thus I expect this comparison provides a good assessment of the difference in the large-scale structure — spiral-arms — between the H I and CO, but not the small-scale shells and holes.

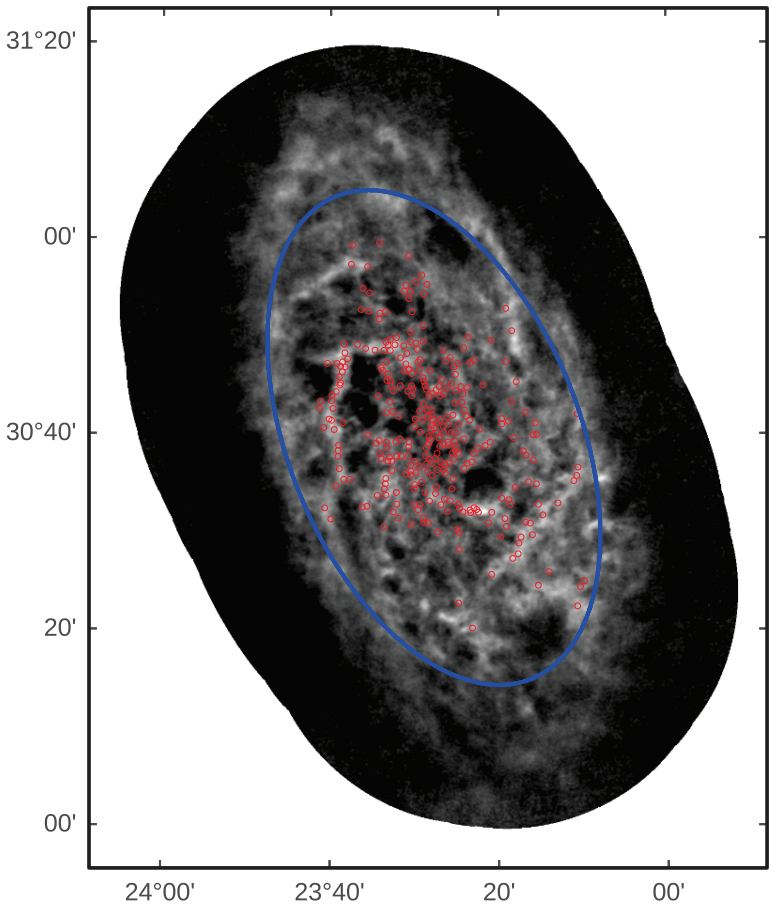


Figure 4.8: Locations of GMCs identified by CPROPS are shown in red circles on the zeroth moment map of HI. The blue contour is set at a radius of 6.6 kpc to remove edge effects in the CO(2-1) data.

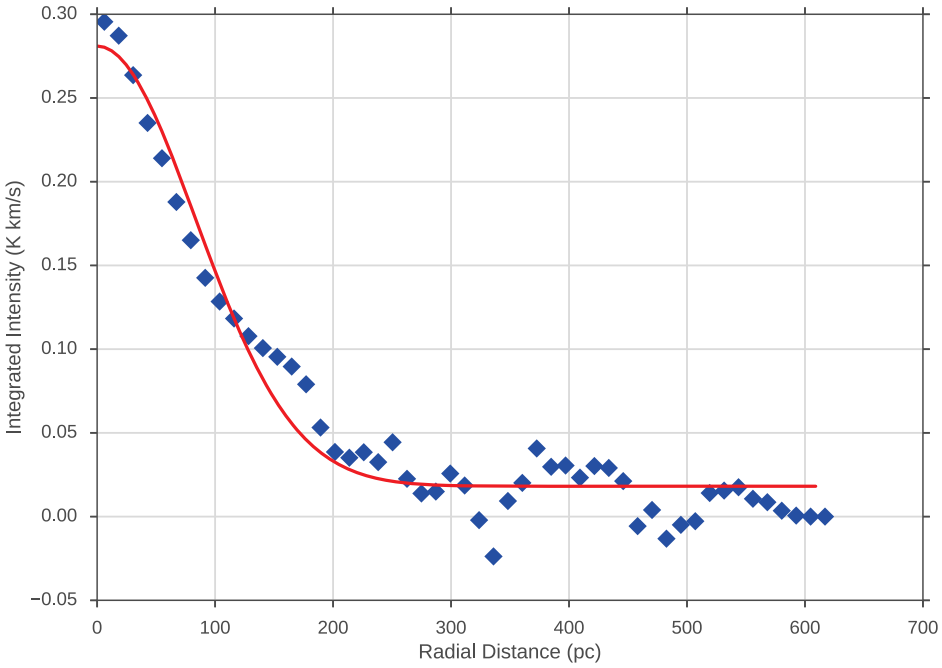


Figure 4.9: The radial profile of CO(2-1) emission defined by the H I skeleton from Figure 3.21.

To address this deficiency in the previous analysis, I can instead construct radial distributions of the H I and CO as a function of their distance from the edges of the adaptive thresholding mask. As described in §3.6, the edges of the mask trace the locations of rapid local changes in H I intensity well. Thus it determines the locations of H I bubbles, as confirmed by a visual inspection of where small holes in the mask occur (e.g., Figure 4.11).

I create common distributions of the CO and H I intensities with respect to their distance from the mask edge by spectrally integrating over 3 km s^{-1} chunks with common velocities in the rotation subtracted cubes. This accounts for the discrepancy in the spectral resolution of the datasets. The CO integrated chunk is then reprojected onto the same grid as the H I data. An adaptive threshold mask is created for the H I integrated chunk using the procedure in §3.6. To only compare regions contained in the CO map, a radial cut of 6.6 kpc is applied to the mask, and all H I emission outside is ignored. The mask is reduced to

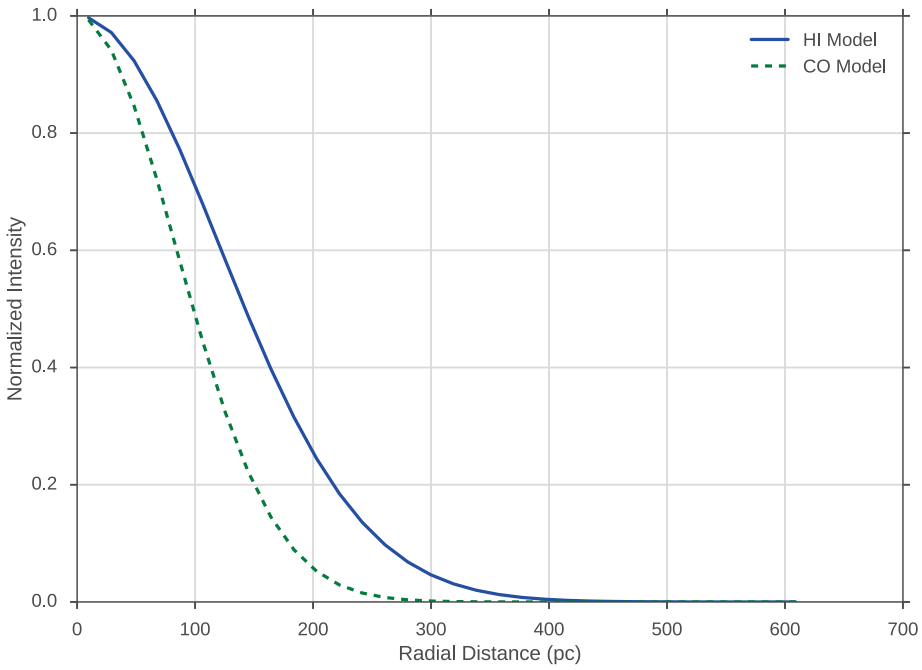


Figure 4.10: The Gaussian model profiles of the H I filamentary network are shown for H I and CO(2-1). Their amplitudes are normalized to unity and the background levels are set to 0 to highlight differences in their widths.

its edges, and a distance transform is applied. I note that the artificial edges from the radial cut-off are also ignored. Distances within the mask holes are assigned negative values, so the resulting distribution will map the variation across the mask edge smoothly. Each pixel's intensity in the CO and H I is matched with its respective distance to create the distribution. This process is repeated for 3 $\mathrm{km\,s^{-1}}$ chunks between -30 $\mathrm{km\,s^{-1}}$ to 30 $\mathrm{km\,s^{-1}}$, defined relative to the systematic velocity in the rotation-subtracted frame.

A portion of the integration from 0 to 3 $\mathrm{km\,s^{-1}}$ around the northern spiral arm is shown in Figure 4.11, overlaid with the edges of the adaptive threshold mask. Multiple partial and full H I bubbles are outlined well by the mask. The CO(2-1) clearly falls within the edge of the mask throughout the region, but does not directly correlate to all of the locations where H I is bright. In fact, the largest CO(2-1) complex in the vertical component of the arm occurs where there is noticably less H I emission than in other regions in this area. This complex is associated with the giant HII region NGC 604. Interestingly, the CO(2-1) follows the contour of the bubble shapes just north of this region, suggesting it is co-located with a possibly expanding shell (Tenorio-Tagle et al., 2000).

This lack of a one-to-one mapping between bright H I and CO(2-1) is also demonstrated in Figure 4.12, where the cumulative distribution functions (CDFs) of H I and CO(2-1) within all 3 $\mathrm{km\,s^{-1}}$ chunks are shown as a function of H I intensity. The CO(2-1) CDF is significantly offset, indicating it is not always co-located with bright H I. Alternatively, if CO(2-1) was only co-located with the brightest H I, the CO(2-1) CDF would rapidly rise above some threshold. Neither case is shown, and the relatively parallel curves between H I intensities of 100 to 200 $\mathrm{K\,km\,s^{-1}}$ suggests CO(2-1) is co-located with a wide range of H I intensities. This may indicate the role of the CNM, since HISA features, with expected widths of $< 1\mathrm{km\,s^{-1}}$ (Heiles and Troland, 2003), will noticeably lower the H I intensity when present (§1.3).

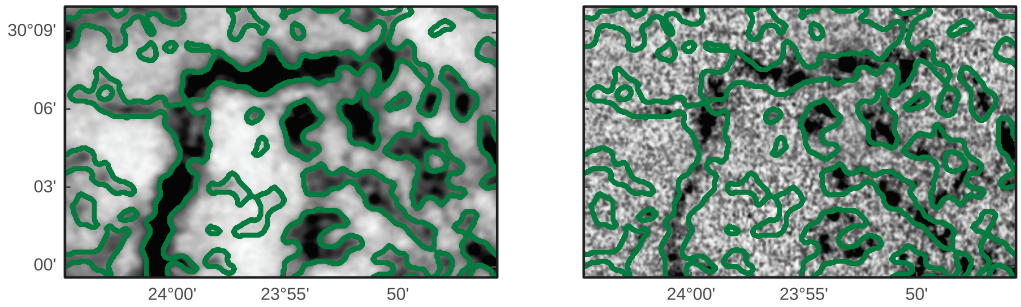


Figure 4.11: The integration from 0 to 3 $\mathrm{km\,s^{-1}}$ is shown for H I (left) and CO(2-1) (right) overlaid with the edges of the adaptive threshold mask (green) derived from the HI. The intensities have each had an arctan transform, normalized to the 85% percentile for the H I and 95% for the CO, to highlight fainter features. The mask highlights multiple full and partial bubbles, including one in the vertical component of the arm near the location of NGC 604. The CO(2-1) emission clearly falls within the mask edges, but not at all of the locations where there is bright HI.

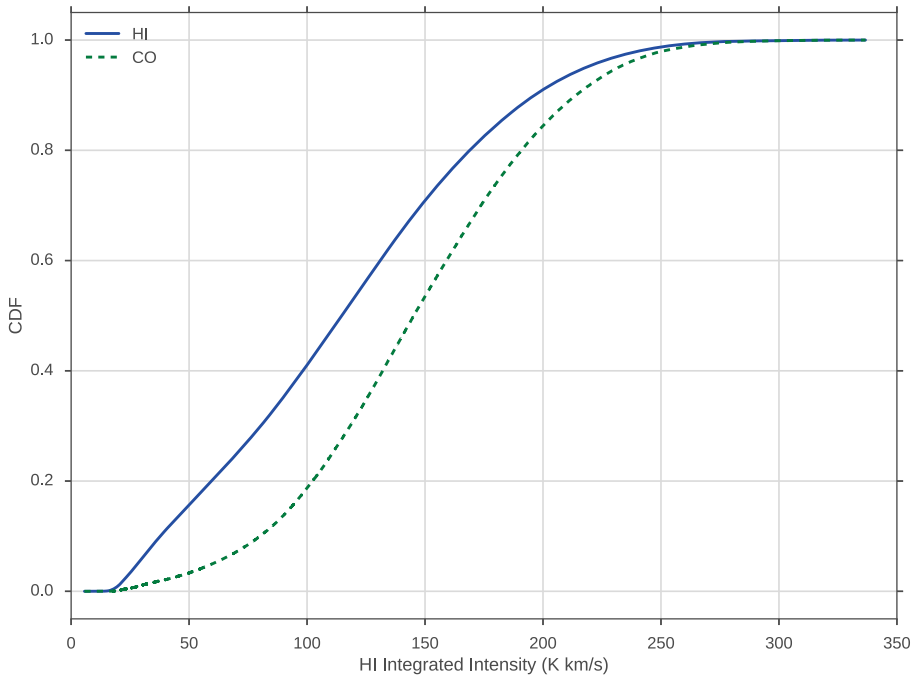


Figure 4.12: The CDFs of H I and CO(2-1) within the adaptive threshold masking over all 3 $\mathrm{km\,s^{-1}}$ chunks. Each curve is normalized by their respective tracer, but plotted against the H I intensity in that pixel. A one-to-one mapping between bright H I and CO(2-1) would show very similar CDFs.

The resulting CO and H I distance distributions are shown in Figure 4.13. There are several interesting features to note. First, the single radial peak of H I in the filamentary networks (Figure 3.23) has been spread out. This shows that mask has variations in width along the filamentary network. Second, neither H I nor CO(2-1) have significant emission within the holes of the mask. These holes can reach quite large sizes. When relating to H I bubbles and their shells, distances beyond $\sim$ 200 pc can typically be ignored, as only a small number of super-bubbles with larger widths are found in M33 (Deul and den Hartog, 1990). Third, there are two peaks in the CO radial distribution, and the greater one is not related to the maximum H I peak. This suggests that the CO emission is preferentially located either nearer to the bubble edges than H I, or tends to be located within the narrower regions of the H I filamentary structure.

Since the mask regions tend to be narrow, there are few pixels whose distance within the mask is larger than 200 pc. This makes importance of the H I peak and the second CO(2-1) peak uncertain. This is an issue since significant emission from at least the H I must occur at these distances. To estimate how important these correlations with distance are, I perform 100 bootstrap iterations, where the H I and CO(2-1) intensities are randomly permuted before being binned by distance. The standard deviation in each bin are shown as the errorbars in Figure 4.13. The H I peak is found to be significant. The second CO peak is consistent with being an increase, however the height and width of this peak are highly uncertain. The first CO peak is unequivocally a real feature.

This first CO peak is expected if molecular cloud formation occurs in wind-driven shells (WDS; §1.4). This is evidence that a significant portion of CO(2-1) emission, the majority of which is likely found in GMCs (Rosolowsky et al., 2007), is found on $\sim$ 100 pc scales from the mask edges. As argued above, the small-scale variations in the mask mostly corresponds to traces of bubble

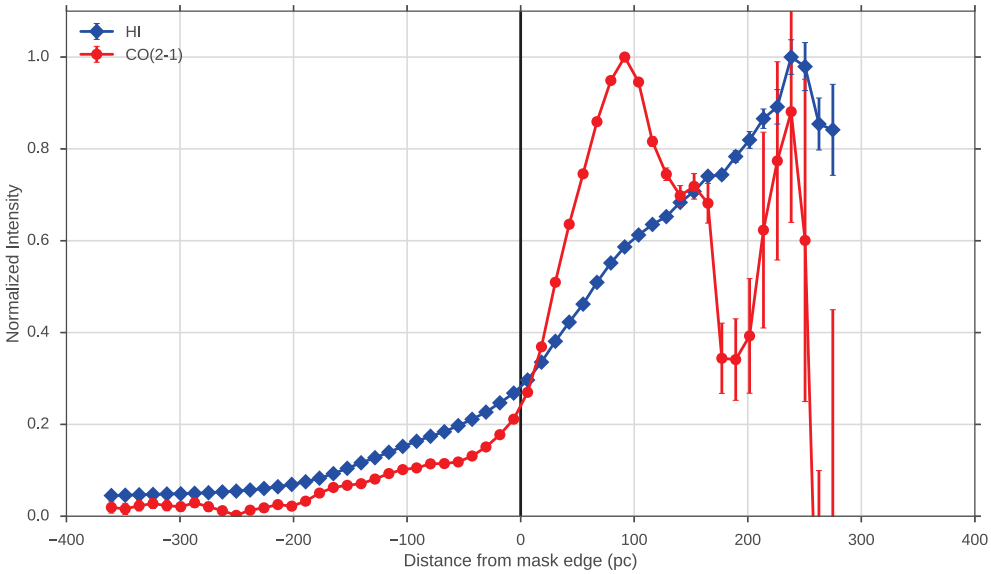


Figure 4.13: The normalized intensities of H I (blue diamond) and CO(2-1) (red circle) are shown as a function of distance from the adaptive threshold mask edges. Negative distances correspond to regions within holes in the HI, as defined by the mask, and positive distances are within the mask — the H I filaments. There are few pixels at the largest positive distances, since the typical width of a filament is only a couple beam widths (Figure 3.23). The errorbars are the standard deviations of 100 bootstrap iterations, where the intensity values have been randomly permuted.

edges. This the expected scale for local converging flows to dominate the cloud formation process (Dobbs et al., 2014) and indicates that WDS may play a crucial role in replenishing the molecular ISM after massive star formation events. Determining the cloud population contained in the first CO peak versus the second, and a direct association between individual GMCs and bubbles, will require a thorough population study and bubble catalog for M33 (§5.1).

Chapter 5

Conclusion

I have presented new observations of the nearby galaxy M33, focusing on analyzing the properties of the atomic medium using the 21-cm H I line and its relation to the molecular medium, traced by CO(2-1). The major findings of this thesis are:

- I present a solution for imaging massive spectral line interferometric datasets. Imaging is completed on a per channel basis, allowing the process to be manually parallelized. While the time required to split the data into individual channels is time consuming ($\sim$ 2 weeks in this case), the imaging and deconvolution only takes a couple hours to complete on each channel. When running on a cluster, the entire time required is several times less than imaging the data as a whole. The imaging of the H I data is completed using multiscale clean (Rich et al., 2008), with Arecibo observations as the starting model to include short-spacing data.

- I fit a rotation curve to the centroid surface using the DISKFIT algorithm. This provides updated values for the disk of M33, which are consistent with previous studies (Corbelli et al., 2014). I find an inclination of 54.6 ± 2.7 deg, position angle of 201.7 ± 0.5 deg, and a systematic velocity of -180.6 ± 0.8 km s^{-1} for M33.

- The rotation curve is subtracted from the H I and CO(2-1) data to examine the residual velocities. The total H I spectrum is well-modeled by two Gaussians with a fixed mean. Following Ianjamasimanana et al. (2012), I tentatively identify the narrow model component as emission from the CNM. The wide component has a width of 25.16 ± 0.21 km s^{-1}, 2.7 times larger than the narrow component. This is far wider than the expected thermal width of the WNM ($\sim$8–10 km s^{-1}), signifying that this component is tracing multiple effects. The components trace similar fractions of the total emission, with a slightly larger fraction (56%) found in the wide component. In contrast, the total CO(2-1) profile requires only one Gaussian component. Its width is found to be 8.3 ± 2.6 km s^{-1}, slightly smaller than the width of the narrow H I component (9.12 ± 0.20 km s^{-1}). This suggests that the narrow H I component and the CO(2-1) emission trace a similar medium, as has been suggested in other works (Pety et al., 2013; Mogotsi et al., 2016).

- An adaptive thresholding technique is used to identify the high-brightness network (HBN) of filamentary structure in the zeroth moment map of H I. Defining the centre of this mask through skeletonization, an average radial profile is created and modeled as a Gaussian, with a fixed mean of zero, and a constant background. The H I filaments have an average deconvolved FWHM of 304.0 ± 1.1 pc, twice the width of the unresolved features found by Braun (1997). The Gaussian component has an average peak of 999.2 ± 4.8 K km s^{-1} and accounts for 54% of the total H I emission. Based on the results of Braun (1997), the HBN is emission from the CNM, and the large-scale H I background results from the WNM. Using the H I filaments, a CO radial profile is found to have a deconvolved FWHM of 190.7 ± 7.3 pc. There is no large-scale background component of CO(2-1) above the noise limit.

- The intensity of H I and CO(2-1) are mapped as a function of distance from the edges of the adaptive threshold mask. Distributions are constructed by integrating over 3 $\mathrm{km\,s^{-1}}$ chunks ranging from -30 to 30 $\mathrm{km\,s^{-1}}$ in the rotation subtracted frame. The mask edges map the locations of the small-scale substructure, such as H I holes and shells. I find that CO is preferentially located closer to the edges than HI, with the most significant peak at $\sim$ 100 pc from the edges. This implies that: (1) CO is located nearer to bubble edges than HI, or (2) CO is preferentially formed in the narrower H I filaments. Molecular cloud formation from converging flows in shells predict a peak in CO emission on $\sim$100–150 pc scales. Thus my results suggest that molecular cloud formation in wind-driven shells may be a dominant process in M33.

- The radial profile of velocity dispersion across the disk is $\sim$ 10 $\mathrm{km\,s^{-1}}$, apart from two large dips at 1 and 4.2 kpc. I argue that these dips are signatures of changes in the large-scale structure of the disk, as they correspond to the radii where M33's optical spiral arms begin and end. These dips in the velocity dispersion are more significant in the Northern half, perhaps implying that the Northern spiral arm impacts the environment more than the Southern arm. Rosolowsky et al. (2007) finds the most massive clouds in M33 are found in the Northern arm and the GMC mass distribution differs between the halves.

5.1 Future Work

The continuation of this project will comprise my PhD. work. In recognition of this, I have attempted to point out where the analyses presented here can be extended and the reasons for doing so. To finish this thesis, I briefly elaborate on how the work presented relates to my future research:

- **Mechanisms of Molecular Cloud Formation** — The primary goal of this project is to understand which processes dominate the formation of molecular clouds. The analysis presented in §4.5, related to local converging flows, is the first step towards this goal. Identifying HISA features using similar techniques to Gibson et al. (2005) will unambiguously identify cold H I clouds, and their presence and morphology in the spiral arms will constrain the agglomeration of small clouds into GMCs. The polarized continuum data, when correctly calibrated and imaged (§2.3), traces synchrotron emission in the 1–2 GHz L-band. Using rotation synthesis, the magnetic field orientation and strength can be used to test mechanisms of magnetic instabilities. The observations presented in this thesis will provide important constraints on how molecular clouds form across a spiral galaxy.

- **Cataloging the Bubbles of M33** — For the past few months, I have worked on developing an algorithm for the automated detection of bubbles within H I cubes. I call this algorithm *BaSiCs* - Bubbles and Shells in Cubes. The new H I data presented here is massive. This makes cataloging bubbles and their shells by-eye, as was done by Bagetakos et al. (2011) with the THINGS galaxies, unfeasible. Providing a robust algorithmic approach will also minimize selection bias in identifying what classifies as a bubble. An example of the bubbles found by BaSiCs is shown in Figures 5.1 and 5.2 for the nearby dwarf galaxies IC 10 and IC 1613. The algorithm works on a per spectral channel basis, identifying regions of strong response to a Gaussian-Laplace transform and fitting elliptical shapes to the boundary based on the same adaptive thresholding technique described in §3.6. These regions are then clustered across the channels based on how correlated their shapes are in adjacent channels. While the algorithm tends to work well for these lower spectral resolution data (1.3 $\mathrm{km\,s^{-1}}$), the high spectral resolution in these new data tends to

artificially fragment bubbles along the line-of-sight. Development of this algorithm as part of my PhD. work will culminate in a catalog of bubbles in M33. This effectively provides a decomposition of the events that inject a substantial amount of energy into the ISM, allowing for the relation between molecular cloud formation and H I shells to be thoroughly studied. The analysis in §4.5 is an initial non-parametric description of this. The connection between WDS and GMCs found here will be extended to a comparison the of the bubble and GMC populations.

- **The Impact of Feedback around NGC 604** — I was awarded additional observational time with the VLA to study the morphology of H I around the giant HII region NGC 604 in the Northern spiral arm. These observations will yield a $\sim$ 4 pc view of H I across the region. Observations with Hubble show an intricate network of expanding HII shells (Martínez-Galarza et al., 2012); the morphology of the H I on a similar resolution to Hubble will provide critical information for understanding how vigorous star formation impacts the local ISM environment.

- **Cloud Formation Across a Spiral Arm** — We have two more successful observing proposals to study the impact of the shock front in the Southern spiral arm. The NOEMA interferometer will produce a 200 pc $\times$ 200 pc map across the arm at $2''$ resolution in CO(2-1). More VLA observations will be taken this fall of the H I at a common resolution to this new CO(2-1) map. At this resolution, we will be able to study the effect of the shock on sub-GMC scales. Understanding how the shock effects the ISM will provide the first observational test for our understanding of the role of spiral arms in the star and cloud formation process on these scales.

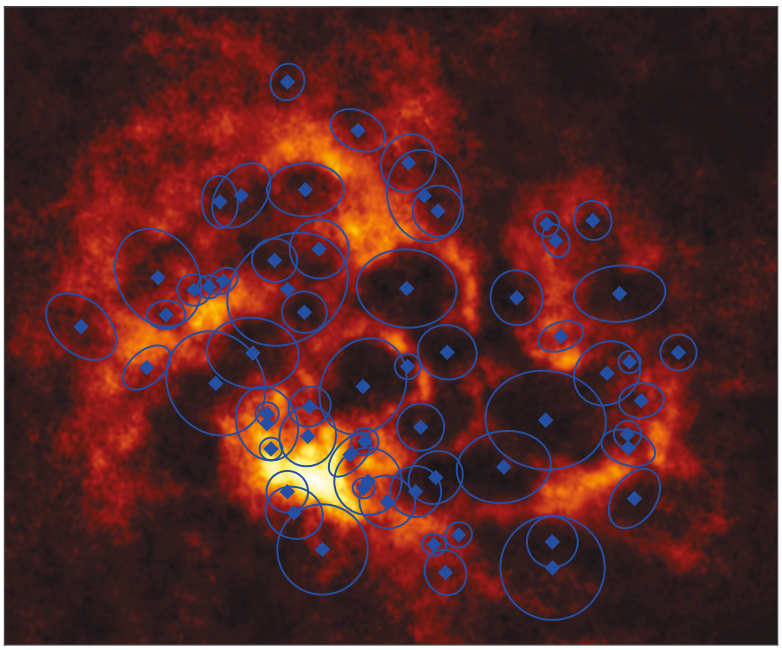


Figure 5.1: Preliminary results of the bubbles detected by BaSiCs in the dwarf galaxy IC10. Data from the LITTLE THINGS survey (Hunter et al., 2012)

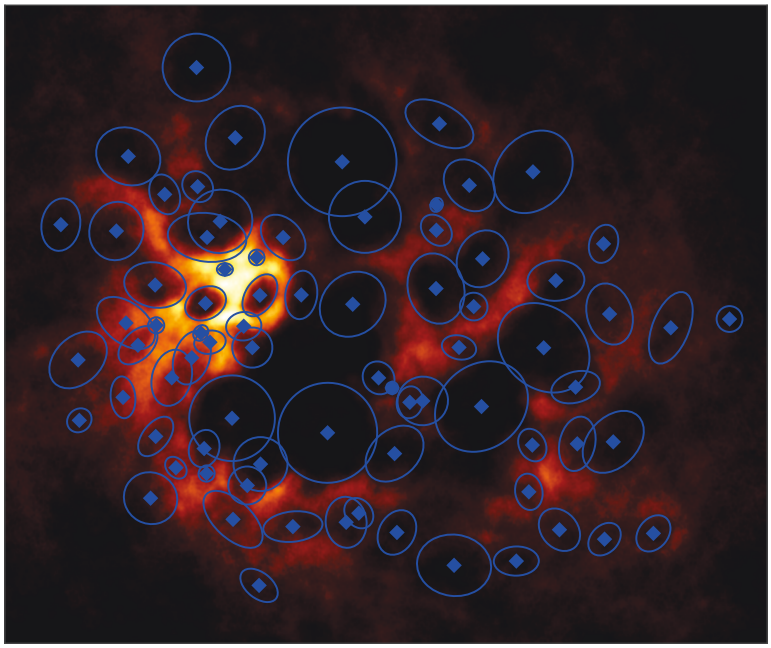


Figure 5.2: Preliminary results of the bubbles detected by BaSiCs in the dwarf galaxy IC1613. Data from the THINGS survey (Walter et al., 2008).

Appendix A

Code Bibliography

I used a number of open-source PYTHON packages to accomplish the research presented in this thesis. I have included this appendix to acknowledge and thank the developers.

- **astropy** (Astropy Collaboration et al., 2013)
- **scikit-image** (van der Walt et al., 2014)
- **spectral-cube** - `http://spectral-cube.readthedocs.io`
- **radio_beam** - `https://github.com/radio-astro-tools/radio_beam`
- **DISKFIT** (Spekkens and Sellwood, 2007; Sellwood and Sánchez, 2010)
- **FilFinder** (Koch and Rosolowsky, 2015)
- **FITS_tools** - `http://fits-tools.readthedocs.io`

The code I developed for this thesis is publicly available: `https://github.com/e-koch/VLA_Lband`. This includes scripts to reproduce the imaging technique described in §2.5. The archival VLA H I pipeline is also publicly available: `https://github.com/akleroy/hi_reduction_scripts`.

Bibliography

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. *Protostars and Planets VI*, pages 27–51, 2014.

Sean M Andrews, D J Wilner, A M Hughes, Chunhua Qi, and C P Dullemond. Protoplanetary Disk Structures in Ophiuchus. *The Astrophysical Journal*, 700(2):1502–1523, August 2009.

Astropy Collaboration, T. P. Robitaille, E. J. Tollerud, P. Greenfield, M. Droettboom, E. Bray, T. Aldcroft, M. Davis, A. Ginsburg, A. M. Price-Whelan, W. E. Kerzendorf, A. Conley, N. Crighton, K. Barbary, D. Muna, H. Ferguson, F. Grollier, M. M. Parikh, P. H. Nair, H. M. Unther, C. Deil, J. Woillez, S. Conseil, R. Kramer, J. E. H. Turner, L. Singer, R. Fox, B. A. Weaver, V. Zabalza, Z. I. Edwards, K. Azalee Bostroem, D. J. Burke, A. R. Casey, S. M. Crawford, N. Dencheva, J. Ely, T. Jenness, K. Labrie, P. L. Lim, F. Pierfederici, A. Pontzen, A. Ptak, B. Refsdal, M. Servillat, and O. Streicher. Astropy: A community Python package for astronomy. *Astronomy & Astrophysics*, 558:A33, October 2013. doi: 10.1051/0004-6361/201322068.

I Bagetakos, E Brinks, F Walter, W J G de Blok, A Usero, A K Leroy, J W Rich, and R C Jr Kennicutt. The Fine-scale Structure of the Neutral Interstellar Medium in Nearby Galaxies. *The Astronomical Journal*, 141(1):23, January 2011.

Steven A Balbus and John F Hawley. A powerful local shear instability in weakly magnetized disks. I - Linear analysis. II - Nonlinear evolution. *Astrophysical Journal*, 376:214–233, July 1991.

H Beuther, E B Churchwell, C F McKee, and J C Tan. The Formation of Massive Stars. *Protostars and Planets V*, pages 165–180, 2007.

F Bigiel, A Leroy, F Walter, E Brinks, W J G de Blok, B Madore, and M D Thornley. The Star Formation Law in Nearby Galaxies on Sub-Kpc Scales. *The Astronomical Journal*, 136(6):2846–2871, December 2008.

F Bigiel, A D Bolatto, A K Leroy, L Blitz, F Walter, E W Rosolowsky, L A Lopez, and R L Plambeck. Unusually Luminous Giant Molecular Clouds in the Outer Disk of M33. *The Astrophysical Journal*, 725(1):1159–1164, December 2010.

L Blitz, Y Fukui, A Kawamura, A Leroy, N Mizuno, and E Rosolowsky. Giant Molecular Clouds in Local Group Galaxies. *Protostars and Planets V*, pages 81–96, 2007.

Alberto D Bolatto, Mark Wolfire, and Adam K Leroy. The CO-to-H2 Conversion Factor. *Annual Review of Astronomy and Astrophysics*, 51:207–268, August 2013.

P W J L Brand and W J Zealey. Cloud structure in the galactic plane - A cosmic bubble bath. *Astronomy and Astrophysics*, 38:363–371, February 1975.

John C Brandt. On the Distribution Of mass in Galaxies. I. The Large-Scale Structure of Ordinary Spirals with Applications to M 31. *Astrophysical Journal*, 131:293–, March 1960.

R Braun, D A Thilker, R A M Walterbos, and E Corbelli. A Wide-Field High-Resolution H I Mosaic of Messier 31. I. Opaque Atomic Gas and Star

Formation Rate Density. *The Astrophysical Journal*, 695(2):937–953, April 2009.

Robert Braun. The Temperature and Opacity of Atomic Hydrogen in Spiral Galaxies. *The Astrophysical Journal*, 484(2):637–655, July 1997.

Robert Braun. Cosmological Evolution of Atomic Gas and Implications for 21 cm H I Absorption. *The Astrophysical Journal*, 749(1):87, April 2012.

Anahi Caldu-Primo, Andreas Schruba, Fabian Walter, Adam Leroy, Karin Sandstrom, W J G de Blok, R Ianjamasimanana, and K M Mogotsi. A High-dispersion Molecular Gas Component in Nearby Galaxies. *The Astronomical Journal*, 146(6):150, December 2013.

Dario Colombo, Annie Hughes, Eva Schinnerer, Sharon E Meidt, Adam K Leroy, Jérôme Pety, Clare L Dobbs, Santiago Garcia-Burillo, Gaelle Dumas, Todd A Thompson, Karl F Schuster, and Carsten Kramer. The PdBI Arcsecond Whirlpool Survey (PAWS): Environmental Dependence of Giant Molecular Cloud Properties in M51. *The Astrophysical Journal*, 784(1):3, March 2014a.

Dario Colombo, Sharon E Meidt, Eva Schinnerer, Santiago Garcia-Burillo, Annie Hughes, Jérôme Pety, Adam K Leroy, Clare L Dobbs, Gaelle Dumas, Todd A Thompson, Karl F Schuster, and Carsten Kramer. The PdBI Arcsecond Whirlpool Survey (PAWS): Multi-phase Cold Gas Kinematic of M51. *The Astrophysical Journal*, 784(1):4, March 2014b.

F Combes, M Boquien, C Kramer, E M Xilouris, F Bertoldi, J Braine, C Buchbender, D Calzetti, P Gratier, F Israel, B Koribalski, S Lord, G Quintana-Lacaci, M Relano, M Röllig, G Stacey, F S Tabatabaei, R P J Tilanus, F van der Tak, P van der Werf, and S Verley. Dust and gas power spectrum in M 33 (HERM33ES). *Astronomy and Astrophysics*, 539:A67, March 2012.

Edvige Corbelli, David Thilker, Stefano Zibetti, Carlo Giovanardi, and Paolo Salucci. Dynamical signatures of a ΛCDM-halo and the distribution of the baryons in M 33. *Astronomy and Astrophysics*, 572:A23, December 2014.

T. Cornwell, R. Braun, and D. S. Briggs. Deconvolution. In G. B. Taylor, C. L. Carilli, and R. A. Perley, editors, *Synthesis Imaging in Radio Astronomy II*, volume 180 of *Astronomical Society of the Pacific Conference Series*, page 151, 1999.

Donald P Cox and Barham W Smith. Large-Scale Effects of Supernova Remnants on the Galaxy: Generation and Maintenance of a Hot Network of Tunnels. *Astrophysical Journal*, 189:L105–, May 1974.

J R Dawson, N M McClure-Griffiths, John M Dickey, and Y Fukui. Molecular Clouds in Supershells: A Case Study of Three Objects in the Walls of GSH 287+04-17 and GSH 277+00+36. *The Astrophysical Journal*, 741(2): 85, November 2011.

J R Dawson, N M McClure-Griffiths, T Wong, John M Dickey, A Hughes, Y Fukui, and A Kawamura. Supergiant Shells and Molecular Cloud Formation in the Large Magellanic Cloud. *The Astrophysical Journal*, 763(1):56, January 2013.

W J G de Blok, F Walter, E Brinks, C Trachternach, S H Oh, and R C Jr Kennicutt. High-Resolution Rotation Curves and Galaxy Mass Models from THINGS. *The Astronomical Journal*, 136(6):2648–2719, December 2008.

E R Deul and R H den Hartog. Small-scale structures in the H I distribution of M33. *Astronomy and Astrophysics (ISSN 0004-6361)*, 229:362–377, March 1990.

E R Deul and J M van der Hulst. A survey of the neutral atomic hydrogen

in M33. *Astronomy and Astrophysics Supplement Series (ISSN 0365-0138)*, 67:509–539, February 1987.

E M Di Teodoro and F Fraternali. 3D BAROLO: a new 3D algorithm to derive rotation curves of galaxies. *Monthly Notices of the Royal Astronomical Society*, 451(3):3021–3033, August 2015.

C L Dobbs, M R Krumholz, J Ballesteros-Paredes, A D Bolatto, Y Fukui, M Heyer, M M M Low, E C Ostriker, and E Vazquez-Semadeni. Formation of Molecular Clouds and Global Conditions for Star Formation. *Protostars and Planets VI*, pages 3–26, 2014.

Clare L Dobbs. The interstellar medium and star formation on kpc size scales. *Monthly Notices of the Royal Astronomical Society*, 447(4):3390–3401, March 2015.

C Druard, J Braine, K F Schuster, N Schneider, P Gratier, S Bontemps, M Boquien, F Combes, E Corbelli, C Henkel, F Herpin, C Kramer, F van der Tak, and P van der Werf. The IRAM M 33 CO(2-1) survey. A complete census of molecular gas out to 7 kpc. *Astronomy and Astrophysics*, 567:A118, July 2014.

B G Elmegreen, J Palous, and S Ehlerova. Environmental dependences for star formation triggered by expanding shell collapse. *Monthly Notices of the Royal Astronomical Society*, 334(3):693–698, August 2002.

Bruce G Elmegreen and John Scalo. Interstellar Turbulence I: Observations and Processes. *Annual Review of Astronomy &Astrophysics*, 42(1):211–273, September 2004.

G Engargiola, R L Plambeck, E Rosolowsky, and L Blitz. Giant Molecular Clouds in M33. I. BIMA All-Disk Survey. *The Astrophysical Journal Supplement Series*, 149(2):343–363, December 2003.

E. B. Fomalont and R. A. Perley. Calibration and Editing. In G. B. Taylor, C. L. Carilli, and R. A. Perley, editors, *Synthesis Imaging in Radio Astronomy II*, volume 180 of *Astronomical Society of the Pacific Conference Series*, page 79, 1999.

Edward B. Fomalont and Melvyn C. H. Wright. *Galactic and Extra-Galactic Radio Astronomy*, chapter Interferometry and Aperture Synthesis, pages 256–290. Springer Berlin Heidelberg, Berlin, Heidelberg, 1974. ISBN 978-3-642-96178-6.

Wendy L Freedman, Barry F Madore, Brad K Gibson, Laura Ferrarese, Daniel D Kelson, Shoko Sakai, Jeremy R Mould, Robert C Jr Kennicutt, Holland C Ford, John A Graham, John P Huchra, Shaun M G Hughes, Garth D Illingworth, Lucas M Macri, and Peter B Stetson. Final Results from the Hubble Space Telescope Key Project to Measure the Hubble Constant. *The Astrophysical Journal*, 553(1):47–72, May 2001.

E Gardan, J Braine, K F Schuster, N Brouillet, and A Sievers. Particularly efficient star formation in M33. *Astronomy and Astrophysics*, 473(1):91–104, October 2007.

Steven J Gibson, A Russell Taylor, Lloyd A Higgs, Christopher M Brunt, and Peter E Dewdney. A Self-Absorption Census of Cold H I Clouds in the Canadian Galactic Plane Survey. *The Astrophysical Journal*, 626(1):195–213, June 2005.

A. Gil de Paz, S. Boissier, B. F. Madore, M. Seibert, Y. H. Joe, A. Boselli, T. K. Wyder, D. Thilker, L. Bianchi, S.-C. Rey, R. M. Rich, T. A. Barlow, T. Conrow, K. Forster, P. G. Friedman, D. C. Martin, P. Morrissey, S. G. Neff, D. Schiminovich, T. Small, J. Donas, T. M. Heckman, Y.-W. Lee, B. Milliard, A. S. Szalay, and S. Yi. The GALEX Ultraviolet Atlas of

Nearby Galaxies. *The Astrophysical Journal Supplement Series*, 173:185–255, December 2007. doi: 10.1086/516636.

P Gratier, J Braine, N J Rodriguez-Fernandez, K F Schuster, C Kramer, E M Xilouris, F S Tabatabaei, C Henkel, E Corbelli, F Israel, P P van der Werf, D Calzetti, S García-Burillo, A Sievers, F Combes, T Wiklind, N Brouillet, F Herpin, S Bontemps, S Aalto, B Koribalski, F van der Tak, M C Wiedner, M Röllig, and B Mookerjea. Molecular and atomic gas in the Local Group galaxy M 33. *Astronomy and Astrophysics*, 522:A3, November 2010.

P Gratier, J Braine, N J Rodriguez-Fernandez, K F Schuster, C Kramer, E Corbelli, F Combes, N Brouillet, P P van der Werf, and M Röllig. Giant molecular clouds in the Local Group galaxy M 33. *Astronomy and Astrophysics*, 542:A108, June 2012.

Lee Hartmann, Javier Ballesteros-Paredes, and Edwin A Bergin. Rapid Formation of Molecular Clouds and Stars in the Solar Neighborhood. *The Astrophysical Journal*, 562(2):852–868, December 2001.

C Heiles. H I shells, supershells, shell-like objects, and 'worms'. *Astrophysical Journal Supplement Series (ISSN 0067-0049)*, 55:585–595, August 1984.

C Heiles. The McKee/Ostriker Model: Paradigm? *Tetons 4: Galactic Structure*, 231:294–, 2001.

Carl Heiles and T H Troland. The Millennium Arecibo 21 Centimeter Absorption-Line Survey. II. Properties of the Warm and Cold Neutral Media. *The Astrophysical Journal*, 586(2):1067–1093, April 2003.

Mark H Heyer, Edvige Corbelli, Stephen E Schneider, and Judith S Young. The Molecular Gas Distribution and Schmidt Law in M33. *The Astrophysical Journal*, 602(2):723–729, February 2004.

A. Hill, R. Reynolds, L. Haffner, K. Wood, and G. Madsen. Modern view of the warm ionized medium. In *Highlights of Astronomy*, volume 10 of *Proceedings of the International Astronomical Union*, pages 574–575, 8 2012. doi: 10.1017/S1743921314012228. URL `http://journals.cambridge.org/article_S1743921314012228`.

D J Hollenbach and A G G M Tielens. Photodissociation regions in the interstellar medium of galaxies. *Reviews of Modern Physics*, 71(1):173–230, January 1999.

E M Hu. High latitude H I shells in the Galaxy. I. *Astrophysical Journal*, 248: 119–127, August 1981.

Deidre A Hunter, Dana Ficut-Vicas, Trisha Ashley, Elias Brinks, Phil Cigan, Bruce G Elmegreen, Volker Heesen, Kimberly A Herrmann, Megan Johnson, Se-Heon Oh, Michael P Rupen, Andreas Schruba, Caroline E Simpson, Fabian Walter, David J Westpfahl, Lisa M Young, and Hong-Xin Zhang. Little Things. *The Astronomical Journal*, 144(5):134, November 2012.

R Ianjamasimanana, W J G de Blok, Fabian Walter, and George H Heald. The Shapes of the H I Velocity Profiles of the THINGS Galaxies. *The Astronomical Journal*, 144(4):96, October 2012.

R Ianjamasimanana, W J G de Blok, Fabian Walter, George H Heald, Anahi Caldu-Primo, and Thomas H Jarrett. The Radial Variation of H I Velocity Dispersions in Dwarfs and Spirals. *The Astronomical Journal*, 150(2):47, August 2015.

R C Kennicutt, D Calzetti, G Aniano, P Appleton, L Armus, P Beirão, A D Bolatto, B Brandl, A Crocker, K Croxall, D A Dale, J Donovan Meyer, B T Draine, C W Engelbracht, M Galametz, K D Gordon, B Groves, C N Hao, G Helou, J Hinz, L K Hunt, B Johnson, J Koda, O Krause, A K Leroy, Y Li,

S Meidt, E Montiel, E J Murphy, N Rahman, H W Rix, H Roussel, K Sandstrom, M Sauvage, E Schinnerer, R Skibba, J D T Smith, S Srinivasan, L Vigroux, F Walter, C D Wilson, M Wolfire, and S Zibetti. KINGFISH—Key Insights on Nearby Galaxies: A Far-Infrared Survey with Herschel: Survey Description and Image Atlas. *Publications of the Astronomical Society of Pacific*, 123(9):1347–1369, December 2011.

Robert C Jr Kennicutt. The star formation law in galactic disks. *Astrophysical Journal*, 344:685–703, September 1989.

Robert C Jr Kennicutt. The Global Schmidt Law in Star-forming Galaxies. *The Astrophysical Journal*, 498(2):541–552, May 1998.

Robert C Jr Kennicutt, Lee Armus, George Bendo, Daniela Calzetti, Daniel A Dale, Bruce T Draine, Charles W Engelbracht, Karl D Gordon, Albert D Grauer, George Helou, David J Hollenbach, Thomas H Jarrett, Lisa J Kewley, Claus Leitherer, Aigen Li, Sangeeta Malhotra, Michael W Regan, George H Rieke, Marcia J Rieke, Hélène Roussel, John-David T Smith, Michele D Thornley, and Fabian Walter. SINGS: The SIRTF Nearby Galaxies Survey. *The Publications of the Astronomical Society of the Pacific*, 115 (8):928–952, August 2003.

Eric W Koch and Erik W Rosolowsky. Filament identification through mathematical morphology. *Monthly Notices of the Royal Astronomical Society*, 452 (4):3435–3450, October 2015.

Mark R Krumholz. The big problems in star formation: The star formation rate, stellar clustering, and the initial mass function. *Physics Reports*, 539 (2):49–134, June 2014.

Mark R Krumholz, Christopher F McKee, and Jason Tumlinson. The Atomic-to-Molecular Transition in Galaxies. I. An Analytic Approximation for Pho-

todissociation Fronts in Finite Clouds. *The Astrophysical Journal*, 689(2): 865–882, December 2008.

Mark R Krumholz, Christopher F McKee, and Jason Tumlinson. The Atomic-to-Molecular Transition in Galaxies. II: H I and H2 Column Densities. *The Astrophysical Journal*, 693(1):216–235, March 2009a.

Mark R Krumholz, Christopher F McKee, and Jason Tumlinson. The Star Formation Law in Atomic and Molecular Gas. *The Astrophysical Journal*, 699(1):850–856, July 2009b.

Mark R Krumholz, Adam K Leroy, and Christopher F McKee. Which Phase of the Interstellar Medium Correlates with the Star Formation Rate? *The Astrophysical Journal*, 731(1):25, April 2011.

Yasutaka Kurono, Koh-Ichiro Morita, and Takeshi Kamazaki. A Study of Combining Technique of Single-Dish and Interferometer Data: Imaging Simulations and Analysis. *Publications of the Astronomical Society of Japan*, 61: 873–894, August 2009.

Min-Young Lee, Snežana Stanimirović, Kevin A Douglas, Lewis B G Knee, James Di Francesco, Steven J Gibson, Ayesha Begum, Jana Grcevich, Carl Heiles, Eric J Korpela, Adam K Leroy, J E G Peek, Nickolas M Pingel, Mary E Putman, and Destry Saul. A HIGH-RESOLUTION STUDY OF THE H I-H2 TRANSITION ACROSS THE PERSEUS MOLECULAR CLOUD. *The Astrophysical Journal*, 748(2):75, April 2012.

A Leroy, A Bolatto, F Walter, and L Blitz. Molecular Gas in the Low-Metallicity, Star-forming Dwarf IC 10. *The Astrophysical Journal*, 643(2): 825–843, June 2006.

Adam K Leroy, Fabian Walter, Elias Brinks, Frank Bigiel, W J G de Blok, Barry Madore, and M D Thornley. The Star Formation Efficiency in Nearby

Galaxies: Measuring Where Gas Forms Stars Effectively. *The Astronomical Journal*, 136(6):2782–2845, December 2008.

Adam K Leroy, Fabian Walter, Frank Bigiel, Antonio Usero, Axel Weiss, Elias Brinks, W J G de Blok, Robert C Kennicutt, Karl-Friedrich Schuster, Carsten Kramer, H W Wiesemeyer, and Hélène Roussel. Heracles: The HERA CO Line Extragalactic Survey. *The Astronomical Journal*, 137(6): 4670–4696, June 2009.

Adam K Leroy, Cheoljong Lee, Andreas Schruba, Alberto Bolatto, Annie Hughes, Jérôme Pety, Karin Sandstrom, Eva Schinnerer, and Fabian Walter. Clumping and the Interpretation of kpc-scale Maps of the Interstellar Medium: Smooth H I and Clumpy, Variable H2 Surface Density. 769(1):L12, May 2013a.

Adam K Leroy, Fabian Walter, Karin Sandstrom, Andreas Schruba, Juan-Carlos Munoz-Mateos, Frank Bigiel, Alberto Bolatto, Elias Brinks, W J G de Blok, Sharon Meidt, Hans-Walter Rix, Erik Rosolowsky, Eva Schinnerer, Karl-Friedrich Schuster, and Antonio Usero. Molecular Gas and Star Formation in nearby Disk Galaxies. *The Astronomical Journal*, 146(2):19, August 2013b.

Alexia R Lewis, Andrew E Dolphin, Julianne J Dalcanton, Daniel R Weisz, Benjamin F Williams, Eric F Bell, Anil C Seth, Jacob E Simones, Evan D Skillman, Yumi Choi, Morgan Fouesneau, Puragra Guhathakurta, Lent C Johnson, Jason S Kalirai, Adam K Leroy, Antonela Monachesi, Hans-Walter Rix, and Andreas Schruba. The Panchromatic Hubble Andromeda Treasury. XI. The Spatially Resolved Recent Star Formation History of M31. *The Astrophysical Journal*, 805(2):183, June 2015.

Mordecai-Mark Mac Low. The Energy Dissipation Rate of Supersonic, Magne-

tohydrodynamic Turbulence in Molecular Clouds. *The Astrophysical Journal*, 524(1):169–178, October 1999.

J R Martínez-Galarza, D Hunter, B Groves, and B Brandl. Ongoing Massive Star Formation in NGC 604. *The Astrophysical Journal*, 761(1):3, December 2012.

C F McKee and J P Ostriker. A theory of the interstellar medium - Three components regulated by supernova explosions in an inhomogeneous substrate. *Astrophysical Journal*, 218:148–169, November 1977.

Sharon E Meidt, Richard J Rand, Michael R Merrifield, Rahul Shetty, and Stuart N Vogel. Radial Dependence of the Pattern Speed of M51. *The Astrophysical Journal*, 688(1):224–236, November 2008.

A Men'shchikov, Ph André, P Didelon, F Motte, M Hennemann, and N Schneider. A multi-scale, multi-wavelength source extraction method: getsources. *Astronomy and Astrophysics*, 542:A81, June 2012.

K M Mogotsi, W J G de Blok, A Caldu-Primo, F Walter, R Ianjamasimanana, and A K Leroy. Hi and CO Velocity Dispersions in Nearby Galaxies. *The Astronomical Journal*, 151(1):15, January 2016.

Kazuyuki Muraoka, Tomoka Tosaki, Rie Miura, Sachiko Onodera, Nario Kuno, Kouichiro Nakanishi, Hiroyuki Kaneko, and Shinya Komugi. 13CO(J = 1-0) On-the-Fly Mapping of the Giant H II Region NGC 604: Variation in Molecular Gas Density and Temperature due to Sequential Star Formation. *Publications of the Astronomical Society of Japan*, 64(1):3, February 2012.

J Palous, J Franco, and G Tenorio-Tagle. The evolution of superstructures expanding in differentially rotating disks. *Astronomy and Astrophysics (ISSN 0004-6361)*, 227:175–182, January 1990.

Narendra Nath Patra, Jayaram N Chengalur, Igor D Karachentsev, Serafim S Kaisin, and Ayesha Begum. Cold H I in faint dwarf galaxies. *Monthly Notices of the Royal Astronomical Society*, 456(3):2467–2485, March 2016.

J E G Peek, Carl Heiles, Kevin A Douglas, Min-Young Lee, Jana Grcevich, Snežana Stanimirović, M E Putman, Eric J Korpela, Steven J Gibson, Ayesha Begum, Destry Saul, Timothy Robishaw, and Marko Krčo. The GALFA-HI Survey: Data Release 1. *The Astrophysical Journal Supplement*, 194(2):20, June 2011.

Jérôme Pety, Eva Schinnerer, Adam K Leroy, Annie Hughes, Sharon E Meidt, Dario Colombo, Gaelle Dumas, Santiago Garcia-Burillo, Karl F Schuster, Carsten Kramer, Clare L Dobbs, and Todd A Thompson. The Plateau de Bure + 30 m Arcsecond Whirlpool Survey Reveals a Thick Disk of Diffuse Molecular Gas in the M51 Galaxy. *The Astrophysical Journal*, 779(1):43, December 2013.

M E Putman, J E G Peek, A Muratov, O Y Gnedin, W Hsu, K A Douglas, C Heiles, S Stanimirović, E J Korpela, and S J Gibson. The Disruption and Fueling of M33. *The Astrophysical Journal*, 703(2):1486–1501, October 2009.

Urvashi Rau, Sanjay Bhatnagar, and Frazer N Owen. Deep wideband single pointings and mosaics in radio interferometry - How accurately do we reconstruct intensities and spectral indices of faint sources? *arXiv.org*, page arXiv:1605.07640, May 2016.

Michael W Regan, Michele D Thornley, Tamara T Helfer, Kartik Sheth, Tony Wong, Stuart N Vogel, Leo Blitz, and Douglas C J Bock. The BIMA Survey of Nearby Galaxies. I. The Radial Distribution of CO Emission in Spiral Galaxies. *The Astrophysical Journal*, 561(1):218–237, November 2001.

J W Rich, W J G de Blok, T J Cornwell, E Brinks, F Walter, I Bagetakos, and R C Jr Kennicutt. Multi-Scale CLEAN: A Comparison of its Performance

Against Classical CLEAN on Galaxies Using THINGS. *The Astronomical Journal*, 136(6):2897–2920, December 2008.

K. Rohlfs and T. L. Wilson. *Tools of radio astronomy*. 2004.

J. D. Romney. Cross Correlators. In G. B. Taylor, C. L. Carilli, and R. A. Perley, editors, *Synthesis Imaging in Radio Astronomy II*, volume 180 of *Astronomical Society of the Pacific Conference Series*, page 57, 1999.

E Rosolowsky, G Engargiola, R Plambeck, and L Blitz. Giant Molecular Clouds in M33. II. High-Resolution Observations. *The Astrophysical Journal*, 599 (1):258–274, December 2003.

Erik Rosolowsky and Adam Leroy. Bias-free Measurement of Giant Molecular Cloud Properties. *The Publications of the Astronomical Society of the Pacific*, 118(8):590–610, April 2006.

Erik Rosolowsky and Joshua D Simon. The M33 Metallicity Project: Resolving the Abundance Gradient Discrepancies in M33. *The Astrophysical Journal*, 675(2):1213–1222, March 2008.

Erik Rosolowsky, Eric Keto, Satoki Matsushita, and S P Willner. High-Resolution Molecular Gas Maps of M33. *The Astrophysical Journal*, 661 (2):830–844, June 2007.

Eva Schinnerer, Sharon E Meidt, Jérôme Pety, Annie Hughes, Dario Colombo, Santiago García-Burillo, Karl F Schuster, Gaëlle Dumas, Clare L Dobbs, Adam K Leroy, Carsten Kramer, Todd A Thompson, and Michael W Regan. The PdBI Arcsecond Whirlpool Survey (PAWS). I. A Cloud-scale/Multi-wavelength View of the Interstellar Medium in a Grand-design Spiral Galaxy. *The Astrophysical Journal*, 779(1):42, December 2013.

Maarten Schmidt. The Rate of Star Formation. *Astrophysical Journal*, 129: 243–, March 1959.

R H M Schoenmakers, M Franx, and P T de Zeeuw. Measuring non-axisymmetry in spiral galaxies. *Monthly Notices of the Royal Astronomical Society*, 292:349–, December 1997.

Andreas Schruba, Adam K Leroy, Fabian Walter, Frank Bigiel, Elias Brinks, W J G de Blok, Gaelle Dumas, Carsten Kramer, Erik Rosolowsky, Karin Sandstrom, Karl Schuster, Antonio Usero, Axel Weiss, and Helmut Wiesemeyer. A Molecular Star Formation Law in the Atomic-gas-dominated Regime in Nearby Galaxies. *The Astronomical Journal*, 142(2):37, August 2011.

J A Sellwood and Ricardo Zánmar Sánchez. Quantifying non-circular streaming motions in disc galaxies. *Monthly Notices of the Royal Astronomical Society*, 404(4):1733–1744, June 2010.

Frank Y Shih. *Image Processing and Mathematical Morphology*. Fundamentals and Applications. CRC PressI Llc, 2009.

S Silich, T Lozinskaya, A Moiseev, N Podorvanuk, M Rosado, J Borissova, and M Valdez-Gutierrez. On the neutral gas distribution and kinematics in the dwarf irregular galaxy IC 1613. *Astronomy and Astrophysics*, 448(1): 123–131, March 2006.

Kristine Spekkens and J A Sellwood. Modeling Noncircular Motions in Disk Galaxies: Application to NGC 2976. *The Astrophysical Journal*, 664(1): 204–214, July 2007.

S. W. Stahler and F. Palla. *The Formation of Stars*. 2005.

Adrienne M Stilp, Julianne J Dalcanton, Steven R Warren, Evan Skillman, Jürgen Ott, and Bärbel Koribalski. Global H I Kinematics in Dwarf Galaxies. *The Astrophysical Journal*, 765(2):136, March 2013.

D Tamburro, H W Rix, A K Leroy, M M Mac Low, F Walter, R C Kennicutt, E Brinks, and W J G de Blok. What is Driving the H I Velocity Dispersion? *The Astronomical Journal*, 137(5):4424–4435, May 2009.

Guillermo Tenorio-Tagle and Peter Bodenheimer. Large-scale expanding superstructures in galaxies. *IN: Annual review of astronomy and astrophysics. Volume 26 (A89-14601 03-90). Palo Alto*, 26(1):145–197, 1988.

Guillermo Tenorio-Tagle, Casiana Muñoz-Tuñón, Enrique Pérez, Jesús Maíz Apellániz, and Gustavo Medina-Tanco. On the Ongoing Multiple Blowout in NGC 604. *The Astrophysical Journal*, 541(2):720–727, October 2000.

D A Thilker, R Braun, and R A M Walterbos. A High-Resolution Mosaic of H I in M33. *Seeing Through the Dust: The Detection of HI and the Exploration of the ISM in Galaxies*, 276:370–, December 2002.

A. R. Thompson. Fundamentals of Radio Interferometry. In G. B. Taylor, C. L. Carilli, and R. A. Perley, editors, *Synthesis Imaging in Radio Astronomy II*, volume 180 of *Astronomical Society of the Pacific Conference Series*, page 11, 1999.

C Trachternach, W J G de Blok, F Walter, E Brinks, and R C Jr Kennicutt. Dynamical Centers and Noncircular Motions in THINGS Galaxies: Implications for Dark Matter Halos. *The Astronomical Journal*, 136(6):2720–2760, December 2008.

Stfan van der Walt, Johannes L. Schnberger, Juan Nunez-Iglesias, Franois Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, and Tony and Yu. scikit-image: image processing in python. *PeerJ*, 2:e453, 6 2014. ISSN 2167-8359. doi: 10.7717/peerj.453. URL `http://dx.doi.org/10.7717/peerj.453`.

S Verley, E Corbelli, C Giovanardi, and L K Hunt. Star formation in M 33: multiwavelength signatures across the disk. *Astronomy and Astrophysics*, 493(2):453–466, January 2009.

Fabian Walter, Elias Brinks, W J G de Blok, Frank Bigiel, Robert C Jr Kennicutt, Michele D Thornley, and Adam Leroy. THINGS: The H I Nearby Galaxy Survey. *The Astronomical Journal*, 136(6):2563–2647, December 2008.

Christine D Wilson and Brenda C Matthews. The Star Formation Histories and Efficiencies of Two Giant H II Regions in M33. *Astrophysical Journal v.455*, 455:125, December 1995.

Christine D Wilson and Nick Scoville. The molecular gas content of the nuclear region of M33. *Astrophysical Journal*, 347:743–749, December 1989.

Christine D Wilson and Nick Scoville. The properties of individual giant molecular clouds in M33. *Astrophysical Journal*, 363:435–450, November 1990.

Christine D Wilson, Constance E Walker, and Michele D Thornley. The Density and Temperature of Molecular Clouds in M33. *The Astrophysical Journal*, 483(1):210–219, July 1997.

Mark G Wolfire, Christopher F McKee, David Hollenbach, and A G G M Tielens. Neutral Atomic Phases of the Interstellar Medium in the Galaxy. *The Astrophysical Journal*, 587(1):278–311, April 2003.

M C H Wright, P J Warner, and J E Baldwin. High resolution observations of neutral hydrogen in M33 - I. The hydrogen distribution. *Monthly Notices of the Royal Astronomical Society*, 155:337–356, 1972.

Melvyn C. H. Wright. *Galactic and Extra-Galactic Radio Astronomy*, chapter Mapping Neutral Hydrogen in External Galaxies, pages 291–319. Springer Berlin Heidelberg, Berlin, Heidelberg, 1974. ISBN 978-3-642-96178-6.

Reiko Yamaguchi, Norikazu Mizuno, Toshikazu Onishi, Akira Mizuno, and Yasuo Fukui. Molecular Clouds in the Supergiant Shell LMC 4: Evidence for Triggered Cluster Formation. *The Astrophysical Journal*, 553(2):L185–L188, June 2001.