Connecting galactic to local scales in the neutral interstellar medium across the Local Group

by

Eric William Koch

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

Doctor of Philosophy

Department of Physics
University of Alberta

Abstract

Star formation drives secular galaxy evolution by linking stars and gas in galaxies. This link forms part of the "baryonic cycle," where stars form from interstellar gas, and a portion of that gas is returned to the interstellar medium upon the star's death. Understanding the baryonic cycle and its role in galaxy evolution requires piecing together how the neutral interstellar medium controls where and when star formation occurs. However, our understanding remains limited because the neutral interstellar medium is affected by processes ranging from the kiloparsec scales of galaxies to the sub-parsec scales where individual stellar systems form. To advance our knowledge of the baryonic cycle and galaxy evolution, observations of the neutral interstellar medium must bridge large to small scales.

In my thesis, I present new observations of the Local Group galaxies that connect large to small scales in the neutral interstellar medium in exquisite detail. The Local Group galaxies provide an external view to trace galaxy-scale processes but are close enough for current telescopes to resolve < 100 pc scales where key components of the baryonic cycle occur. I present the first part of an on-going Local Group survey of atomic hydrogen taken with the Very Large Array. These observations provide high-spatial and -spectral resolution maps of M31 and M33 produced from my new techniques for handling massive interferometric data sets. I use these new observations to demonstrate the complex kinematics of the atomic interstellar medium and the large bias implicit in

using approximate line shape measurements. My work demonstrates that detailed spectral modeling is critically needed to guide our interpretation of the atomic interstellar medium. Applying detailed spectral modeling, I show that 21-cm HI emission is best modeled as a set of optically-thin Gaussians. I further show that previous results reporting opaque HI on 100 pc scales are strongly rejected by the new observations.

I then compare tracers of the atomic and molecular interstellar medium and find strong correlations between their kinematics. The molecular interstellar medium is the direct fuel for star formation, and this correlation indicates a continued role for the atomic interstellar medium throughout the star formation process. However, these correlations are apparent when only the atomic gas spectrally associated with the molecular gas is considered. Previous studies that use all of the atomic gas along the line-of-sight are unlikely to find this association. These results further suggest the need for careful spectral modeling to study processes that link the atomic and molecular media, including how molecular gas is formed from the atomic gas.

Finally, I demonstrate the difficulties in recovering the source of large-scale turbulence in nearby galaxies, which is a key but poorly constrained component in modern star formation theories. I show that features in the spatial power spectrum, previously interpreted as large-scale galaxy properties, are not physical and instead result from the instrument response function. Because of this, the source of turbulent driving remains ambiguous.

By combining high-spectral resolution observations of the atomic interstellar medium with detailed modeling, my work opens new avenues for exploring the neutral interstellar medium in nearby galaxies.

Preface

This thesis is original work by Eric William Koch, conducted under the supervision of Erik W. Rosolowsky.

Chapter 3 of this thesis is published in Koch, E.W et al. (2018). "Kinematics of the atomic ISM in M33 on 80 pc scales," *MNRAS*, 479, 2505–2533. E.W. Rosolowsky led the proposal for the VLA observations. A preliminary reduction of these observations, though later discovered to be flawed, was presented in my MSc. thesis (2016). I planned the observations with E.W. Rosolowsky and A.K. Leroy, performed the data reduction, imaging, and analysis. I led writing the manuscript with input from all authors.

Chapter 4 of this thesis is a manuscript in preparation and will be submitted to *Monthly Notices of the Royal Astronomical Society* as Koch, E.W et al. 2020, "A dearth of tophats in M31 & M33: HI spectra strongly prefer multigaussian over opaque model fits." A.K. Leroy led the proposal for and planned the M31 VLA observations. I performed the data reduction, imaging, and analysis. I led writing the manuscript in collaboration with J. Chastenet, I. Chiang, A.K. Leroy, E.W. Rosolowsky, K.M. Sandstrom, and D. Utomo.

Chapter 5 of this thesis is published in Koch, E.W et al. (2019). "Relationship between the line width of the atomic and molecular ISM in M33," *MNRAS*, 485, 2324–2342. I used the VLA M33 HI observations from Chapter 3 (Koch et al., 2018c) and CO(2-1) observations taken by the IRAM 30-m telescope that are published in Gratier et al. (2010) and Druard et al. (2014).

I performed the analysis and led writing the manuscript with input from all authors.

Chapter 6 of this thesis is published in Koch, E.W et al. (2020). "Spatial power spectra of dust across the Local Group: No constraint on disc scale height," *MNRAS*, 492, 2663–2682. Using dust modeling from Utomo et al. (2019a), I performed the analysis and led writing the manuscript with input from all authors.

This thesis makes use of observations taken with the NSF's Karl G. Jansky Very Large Array and the Green Bank Telescope. The National Radio Astronomy Observatory and the Green Bank Observatory are facilities of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

This research was enabled in part by support provided by WestGrid, Compute Canada, and CANFAR.

For my ever-supportive family.

And Sarah, my partner in all things.

Acknowledgements

Thank you to my supervisor, Erik Rosolowsky, who ended up being stuck with me for the past 9 years at 2 different universities. First and foremost, thank you for being a constant source of advice throughout my studies, starting with a nervous first-year undergraduate over 9 years ago who really wanted to "do research" (without any clue what that meant), to today, and likely well into the future. I appreciate the freedom I have had to guide my research with your constant support, encouragement, and guidance.

Over the past several years, I have had the opportunity to work with a wonderful group of collaborators, including several who have generously mentored me along the way. Thank you to Jason Loeppky, Stella Offner, Ryan Boyden, Adam Ginsburg, Blakesley Burkhart, Amanda Kepley, Andreas Schruba, Jonathan Braine, Megan Johnson, Jay Lockman, Snežana Stanimirović, and Julianne Dalcanton. A special thank you to the z0mgs collaboration, including Karin Sandstrom, Jérémy Chastenet, I-Da Chiang, and Dyas Utomo, who have yet to kick me off their weekly telecons. And finally, thank you to Adam Leroy, who has been a mentor, collaborator, and host during my 3 month stay of The Ohio State University in 2019.

Thank you to the many members of the Astrophysics Group at the University of Alberta. First, thank you to my committee members, Dmitri Pogosyan and Gregory Sivakoff, for their guidance through my MSc and PhD. It has been a pleasure working with you for the past 6 years (and a summer). Dur-

ing that time, I developed friendships with and had the opportunity to learn from my fellow graduate students and transient postdocs. Thank you in particular to Alex Tetarenko, Bailey Tetarenko, Arash Bahramian, Kenny Van, Soumen Deb, Joseph Nofech, Maria Pettyjohn, Andrew Hughes, Zhuo Chen, Dario Colombo, and Veselina Kalinova.

Thank you to my family for your love, encouragement, and support that have allowed me to follow my dreams.

Most of all, thank you my wonderful and amazing Sarah. Words cannot express how much your constant love and support mean to me.

Table of Contents

List of Tables

List of Figures

Abbreviations

AGN Active Galactic Nuclei.

ALMA Atacama Large Millimeter/Sub-millimeter Array.

BIC Bayesian Information Criterion.

CASA Common Astronomy Software Application.

CGM CircumGalactic Medium.

CMB Cosmic Microwave Background.

CNM Cold neutral Medium.

DGR Dust-to-Gas Ratio.

EBHIS Effelsburg-Bonn HI Survey.

FFT Fast Fourier Transform.

FWHM Full-Width-at-Half-Max.

GBT Green Bank Telescope.

GHz GigaHertz (1 GHz=10^9 Hz).

GMC Giant Molecular Cloud.

HIM Hot Ionized Medium.

HISA HI Self-Absorption.

HVC High Velocity Cloud.

HWHM Half-Width-at-Half-Max.

I/O Input/Output.

IC Index Catalogues.

IGM InterGalactic Medium.

IR InfraRed.

IRAC InfraRed Array Camera.

IRAM Institute de Radioastronomie Millimetrique.

ISM InterStellar Medium.

Jy Jansky (1 Jy=10^{-23} erg s^{-1} cm^{-2} Hz^{-1}).

kpc kiloparsec (1 kpc=3.086^{21}).

LMC Lmall Magellanic Cloud.

LTE Local Thermodynamic Equilibrium.

M31 Messier 31.

M33 Messier 33.

MCMC Markov Chain Monte Carlo.

MHD Magneto-HydroDynamics.

MIPS Multiband Imaging Photometer.

MRI Magneto-Rotational Instability.

MS Measurement Set.

MW Milky Way.

NGC New General Catalogue.

NRAO National Radio Astronomy Observatory.

PACS Photoconductor Array Camera and Spectrometer.

pc parsec (1 pc=3.086^{18}).

PSF Point Spread Function.

PV Position-Velocity.

RFI Radio Frequency Interference.

RRL Radio Recombination Line.

S/N Signal-to-Noise.

SD Single Dish.

SED Spectral Energy Distribution.

SFR Star Formation Rate.

SMC Small Magellanic Cloud.

SN SuperNova.

SPIRE Spectroscopic and Photometric Imaging Receiver.

UV Ultra-Violet.

VLA (Karl G. Jansky) Very Large Array.

WIM Warm Ionized Medium.

WNM Warm neutral Medium.

Chapter 1

Introduction

Galaxies recycle. Throughout cosmic time, this recycling—called the baryonic cycle—links generations of stars through the gas they are born from and the enriched material from nuclear burning they expel back into interstellar space upon their death. The baryonic cycle is responsible for many of the properties we observe today in nearby galaxies.

Galaxies are a discrete collection of three components: dark matter, stars, and the interstellar medium (ISM; gas, dust, and high energy particles). The galaxy mass is dominated by the dark matter halo, which set the gravitational potential that physically defines the extent of galaxies. The latter two components are the baryons we directly observe.

Galaxies come in many shapes and sizes, with differing amounts of stellar and ISM masses. Understanding how this diverse population formed requires piecing together its evolution throughout cosmic time. This evolution is driven by several processes: mergers between galaxies; feedback from active galactic nuclei (AGN); accretion from the circumgalactic (CGM) and intergalactic media (IGM); and finally, star formation. Galaxy accretion and star formation are two components related through the on-going baryonic cycle, which ultimately controls the stars and the ISM in all galaxies. In the absence of galaxy mergers and AGN, the baryonic cycle drives galaxy evolution.

While we know its importance and the broad details of how it operates, there remain many unknowns about how the baryonic cycle produces ob-

served galaxy populations. We know that nearby star-forming galaxies will deplete their current ISM mass in $\sim$ 2 Gyr (Leroy et al., 2013), far shorter than the current 13.6 Gyr age of the universe. Furthermore, star formation rates (SFRs) earlier in the universe were much higher, with a peak measured at redshift $z \sim 2$ (Madau & Dickinson, 2014). Because of this, gas must continuously accreted onto galaxies for star formation to continue to the present day. Galaxies with little or no current star formation, elliptical and lenticular galaxies, are evidence of star formation "quenching" when gas accretion has been disrupted or stopped (Kennicutt & Evans, 2012). This diversity among galaxy populations, and the processes that cause them, requires piecing together the baryonic cycle from gas accretion to the formation of stars. However, it remains difficult to make these connections, since accretion onto the galaxy occurs on $\sim$ kpc scales[1] (Putman et al., 2012), yet the formation of single stellar systems occurs on $\sim$ 0.1 pc scales (André et al., 2014). This large range of spatial scales naturally produces a large range in the associated time scales, where the $\sim$ 2 Gyr depletion time requires slow but consistent accretion rates on large scales, while the free-fall time for gravitational collapse of individual stellar systems is $\sim$ 100 kyr. Simultaneously accounting for these ranges is the primary issue in connecting star formation to galaxy evolution through the baryonic cycle.

My thesis is motivated by the missing pieces in the baryonic cycle which require bridging large to small scales in the ISM. In particular, I explore relations between the atomic and molecular ISM in the Local Group galaxies, which provide an external view to trace large galactic scales ($\sim$ kpc) but are close enough to resolve small scales ($\sim$ 10s pc) approaching scales resolved within the Milky Way. Because of their proximity, the Local Group galaxies are an ideal laboratory to answer key questions about the baryonic cycle.

In this chapter, I introduce the galactic ecosystem and the baryonic cycle in greater detail (§1.1). I then motivate how a gap in our understanding of

[1] "pc" refers to the astronomical distance unit, parsec. For reference, 1 parsec=3.3 light years.

the ISM and star formation has developed from a physical scale gap missed in most observations (§1.2), where small-scales are easily accessible in the Milky Way and galactic-scales are accessible from nearby galaxies. Given the current instrumentation available, the Local Group galaxies are the ideal targets to fill this gap in our knowledge. The remainder of §1.2 introduces the background for the three questions I explore in this thesis:

1. How do models of the 21-cm HI line shape affect our interpretation of the neutral ISM? (§1.2.1; Chapters 3 & 4)

2. Are atomic (HI) and molecular (H_2) ISM kinematics coupled on 80 pc scales? (§1.2.2; Chapter 5)

3. Is the source of large-scale turbulent driving imprinted on the ISM density structure? (§1.2.3; Chapter 6)

1.1 The galactic ecosystem & baryonic cycle

The interdependent processes that control the baryonic cycle and the secular evolution of galaxies produce a "galactic ecosystem." The internal ecosystem is driven by the baryonic cycle, linking the stellar and ISM baryonic components in a galaxy. Due to this link, the baryonic cycle takes place within the discs of galaxies, typically near their mid-plane ($\sim \pm 100$ pc), since most of the neutral ISM mass is located there (Ferrière, 2001).

We know the major stages in how the baryonic cycle operates. These can be summarized in two steps: stars form from interstellar gas, and a portion of that gas is returned back into the ISM upon the star's death. The cycle is completed when the gas from stellar death cools and mixes back into the cold ISM where star formation takes place. Figure 1.1 shows an overview of these steps in the baryonic cycle. However, there is a broad range in both physical and time scales involved with these steps. Furthermore, different steps of the baryonic cycle influence each other, creating feedback loops, and external mechanisms operating at galaxy scales (or beyond) influence this cycle, too.

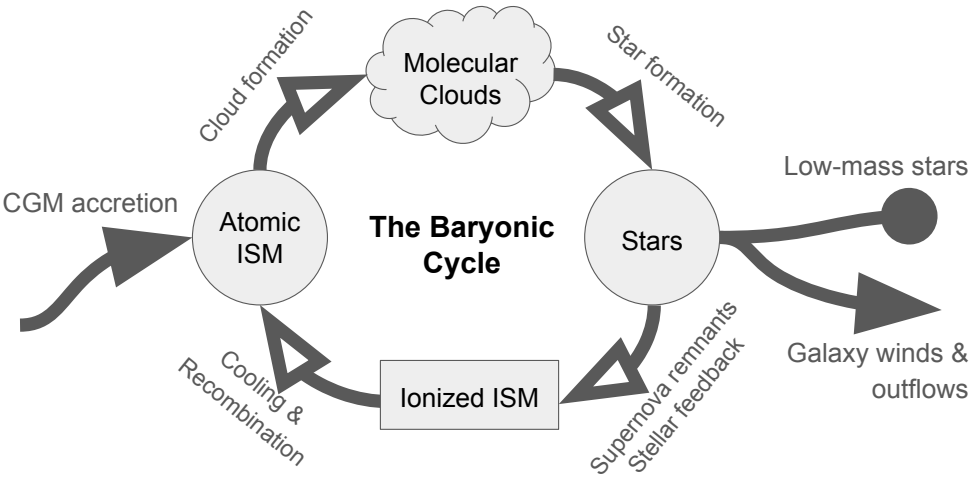

Figure 1.1: Schematic of the baryonic cycle. Open symbols indicate processes of the cycle that occur within the galactic disc, and closed symbols are processes that add or remove baryons from the cycle.

In the remainder of this section, I split the cycle shown in Figure 1.1 into three parts: converting the ISM to stars (§1.1.1), star formation's effect on its environment and the death of massive stars (§1.1.2), and external mechanisms that feed and remove the ISM from galactic discs (§1.1.3).

1.1.1 Converting gas to stars: the role of star formation

The baryonic cycle is linked through the different ISM properties within a galaxy (Figure 1.1). These properties vary dramatically in the ISM, with temperature and density varying by > 7 order-of-magnitude due to different processes affecting the environment. To classify these different ISM conditions, the ISM is classified into "phases," which are defined by the phase of hydrogen in those conditions since hydrogen dominates by mass and number in the ISM (Ferrière, 2001). The notation used is HII for ionized hydrogen, HI for neutral hydrogen, and H_2 for molecular hydrogen[2]. Though other atoms affect ISM

[2]This gives two H "twos," requiring some context to understand if ionized (HII) or molecular hydrogen (H_2) is being discussed.

properties, hydrogen plays a key role throughout.

The hottest and most diffuse ISM phases are the ionized media, often classified as a hot interstellar medium (HIM) and a warm ionized medium (WIM). The HIM is the hottest ($\sim 10^6$ K) and most diffuse (0.003 cm^{-3}) ISM component produced by supernovae shocks. About 20% of the ISM volume is filled by the HIM in the Milky Way (Ferrière, 2001). The WIM is results primarily from massive O and B stars, which produce strong UV radiation with energies exceeding the 13.6 eV needed to ionize hydrogen. This UV radiation produces HII regions surrounding these massive stars, which have temperatures of ~ 8000 K from the balance between photoelectric heating and recombination (Draine, 2011). There is also a diffuse and widespread WIM component filling $\sim 10\%$ of the ISM volume (Kulkarni & Heiles, 1987). This diffuse WIM requires approximately the ionizing energy output from most O and B stars in a galaxy *in addition to* the energy to HII regions. Because of this, it remains unclear whether a second ionizing source is required to produce the WIM (Ferrière, 2001). About 90% of the ionized gas mass is within the WIM (Haffner et al., 2009).

Below temperatures of $\lessapprox 10^4$ K, protons and electrons rapidly combine to form atomic hydrogen[3] (HI). This is the atomic ISM, which has two distinct phases: a warm neutral medium (WNM; $n \sim 1$ cm^{-3}; $T \sim 6000$ K) and a cold neutral medium (CNM; $n \sim 100$ cm^{-3}; $T \sim 100$ K). These states result from the stable pressure equilibrium solutions when heating and cooling processes balance (Field et al., 1969; Wolfire et al., 1995; Wolfire et al., 2003). Some fraction of the atomic gas is perturbed into an intermediate and unstable state due to external processes, such as turbulence (Bialy et al., 2017). In our solar neighbourhood, HI is the dominant ISM phase (Ferrière, 2001), though this drastically changes in different locations in the Milky Way and other nearby galaxies as the galactic environment changes (e.g., metallicity, interstellar radiation field, mid-plane disc pressure; Kennicutt & Evans, 2012).

One further state change of hydrogen is the conversion from atomic (HI)

[3]Without a strong nearby ionizing source, as is the case in HII regions.

to molecular (H_2) hydrogen. In small regions within a galaxy (< 10 pc), the ISM density can exceed $\sim 10^2$ cm^{-3}, allowing H_2 to rapidly form. Since the CNM reaches similar densities, it is likely the precursor material where the molecular ISM is formed from. In the Milky Way and other nearby galaxies, these collections of H_2 are often discrete and are surrounded by an envelope of HI. Due to their discrete nature, these collections are referred to as molecular clouds, or when their mass is $\gtrapprox 10^4$ M$_\odot$, giant molecular clouds (GMCs).

This transition to form molecular clouds from the CNM is a long standing problem. Some mechanism must act to gather the material together, increasing the density and, as a result, allow the HI to cool and form H_2. Many mechanisms have been proposed over the last 50 years, including: converging flows, agglomeration (or collisions) of smaller clouds, gravitational instability in galactic discs, and magnetic instabilities (see review by Dobbs et al., 2014). Several mechanisms have been explored in numerical simulations (e.g., Shetty & Ostriker, 2006; Dobbs, 2008), though there remains few observational constraints beyond studies of individual regions (e.g., Dawson et al., 2011). The mechanism of molecular cloud formation and the timescale it operates on is a key unknown in the baryonic cycle.

The H_2 in molecular clouds is the coldest and densest phase of the ISM, with $T \sim 10$ K and $n > 10^3$ cm^{-3}. The combination of low temperature and high density is an important property to form stars since, eventually, gravity must cause the collapse to convert cold gas into a star. Low temperatures are necessary such that thermal motions are small and do not support against gravity on the ~ 0.1 pc scales where individual stellar systems form. There is strong observational evidence that star formation *only* occurs in molecular clouds in the nearby universe (Leroy et al., 2008) (though see Krumholz, 2013). Figure 1.2 shows the nearby galaxy NGC 5055 in two ISM tracers and two star formation tracers. The HI is detected almost ubiquitously throughout the galaxy. The H_2 (as traced by carbon monoxide; CO), on the otherhand, is centralized into the centre of the galaxy. Both star formation tracers, showing star formation still embedded within molecular clouds at earlier times and

exposed at later times, are distributed far more closely to the H_2 than HI. This is similarly observed in nearby molecular clouds, where all stellar systems are found to form in the densest concentrations within the molecular clouds (André et al., 2014).

While gravitational collapse must occur to form stars, the star formation rate from observations is much lower than would be implied from gravitational collapse alone (e.g., Kennicutt, 1989). Observations find that the depletion time to convert all current molecular gas in a galaxy into stars is ~ 2 Gyr, *far* longer than the ∼few Myr free-fall time for a molecular clouds (Leroy et al., 2013) Star formation is therefore *inefficient*.

Much of the star formation research in the past several decades has sought some additional mechanism to explain this inefficiency through different mechanisms that support against gravity. Which mechanism is responsible has evolved from magnetic fields to turbulence and stellar feedback over the last few decades (Kennicutt & Evans, 2012). These changes in the dominant processes in star formation result from improved observations and increasingly realistic numerical simulations. However, each of these mechanisms—gravity, magnetic fields, turbulence, and stellar feedback—*is* likely to be important for star formation, though likely at different scales or stages. Furthermore, each of these mechanisms may be mutually influenced by the others, making it difficult to draw general conclusions about their importance at different times. Because of this mutual influence between ISM processes, an encompassing star formation theory remains elusive.

Whatever the processes involved in star formation are, it remains an obvious bottleneck in the baryonic cycle. Starting from densities of 100 cm^{-3} in the cold neutral medium, gravity must eventually cause large collections of gas to collapse into stars, where the core densities reach $\gtrapprox 10^{26}$ cm^{-3} (for a 1 $M_\odot$ main-sequence star; Carroll & Ostlie, 2007). This is a dramatic change that occurs through star formation, as the density increases by ~ 24 *orders of magnitude*. Another way of viewing this extreme change is the initial volume that the matter for an 1 $M_\odot$ star would fill. This change in density is equivalent

to an initial volume of ~ 0.5 pc^3 that must decrease by a factor of 10^{22} into the volume of the Sun[4]. The formation of each star represents this extreme change[5].

Two key results from the aftermath of star formation provide important constraints for any successful star formation theory. The first is that molecular clouds only live for $\lessapprox$10s of Myr. By comparing the locations of molecular clouds with young stellar clusters, or tracers of recent star formation, the physical separations strongly suggest that stellar feedback rapidly disperses molecular clouds soon after star formation starts (e.g., Kawamura et al., 2009; Chevance et al., 2020b). This is critical for understanding the inefficiency of star formation. A short cloud lifetime strongly constrains the entire process of from forming a molecular cloud to creating stars to a time scale of ~ 10 gravitational free-fall times. Cloud formation mechanisms must be able to produce a star-forming environment in some fraction of this time, roughly constrained to be ∼few Myr from observations (Kawamura et al., 2009). This suggests there should be strong *kinematic* evidence of the mechanism required to gather atomic gas into the $> 10^4$ M$_\odot$ GMCs in nearby galaxies.

The second result is that the initial mass function (IMF) of stars is remarkably constant (Offner et al., 2014), despite the properties of molecular clouds (e.g., Colombo et al., 2014) and stellar clusters (e.g., Adamo et al., 2015) varying with galactic environment (Jeffreson & Kruijssen, 2018). Most stars are thought to form in loose clusters that relax and disperse within a few galactic orbital times (Lada & Lada, 2003), producing the "field" stars in galaxies (e.g., the Sun). Denser stellar clusters, however, remain bound for much longer periods of time; the most extreme, though ancient, examples of these are globular clusters. For star formation theories, the challenge remains to determine how the properties of the ISM vary in different galactic environments, yet form

[4]The ISM is highly structured (André et al., 2014), however, so it is more realistic that portions of the material would have densities a few orders of magnitude larger. Even so, a vast change in density is still needed.

[5]The first stars (population III) in the universe are thought to be much more massive, and larger, than star formed at later times (population I; Krumholz, 2017). The density change is, nevertheless, extremely large.

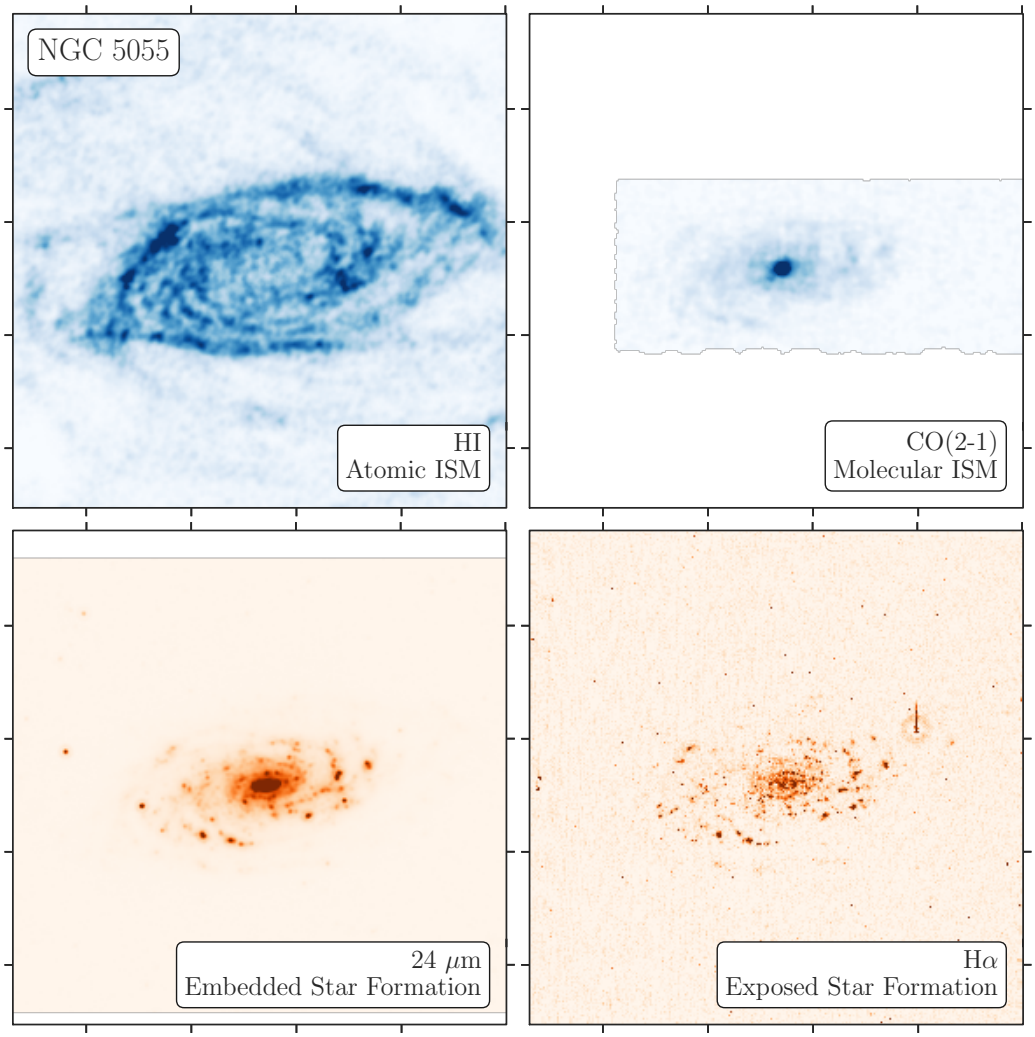

Figure 1.2: The nearby galaxy NGC 5055 shown in HI (Walter et al., 2008), CO(2-1) (Leroy et al., 2009), *Spitzer* MIPS 24 μm (Kennicutt et al., 2003), and Hα (Regan et al., 2001). This galaxy is a clear example of the strong correlation between star formation tracers and molecular gas (traced by CO), and lack of correlation with HI, which is detected more ubiquitously across the galaxy (Kennicutt & Evans, 2012).

stars in roughly the same mass distribution.

1.1.2 Stellar death, feedback and enrichment

Once stars form and reach the main sequence, most remain in an equilibrium state until the accessible fuel (H) is exhausted. Given the extreme temperature sensitivity of nucleosynthesis, the rate that massive stars burn their fuel is significantly faster than low-mass stars. Because of this, $M > 8$ $M_\odot$ stars have short lifetimes of < 30 Myr[6]. This means that, when present, massive stars live most of their lives near the molecular cloud they formed from. The significant energy output from these stars influences their natal environments, perhaps enough to disperse the entire molecular cloud. This is the primary role of stellar feedback in star formation theories: providing a mechanism that rapidly destroys and unbinds the molecular clouds to stop star formation, thereby making it inefficient (Krumholz, 2014).

For the most massive stars that undergo a supernova (SN; $M \gtrapprox 8$ $M_\odot$), this energy may be directly deposited into the nearby dense ISM (Walch et al., 2015). The average momentum per unit mass deposited by a supernovae explosion (SNe) is $\sim 3 \times 10^3$ km s^{-1}, which, in addition to destroying the natal molecular clouds, is expected to drive turbulence in galaxies (Mac Low & Klessen, 2004b; Tamburro et al., 2009).

Lower mass stars also affect their environment, for example, through protostellar outflows (Krumholz et al., 2014). These feedback sources are not explosive mechanisms like for massive stars, but still impact their natal environments with a cumulative energy output capable of driving turbulence on sub-cloud scales or unbinding the whole molecular cloud (e.g., Arce et al., 2010). However, the lifetimes of low-mass stars greatly exceed a molecular cloud lifetime, and so the stellar deaths of low-mass stars do not influence their natal molecular clouds.

Stars of all masses eventually return some portion of their mass back into

[6]Using the relation $\tau_{\rm MS} = 10\,(M/M_\odot)^{-3}$ Gyr (Carroll & Ostlie, 2007).

the ISM[7]. This ejected material is enriched through nucleosynthesis, decreasing the fraction of hydrogen and increasing various heavier elements. The fraction of non-hydrogen and helium atoms is referred to as "metallicity." This enriched material eventually cools and mixes back into the ISM, causing the next generation of stars to differ in chemical composition from previous generations.

Though hydrogen remains the dominant ISM component in mass and number, the observed abundances of elements such as C, O, Mg, Si, and Fe comprise an additional and crucial ISM component: dust grains. Dust grains consist of lattice like structures of the above elements, including large carbon-based molecules, like polycyclic aromatic hydrocarbons (i.e., flat multi-ring C molecules; Draine & Lee, 1984). These dust grains have two important roles in the neutral ISM: (1) They catalyze the formation of H_2 far faster than gas phase reactions, and as such, are expected to have produced essentially all H_2 in the nearby universe. (2) Dust absorbs UV and optical light (e.g., extinction), where the grains thermally radiate into the IR, and provide an important shielding source for H_2 and other molecules found in molecular clouds (Ferrière, 2001).

1.1.3 Galaxies are not closed systems: inflow and outflow in galaxies

While there are complexities to the baryonic cycle presented in the above two sections, most of those processes occur within the galactic disc. However, galaxies are not closed systems. Mass can both be added and lost through different processes, with the galactic halo and surrounding circumgalactic medium providing an interface onto the galaxy (Putman et al., 2012). In turn, these components interact with the wider intergalactic medium (IGM). Further environmental interactions occur through galaxy mergers and the influence of galaxy clusters, as briefly mentioned at the beginning of this chapter.

[7]Though for red dwarfs, whose main-sequence lifetime exceeds the Hubble time, their mass can essentially treated as "lost." This is most of the mass that goes into stars.

Star-forming galaxies in the nearby universe will deplete their molecular ISM in $\sim$ 2 Gyr at the current levels of star formation (Leroy et al., 2013). Because of this, on-going star formation in galaxies require the ISM to be replenished from the surrounding CGM. There are several pieces of observational evidence for galaxy accretion, including high velocity clouds (Dickey & Lockman, 1990) and other off-rotation atomic and ionized gas components found towards nearby galaxies (Fraternali et al., 2004). To feed the inner galactic discs, where star formation is typically concentrated, ISM replenishment requires a radial mass transport (Krumholz et al., 2018), which has been measured in a small number of face-on spiral galaxies (Schmidt et al., 2016). Measuring the accretion rate onto and within galaxies remains difficult, and as such, observational constraints are only available for a few nearby galaxies.

Acting against galaxy accretion are galaxy winds and outflows, driven by a stellar feedback and/or AGN, where material is driven off of the disc and into the halo or beyond. Particularly for AGN, this effect is well-known based on observations of radio jets, and the strong correlation between black-hole mass and stellar velocity dispersion in galactic bulges suggests this feedback plays an important role in galaxy evolution (e.g., Veilleux et al., 2005).

Additionally, recent observations from nearby starbursts resolve multi-phase ISM structure in galactic winds (Bolatto et al., 2013b; Leroy et al., 2015), where the intense feedback from either source is sufficient to remove bulk ISM components in close proximity (Krieger et al., 2019).

Combined with the ISM and star formation within the galactic disc, this inflow and outflow increase the difficulty in understanding galaxy evolution through the baryonic cycle. Primarily, this difficulty results from considering the galaxy as an open system and introducing larger than galaxy scales into the cycle.

1.2 Bridging the neutral ISM from extragalactic to Galactic scales with Local Group galaxies

One issue that is broadly faced in ISM and star formation studies is the dichotomy of physical scales that are observed in detail. Due to a natural bias of using telescope time to observe what is "bright, " our understanding of the ISM is guided by observations in the Milky Way, which provide extraordinary detail at small physical scales ($\ll 1$ pc), and of external galaxies, which test a broad range of ISM environments both between and within galaxies. From these two regimes, there remains a gap in the physical scales typically observed between 10s to ~ 100 pc. In the Milky Way, it is difficult to provide large-scale context due to our place within the disc and the ambiguity in kinematically determined distances. Extragalactic observations provide the large-scale context to link with the galactic environment, but often fail to probe these smaller scales due to the limitations in telescope resolution or unrealistic amount of observing time to reach the required sensitivity.

Though purely observational, this gap in spatial scales has propagated to a gap in ISM and star formation theories. Following Krumholz (2014), ISM and star formation theory can be categorized as bottom-up (motivated by Milky Way observations) or top-down (motivated by extragalactic observations). Neither class of model can be extrapolated to produce consistent results, though a few relations do successfully connect the two regimes (e.g., gas surface density and the SFR per free-fall times; Krumholz, 2014). These agreements suggest that recent models are on the right path to build a self-consistent framework, but their disagreements suggest they still miss some ingredients.

An obvious solution to this dichotomy is to fill the gap in physical scales probed by observations. This is where the Local Group galaxies play a key role. As the nearest external galaxies, the Local Group galaxies provide the benefits of both Milky Way and extragalactic observations in the range of spatial scales

that can be explored. The key drawback is the small sample size: there are two other major galaxies in the Local Group (M31 and M33), along with ≈ 6 gas-rich and massive star-forming dwarf galaxies[8] (LMC, SMC, NGC 6822, IC10, IC1613 & WLM). Each galaxy is either a different galaxy type, or has slightly different properties and environment, providing a reasonable span over galaxy types despite the small sample size (Karachentsev et al., 2004). In this thesis, I focus on the Large Magellanic Cloud (LMC), Small Magellanic Cloud (SMC), Messier 31 (M31) and Messier 33 (M33; Table 1.1).

The closest gas-rich dwarfs, the LMC and SMC, provide the best match in physical scales to the Milky Way, and neutral ISM observations bridging the spatial gap have been available for ~ 20 years (Stanimirovic et al., 1999a; Kim et al., 2003). Recent ALMA observations have pushed down to ~ 0.1 pc resolution in several molecular lines (e.g., Indebetouw et al., 2013), that can be directly compared to surveys of nearby (< 500 pc) molecular clouds in the solar neighbourhood. While both galaxies provide excellent probes matched directly to Milky Way observations, their galactic environments are quite different, both in their galaxy structure and their low-metallicity environment.

About an order-of-magnitude farther in distance are the two other major galaxies in the Local Group: M31 and M33. M31 is the other major galaxy in the Local Group, with a similar mass to the Milky Way. Because of our external perspective, M31 provides an interesting analog to compare with the Milky Way's galactic structure, star formation, and ISM properties. One clear difference from the Milky Way, however, is the ring-like structure which may be due to a past major disturbance (Williams et al., 2015), possibly with the dwarf galaxies M32 and NGC 205 (Gordon et al., 2006), though it remains unclear whether this interaction is the sole source of the rings (Lewis et al., 2015).

M33 is a late-type dwarf spiral, with a total mass $\sim 40\%$ of M31, and provides a interesting comparison to the more massive spiral M31. Despite

[8]This number depends on the distance used to define the Local Group. I have used $D < 1$ Mpc here.

having a much smaller stellar mass (Table 1.1), M33 currently forms stars at a higher rate than M31 and has a similar total H_2 mass. There is further a large metallicity gradient in M33 (Rosolowsky & Simon, 2008; Bresolin, 2011), such that the outer disc has a metallicity < 0.5 $Z_\odot$, a factor of 2 lower than the Milky Way. This change in metallicity makes M33 ideal for testing how ISM properties vary with galactic environment (e.g., Bialy & Sternberg, 2019). M33 is a "flocculent" structure with spiral arm fragments rather than the well-defined grand spiral structure. However, M33 has a moderate inclination which provides our closest top-down view of a spiral galaxy.

	M31	M33	LMC	SMC
Type (dV91)	SA(s)b	SA(s)cd	SB(s)m	SB(s)m pec
Distance (kpc)	744 (V10)	840 (F01)	50 (K14)	62 (G14)
Diameter (D_{25}, kpc; dV91)	39.2	17.8	9.9	5.8
Systemic velocity (km s^{-1})	-306 (C10)	-179 (K18)	293 (vdM02)	148 (DT19)
Inclination ($^\circ$)	78 (C10)	55 (K18)	35 (vdM02)	51 (DT19)
Rotation velocity (km s^{-1})	240 (C10)	110 (K18)	65 (vdM02)	55 (DT19)
$M_{\rm total}$ (M$_\odot$)	1.3×10^{12} (C10)	5.0×10^{11} (C14)	8.7×10^{9} (vdM02)	2.4×10^{9} (DT19)
$M_{\rm stellar}$ (M$_\odot$)	1.3×10^{11} (C10)	4.8×10^{9} (C14)	2.7×10^{9} (vdM02)	4.8×10^{8} (DT19)
$M_{\rm HI}$ (M$_\odot$)	5.8×10^{9} (Ch. 2)	1.8×10^{9} (K18)	5.0×10^{8} (K03)	4.7×10^{8} (DT19)
$M_{\rm H_2}$ (M$_\odot$)	3.6×10^{8} (N06)	3.1×10^{8} (D14)	6.3×10^{7} (J16)	2.0×10^{7} (J16)
$M_{\rm H_2}/M_{\rm HI}$	0.06	0.17	0.12	0.04
SFR (M$_\odot$ yr^{-1})	0.3 (R16)	0.45 (V09)	0.20 (S99)	0.033 (S99)

B09 (Braun et al., 2009) ⋄ C10 (Corbelli et al., 2010) ⋄ C14 (Corbelli et al., 2014) ⋄ D14 (Druard et al., 2014) ⋄ DT19 (Di Teodoro et al., 2019) ⋄ F01 (Freedman et al., 2001) ⋄ G14 (Graczyk et al., 2014) ⋄ J16 (Jameson et al., 2016) ⋄ K03 (Kim et al., 2003) ⋄ K14 (Klein et al., 2014) ⋄ K18 (Koch et al., 2018c) ⋄ N06 (Nieten et al., 2006) ⋄ R16 (Rahmani et al., 2016) ⋄ S99 (Stanimirovic et al., 1999a) ⋄ dV91 (de Vaucouleurs et al., 1991) ⋄ vdM02 (van der Marel et al., 2002) ⋄ V09 (Verley et al., 2009) ⋄ V10 (Vilardell et al., 2010)

Table 1.1: Local Group galaxy properties. Total galaxy masses include the dark matter halo.

The current limitation for tracing the atomic and molecular ISM in nearby galaxies is the resolution of cm-interferometers. With massive resolution and sensitivity gains for mm-inteferometers in the previous decade (ALMA, SMA, NOEMA[9]), observations down to ~ 1 pc scales are accessible for targets within $D \sim 1$ Mpc (e.g., Tokuda et al., 2020).

On the other hand, world-class cm-interferometers, like the Very Large Array (VLA), are limited to $\sim 6''$ angular scales for imaging with high fidelity (Chapter 2). At $D = 1$ Mpc, this means that ~ 25 pc scales can be traced with the 21-cm HI line, roughly matching the size of giant molecular clouds (e.g., Dobbs et al., 2014). This limitation makes the Local Group galaxies the *only* targets where both neutral ISM phases can be traced to GMC scales. As such, these are the key targets for bridging the scale gap in ISM studies, until the next generation cm-interferometers[10] provide the order-of-magnitude gains the the mm-regime has recently experienced.

In the remainder of this section, I provide a more detailed background of the specific unknowns that lead into the key questions I explore in this thesis. Each of these questions addresses explores the neutral ISM properties within the observational gap between large and small scales. These questions fall broadly into three categories: (1) observational interpretations of the 21-cm line, (2) the relationship between the atomic and molecular ISM, and (3) the source of large-scale turbulent driving.

1.2.1 Modeling and interpretation of the HI line shape in nearby galaxies

Perhaps more than any other single spectral-line, the atomic medium traced through the 21-cm HI line traces, or is affected by, most of the baryonic cycle. The 21-cm HI line traces several processes directly, including the thermodynamics of the atomic ISM (Field et al., 1969; Wolfire et al., 2003; Heiles & Troland, 2003), molecular cloud formation and the outer envelope of molecu-

[9] Atacama Large Millimeter/submillimeter Array (ALMA); SubMillimeter Array (SMA); NOrthern Extended Millimeter Array (NOEMA).

[10] i.e., the Square Kilometer Array (SKA) and next-generation Very Large Array (ngVLA).

lar clouds (Fukui et al., 2009; Dobbs et al., 2014), the large-scale turbulence in galaxies (Elmegreen et al., 2001), the galactic ISM structure (i.e., disc scale height and length; Dickey & Lockman, 1990), the inflow of cool gas from the CGM (Fraternali et al., 2002a), and the galactic rotation curve used to infer the dark matter halo (e.g., Corbelli et al., 2010; Corbelli et al., 2014). Indirectly, 21-cm HI is affected by stellar or AGN feedback (e.g., galactic winds; Tenorio-Tagle & Bodenheimer, 1988), which can drive outflows, turbulence and alter the HI structure by producing "bubbles" or holes in the atomic gas. These processes affect different components of the baryonic cycle (Figure 1.1).

Due to these different processes, spectral modelling of 21-cm HI spectral lines remains a challenge. From thermal and turbulent motion, the HI line shapes are expected to be close to Gaussian, however, multiple features along a line-of-sight from multiple emitting sources or due to turbulent effects (e.g., Lazarian & Pogosyan, 2000) complicate the spectrum. Similarly, opacity effects will alter the shape of HI features. These spectral properties are intricately related to the different physical processes affecting the atomic ISM, and due to the limited information from spectral line observations, it can be difficult to disentangle the line profiles and recover the physical properties of the atomic ISM. In this section, I review spectral modeling of the 21-cm line in nearby galaxies to highlight how different models produce large discrepancies in the inferred properties of the atomic ISM.

The 21-cm(1420 MHz) HI spectral line results from a hyperfine transition in atomic hydrogren's ground state. The interaction of the magnetic moments of the proton and electron produces two different states with a higher energy with parallel spins and a lower energy with anti-parallel spins. To reach the higher energy level, HI is collisionally excited mostly by other HI atoms.

The strength of a radiative transition is measured with the emission coefficient, or Einstein-A coefficient. Between two magnetic dipoles, the emission coefficient 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 (Stahler & Palla, 2005). Plugging in the 1420 MHz frequency, we find that the 21-cm HI line has a tiny emission coefficient of $A \approx 2.85 \times 10^{-15}$ s^{-1}. Expressed instead as $\tau_{1/2} = A^{-1}$, the half-life for radiative transitions if 10^7 yr for HI to emit. Because of this long time scale, most HI will collisionally de-excite and radiative de-excitation will be rare. What makes the 21-cm HI bright enough to observe is the massive number of HI atoms.

The level populations of 21-cm in the upper and lower hyperfine states are defined in terms of the spin temperature $T_{\rm spin}$ such that:

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

where u is the upper parallel spin state and l is the lower anti-parallel state. The upper state is a triplet ($g_u = 3$) and the lower is singlet ($g_l = 1$), resulting in a 3:1 ratio. Across the range of HI spin temperatures ($\sim$ 10–8000 K), we expect this 3:1 ratio to hold, since $h\nu/k = 0.0682$ K is far smaller than the CNM ($\sim$ 100 K) and CMB temperatures ($\sim$ 2.8 K).

Assuming local thermodynamic equilibrium (LTE), the line absorption coefficient (κ_ν) can then be written as

$$\kappa_\nu = \frac{3c^2}{32\pi} \frac{A_{10} n_{\rm H} h}{\nu_{10} k T_{\rm s}} \phi(\nu) \tag{1.3}$$

where $\phi(\nu)$ is the spectral line profile normalized to unity. The spectral line profile is set by the the dominant source of internal motion within the emitting volume. For typical properties in the atomic ISM, thermal motion[11] will dominate producing a Gaussian line shape:

$$\phi(\nu) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\nu - \nu_0)^2}{2\sigma^2}\right), \tag{1.4}$$

where σ is the line width and ν_0 is the line centre. Note that $\phi(\nu)$ can be equivalently expressed as the Doppler-shifted velocity v. Using the radio definition, $v = c(\nu_0 - \nu)/\nu_0$ with c as the speed of light.

[11] Turbulence as a broadening source will be mentioned below.

The optical depth is defined as the integral of the line absorption coefficient along the line-of-sight depth s:

$$\tau(\nu)(s) = \int_{s_0}^{s} \kappa_\nu(s) \mathrm{d}s. \tag{1.5}$$

If we assume that HI is isothermal along the line-of-sight (T_s is constant), the integrand does not depend on the path length s, and the optical depth is

$$\tau(\nu) = \frac{3c^2}{32\pi} \frac{A_{10}h}{\nu_{10}kT_\mathrm{s}} N_\mathrm{H} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\nu-\nu_0)^2}{2\sigma^2}\right), \tag{1.6}$$

where $N_\mathrm{H} = n_\mathrm{H}(s - s_0)$ is the column density integrated over the path length. Since the path length is not often constrained in 21-cm HI observations, the column density is the primary measure of interest. Equation 1.6 shows a key relation between optical depth and spin temperature: $\tau_\nu \propto T_\mathrm{s}^{-1}$.

Formally, the specific intensity (I_ν) is measured in observations. However, it is convenient to instead express specific intensity as a brightness temperature using the Rayleigh-Jeans approximation:

$$T_\mathrm{b} = \frac{c^2}{2k\nu^2} I_\nu. \tag{1.7}$$

This approximation is valid at centimetre wavelengths.

In terms of brightness temperature, the radiative transfer equation can then be written as

$$T_\mathrm{b}(\nu) = T_\mathrm{s}\left[1 - e^{-\tau(\nu)}\right] + T_\mathrm{c} e^{-\tau(\nu)}, \tag{1.8}$$

where T_c background temperature. Against a bright background source $T_\mathrm{c} \gg T_\mathrm{s}$, an absorption line will be described by:

$$T_\mathrm{b}(\nu) = T_\mathrm{c} e^{-\tau(\nu)}. \tag{1.9}$$

The strength of the absorption line depends directly on the optical depth, τ. Since $\tau \propto 1/T_\mathrm{s}$, HI absorption preferentially traces cold gas with low T_s. I note that T_c can be any background source, including a local background of bright high T_s, where foreground cold gas with low T_s will self-absorb[12].

[12]However, this requires T_s to vary along the line-of-sight, altering the radiative transfer equation (e.g., Dickey et al., 2003; Heiles & Troland, 2003).

In the absence of a background source, the HI emission line is described by:

$$T_{\rm b}(\nu) = T_{\rm s}\left[1 - \exp(-\tau(\nu))\right]. \tag{1.10}$$

In the optically-thin limit, $1 - \exp(-\tau(\nu)) \approx \tau(\nu)$, and since $\tau_\nu \propto T_{\rm s}^{-1}$ (see above), the brightness temperature is independent of $T_{\rm s}$. From Equation 1.6, the optically-thin HI emission will be (1) described by a single Gaussian line shape, and (2) $T_{\rm b} \propto N_{\rm H}$. The latter is an extremely important result since the total HI column density $N_{\rm H}$ can be measured by integrating over the HI spectrum. Taking this one further step, the column density can be converted to the HI mass with the distance to the source (D; Mpc) and telescope beam size (θ; $'$):

$$M = 0.34 D^2 \theta^2 \int T_{\rm b} {\rm d}v, \tag{1.11}$$

where mass is in $M_\odot$, $T_{\rm b}$ in K and the spectral axis in velocity v with units of km s^{-1} (Rohlfs & Wilson, 2004).

When the emission is opaque ($\tau > 1$), the line shape is altered from a Gaussian in Equation 1.10 (Rohlfs et al., 1972; Braun et al., 2009). Effectively, the observed line shape will have Gaussian-shaped line wings and a flattened top over the spectral range where $\tau > 1$. In this case, Equation 1.10 can constrain both $N_{\rm H}$ and $T_{\rm s}$.

Where does the difficulty in modeling the HI spectral line come from? First, $T_{\rm s}$ likely changes along a line-of-sight when intersecting different atomic ISM phases, and it may not be possible to distinguish these changes from the observed spectral line shape alone. Second, other processes, such as turbulence, will alter the line shape, whose influence may further be degenerate with the thermal atomic ISM properties. Distinguishing these processes is the primary goal behind most HI spectral line modeling.

The atomic ISM phases affect the HI line shape due to the difference in opacity of the WNM and CNM. This difference follows from the classical interpretation of the two-phase atomic ISM (Field et al., 1969; Wolfire et al., 2003), where the WNM and CNM have a temperature and density that vary by $\sim$ 2 orders of magnitude while in pressure equilibrium. In reality, the

atomic ISM phases are more complex as some fraction of the atomic ISM is continuously perturbed into an intermediate unstable regime (the UNM) by turbulence (Heiles & Troland, 2003; Bialy et al., 2017; Murray et al., 2018). Due to its low density, the WNM is optically thin and fills a much large volume relative to the CNM (Ferrière, 2001). Colder and denser atomic HI, comprising the CNM and cool UNM, fills a smaller volume more similar to the molecular ISM (Ferrière, 2001). On molecular clouds scales ($\sim$ 20 pc), the structure ofHI emission resolves into a combination of the atomic HI phases that can be distinguished by the narrow and wide line widths, based on the thermal line width, of individual components in the spectrum (e.g., Haud, 2000). However, velocity blending along the line-of-sight and turbulent broadening limits the completeness of separating the phases with this method.

The cold and denser HI is more likely to be optically-thick, affecting how this component is distinguished in observations. This opaque HI contributes to the HI emission like the optically-thin HI, but it can be cleanly separated by measuring it in absorption against a bright background source. Figure 1.3 shows a simplified schematic of HI emission and absorption, similar to what was observed in the Milky Way (Bihr et al., 2015). At coarser spatial resolution ($>$ 100 pc), typical of nearby galaxy observations, the HI is often assumed to be optically-thin since the cold and/or opaque HI likely has a small beam filling factor[13]. These variations in the HI line opacity make it difficult to distinguish between the different atomic ISM phases in observations. However, measuring the fraction of gas in these phases is critical as it constrains fundamental properties, including the total atomic ISM mass.

Our ability to distinguish the opaque CNM and thin WNM further depends on geometry of the target along our line-of-sight. In the Milky Way, our position within the Galactic disc makes HI detectable towards most lines-of-sight (see review by Dickey & Lockman, 1990). Because of this, there are many bright background continuum sources where opaque HI can be measured

[13]The beam filling factor is the ratio of the true source size to the resolution in the observations.

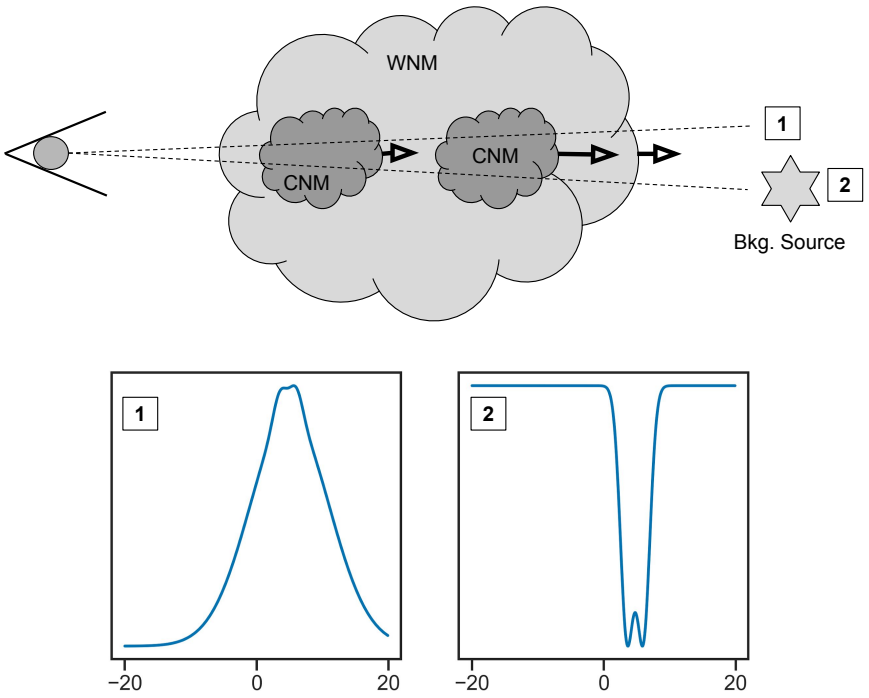

Figure 1.3: A simplified schematic of 21-cm HI emission and absorption against a bright background continuum source. Values shown are qualitative only. Similar lines-of-sight are traced, where (2) intersects a background source and (1) does not, and their resulting spectra are shown in the bottom panels. The top picture shows three emitting "clouds" moving away from the observer at slightly different velocities. Towards line-of-sight (1), all three clouds combine to produce the emission spectrum. Due to their similar velocities, and the much larger line width of the "WNM" cloud, the components are highly blended in emission, making it difficult to separate and model the emission from each cloud. However, the two small "CNM" clouds have a higher optical depth that is preferentially traced in absorption, as shown in the bottom right spectrum. Milky Way observations often find cases similar to this example (though typically far more complex), and the combination of emission and absorption is used to determine the atomic ISM phases (e.g., Dickey et al., 2003; Heiles & Troland, 2003). This technique is also used in nearby galaxies, though their smaller angular size on the sky provides far fewer lines-of-sight that intersect a bright background source. Observations of HI emission in nearby galaxies further suffer from coarser physical resolution. This often results in the assumption that the HI emission is primarily from the WNM, since it fills a larger volume in the ISM, and therefore tends to fill the telescope beam (Ferrière, 2001).

(Heiles & Troland, 2003; Murray et al., 2018; Wang et al., 2020a). Furthermore, there are many lines-of-sight where foreground opaque HI self-absorbs against a bright WNM background, referred to as HI self-absorption (HISA; Gibson et al., 2000; Gibson et al., 2005; Kavars et al., 2005; Soler et al., 2019; Wang et al., 2020b). In both cases, opaque CNM that produces clear absorption lines are the easiest to detect.

External galaxies have a much smaller footprint on the sky where absorption can be clearly distinguished, and so estimates of the CNM fraction is more difficult to constrain. The best constraints from absorption come from the Local Group galaxies as their larger angular size intersects more bright background continuum sources (Braun & Walterbos, 1992; Dickey & Brinks, 1993; Dickey et al., 1994, 2000; Jameson et al., 2019). Self-absorption features are not obvious from current observations, though there are hints of its influence (Liu et al., 2019). If cold opaque HI is beam-filling[14], however, the HI *emission* line shape may be observed as a "flat-top," where the HI becomes optically-thick over a range of velocities (Rohlfs et al., 1972; Braun et al., 2009; Braun, 2012). This offers the potential to map opaque HI and recover the total atomic ISM mass in nearby galaxies if such features are observed.

Affecting the recovery of all thermal properties is line broadening due to turbulence. On small scales, turbulence can have a drastic effect on the HI line shape, producing distinct spectral features in HI spectra that arise from the same emitting volume (Lazarian & Pogosyan, 2000). On larger scales, these features are averaged over and tend to produce a near-Gaussian shape that is broadened with respect to the thermal line width. These larger line widths are most noticeable in cold gas (atomic and molecular), where the thermal line widths are small ($\lessapprox 0.8$ km s^{-1} at < 100 K), and the turbulence is super-sonic (e.g., Solomon et al., 1987). This results in a strong degeneracy between the thermal and turbulent properties of the medium, though these sources may be distinguished by comparing line widths near strong turbulent driving sources or changes in the average galactic environment (e.g., Tamburro

[14]The sources is at least minimally-resolved at the resolution of the observations.

et al., 2009).

One expected source of turbulent driving is stellar feedback from massive star formation, especially supernovae (Mac Low & Klessen, 2004b). This feedback produces distinct HI "holes" that, combined with the feedback source's energy and the energy required to produce the whole, can be used to estimate the turbulent energy injection into the ISM (Heiles, 1979; Bagetakos et al., 2011) and the efficiency predicted by ISM models (McKee & Ostriker, 1977; Ostriker et al., 2010).

While the latter sources affecting line shape mostly account for the bright HI emission in galaxies, the HI line also traces fainter structure related to inflow onto and outflow from the galaxy disc. With sufficient resolution, inflow and outflow structures are resolved into high-velocity clouds (HVCs), named for their discovery near the Milky Way at velocities significantly different than what would be predicted from galactic rotation (Dickey & Lockman, 1990). Recent observations show a population of clouds around most Local Group galaxies (e.g., Thilker et al., 2004; Lockman et al., 2012; Keenan et al., 2016), including clouds that appear to have been ejected from the SMC (McClure-Griffiths et al., 2018).

Without sufficient resolution, a low level of HI is often detected in nearby galaxies (Fraternali et al., 2002a), which includes a combination of unresolved HVCs (Stanimirović et al., 2006), outflows driven by stellar or AGN feedback (Barnabè et al., 2006; Sancisi et al., 2008), and the interface with a galaxy's gaseous halo and CGM (Putman et al., 2009; Heald et al., 2011; Putman et al., 2012). These components together create low-level "wings" in an HI spectrum with equivalent Gaussian widths far larger than the bright HI emission from the galactic disc.

In extragalactic HI observations, these different effects are mixed together to produce the HI spectrum. Spectra from lower resolution observations ($\sim$ 100s pc) average over many different effects but provide a useful measure of the average HI kinematics in galaxies. To be able to separate different effects at work, however, HI observations below typical disc scale heights ($\sim$ 100 pc;

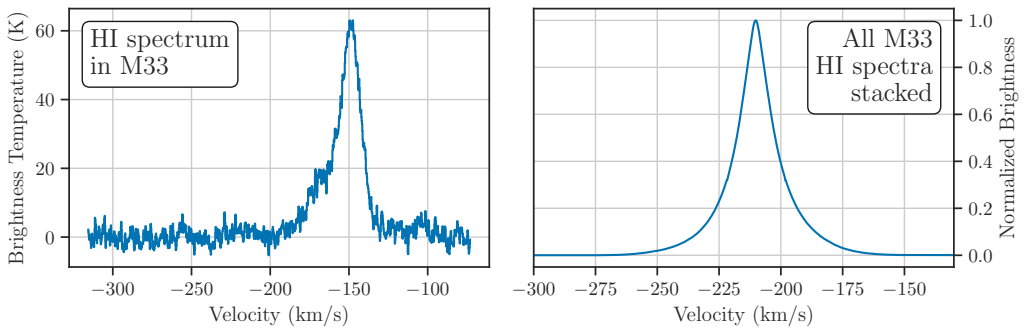

Figure 1.4: Example HI spectra in M33 (see Chapter 2). The left shows an individual spectrum while the right shows the stacked spectrum. The individual spectrum has bright component with a blue-shifted shoulder and faint emission on either side. The stacked spectrum shows the "universal" stack shape (Petric & Rupen, 2007) with a narrow peak and wide line wings which results from the stacking process (Koch et al., 2018c, Chapter 3).

Ferrière, 2001) are required.

Given the limitations of HI observations in nearby galaxies, different approaches have been used to interpret the HI line shape. While most of these methods are valid models to describe the observations, their interpretation of the atomic ISM properties varies widely.

The simplest HI kinematic measurement uses the second moment of an HI spectrum to estimate the line width. If the spectrum is a single Gaussian, the line width from the second moment is the Gaussian rms line width. Tamburro et al. (2009) explore this measurement in several nearby galaxies and find typical HI line widths 8–12 km s^{-1}. Assuming the volume-filling WNM dominates the HI emission, these line widths exceed the expected thermal line width of 6 km s^{-1} for 6000 K gas, thereby requiring additional turbulent broadening from turbulence. Tamburro et al. (2009) compare the turbulent energy dissipation rate from the measured line widths and find that energy released by supernovae are sufficient to produce the observations. In the outer regions of galaxies, where the supernovae rate is low, they suggest that turbulence could be maintained by the magneto-rotational instability (MRI; Mac Low & Klessen, 2004a). The second moment is appropriate when limited sensitivity

makes the line shape uncertain, though, it likely overestimates the line width (Mogotsi et al., 2016; Koch et al., 2018c), suggesting other methods should be used where the data allow.

To overcome limitation in sensitivity, many studies increase the signal-to-noise ratio (S/N) by spectral stacking, where a large number of spectra are aligned to a commonly-defined line centre. Once aligned, these spectra add coherently to boost the sensitivity. Several HI extragalactic studies use this method, and often find a similar stacked profile shape combining a central near-Gaussian peak with line wings enhanced relative to a Gaussian (Young & Lo, 1996; Braun, 1997; Petric & Rupen, 2007; Ianjamasimanana et al., 2012; Stilp et al., 2013a). Petric & Rupen (2007) suggest that this "universal" shape demonstrates similar HI properties in different galaxies, though Koch et al. (2018c, Chapter 3) show that this shape is due to the stacking method itself.

The interpretation of this universal HI stacked line shape has produced two significantly different interpretations depending on how the profile is modelled. The first model uses two Gaussians to fit the stacked spectrum. This combination of Gaussians tends to produce a bright and narrow component with widths of 3–6 km s^{-1}, and a wide and faint component with widths of 9–25 km s^{-1} (Young & Lo, 1996; Young et al., 2003; Ianjamasimanana et al., 2012; Stilp et al., 2013a; Utomo et al., 2019a). Braun (1997) similarly interpret stacked profiles as a narrow and wide component, though the wide component is modelled as a Lorentzian profile. Several of these studies interpret the narrow and wide components with the CNM and WNM, respectively (Young & Lo, 1996; Young et al., 2003; Ianjamasimanana et al., 2012), where excess from the thermal line widths is assumed to be from turbulence. Ianjamasimanana et al. (2012) use the fraction of each component's integral to estimate the CNM-to-WNM ratio in nearby galaxies and find values ranging from 0.14–0.99.

The CNM and WNM fractions inferred by Ianjamasimanana et al. (2012) vary drastically between different galaxies. Stilp et al. (2013a) demonstrate that this is partially due to the covariance between the wide and narrow com-

ponents and that the area under each component is sensitive to small changes in the profiles, making interpretations of these components to the CNM and WNM uncertain. Instead, Stilp et al. (2013a) introduce a non-parametric line width and line wing definitions that does not rely on an explicit CNM/WNM interpretation and does not contain correlated parameters Using these non-parametric measures, they find equivalent Gaussian widths range from 6–10 km s^{-1}. From this and measures of the line wing strength, they estimate the turbulent kinetic energy as the primary component producing the stacked line profiles. Similar to Tamburro et al. (2009), they find that turbulence in the inner regions of galaxies can be driven by supernovae, though they find weak correlations between radial profiles of the star formation rate versus the line width and wing statistics.

Finally, HI spectra are also directly fit, matching many Milky Way studies. Unlike the Milky Way studies, however, most extragalactic HI models are limited to a single Gaussian model (Boulanger & Viallefond, 1992; Combes & Becquaert, 1997; Mogotsi et al., 2016) or two-Gaussian model (Young & Lo, 1996; de Blok & Walter, 2006; Warren et al., 2012). These two-Gaussian models often find a narrow components imposed on top of a wide component. Some studies also use a Gauss-Hermite polynomial fit, which can model skewed profiles (Young et al., 2003; Warren et al., 2012; Stilp et al., 2013a). However, Young et al. (2003) note that it is unclear how to map from this model to the physical gas conditions and the aforementioned studies only use this method to improve estimates of the line centre.

Compared to the two-Gaussian CNM/WNM interpretation (Ianjamasimanana et al., 2012), far fewer fits to individual spectra require the two-Gaussian model (Young & Lo, 1996; Warren et al., 2012), apart from galaxies with a high inclination where lines-of-sight intersect ~ 2 bright components through the disc (de Blok & Walter, 2006). Because of this, Warren et al. (2012) conclude that cold HI, as seen from emission, is not significant in the low-mass galaxies included in their sample.

For Local Group galaxies, the spectral resolution and sensitivity of HI ob-

servations is sufficient to recover more spectral complexity at high (< 100 pc) resolution, where many spectra cannot clearly be modelled by just one or two Gaussians, and the spectral shapes become more similar to Milky Way HI observations[15]. From this complexity, Braun et al. (2009) suggest that some HI spectra in M31 have a distinct flattened-top, which they posit results from opaque cold HI. Braun (2012) extend this identification to the LMC and M33. If this is the case, it provides a powerful method for measuring the HI opacity since it relies only on HI emission, not absorption (e.g., Dickey et al., 2003). Braun et al. (2009) show that this flat-top shape can be reproduced by a simple model where a single gas component, broadened by non-thermal (e.g., turbulent) motion, is optically-thick over some column density (Equation 1.10; see also Rohlfs et al., 1972). From this modelling, they infer that the optically-thin approximation underestimates the atomic gas mass by $\sim 35\%$ across M31, M33, and the LMC. However, Rohlfs et al. (1972) note that there is ambiguity between a flat-top opaque model and spectra where two Gaussian components are blended together.

These diverging interpretations produce large differences in inferred physical properties of the atomic ISM in nearby galaxies. Because of this, it is difficult to connect with the Milky Way HI properties observed on small scales and test predictions of how the atomic ISM varies with galactic environment (Bialy & Sternberg, 2019).

Chapters 3 & 4 explore the HI line shape in detail using our new HI VLA observations of M31 and M33. In particular, I show that the high spectral resolution of these new data unveil spectral complexity previously only seen in the Milky Way and Magellanic Clouds, and that the kinematic methods used in most extragalactic HI studies provide a coarse, and sometimes biased, measurement that does not capture this complexity.

[15]Though we do not expect these to be entirely similar, even on matched scales, due to the different observed geometries (i.e., within verse external to the galaxy).

1.2.2 From HI to H_2: kinematics of the neutral ISM in molecular cloud evolution

Molecular clouds are the birthplace of stars in the nearby universe (e.g., Bigiel et al., 2008, 2010; Schruba et al., 2011). Observations over the last decade have firmly established the strong relationship between recent star formation and molecular gas (the Kennicutt-Schmidt "Law"; Schmidt, 1959; Kennicutt, 1998) and the weak correlation with atomic gas (Bigiel et al., 2008; Kennicutt & Evans, 2012). These properties of molecular clouds are well-studied in the Milky Way (e.g., Larson, 1981; Solomon et al., 1987; Colombo et al., 2019) and nearby galaxies (e.g., Wilson & Scoville, 1990; Rosolowsky et al., 2003; Nieten et al., 2006; Leroy et al., 2009; Gratier et al., 2010; Corbelli et al., 2017), however, the role of HI in molecular cloud evolution remains poorly explored in observations (Dobbs et al., 2014).

In most of these studies, H_2 is traced by the low rotational transitions of carbon monoxide (CO). The key difference is that H_2, as a homonuclear diatomic molecule, has no permanent dipole moment, and the $T \sim 10$ K in molecular clouds is insufficient to excite the allowed vibrational transitions (Draine, 2011). CO, however, has a permanent dipole moment with low-energy rotational transitions which are easily excited in molecular cloud conditions. Combined with the relatively high abundance of C and O, this makes CO the most abundant molecule behind H_2 in molecular clouds[16] and bright enough to be observed throughout the universe. Physically, the bright CO lines provide the key cooling channel in molecular clouds and balance heating from cosmic rays (e.g., Ferrière, 2001; Krumholz, 2014).

Molecular clouds have three stages in their evolution: (1) formation and no star formation, (2) early star formation still embedded within the cloud, and (3) on-going and exposed star formation, whose feedback destroys the cloud (Chevance et al., 2020a). These stages must occur over within a cloud lifetime ($\sim$ 10–50 Myr; Kawamura et al., 2009; Corbelli & Schneider, 1997;

[16]The second most abundant atom is He, which is a noble gas and therefore is not very reactive nor directly observable.

Kruijssen et al., 2019; Chevance et al., 2020b), implying that cloud formation and the onset of star formation is rapid. Additionally, molecular clouds may continuously grow by accreting from the surrounding HI throughout their lives and influence the star formation efficiency, until feedback is sufficiently strong to destroy the cloud (Lee et al., 2016).

Through the molecular cloud evolution stages, HI is either an active or passive component to the processes within the cloud. First, H_2 must form from HI, most likely from the CNM whose properties are closest to the molecular medium. Where H_2 forms is set by the local volume density, where UV shielding allows for efficient H_2 formation (Krumholz et al., 2009; Krumholz, 2013; Sternberg et al., 2014; Bialy et al., 2017), and the mid-plane disc pressure set by the galactic environment (Blitz & Rosolowsky, 2006; Ostriker et al., 2010), but the mechanism that gathers the CNM to increase the density remains unclear from observations (Dobbs et al., 2014). Once the cloud has formed, an HI envelope should persist around the cloud following predictions of shielding-based H_2 formation models (Krumholz et al., 2009; Sternberg et al., 2014), which affects the cloud's gravitational boundedness (Elmegreen & Elmegreen, 1987). After the onset of star formation, stellar feedback has a profound effect on the HI envelope while the cloud is destroyed due to its lower density than the H_2. This expectation follows from various numerical simulations that show stellar feedback strongly affecting lower density material in clouds (Dale et al., 2014; Seifried et al., 2018), where over-pressured regions from heated gas preferentially leak via low column density paths (Dale, 2017).

While the lack of correlation between HI and star formation surface density is well-studied (Bigiel et al., 2008), the joint kinematics of HI and CO (tracing H_2) remains poorly explored, yet has the potential to solve key unknowns about the role of HI in molecular clouds from formation to destruction. In the Milky Way, several works note a clear layered appearance of molecular clouds (traced by CO) surrounded by bright HI (Grabelsky et al., 1987; Elmegreen & Elmegreen, 1987). Due to complexities in the HI line shape (§1.2.1), this relationship is primarily based on visual inspection. More recent studies find

HI self-absorption towards molecular clouds with similar kinematic properties between HI and H_2 (Li & Goldsmith, 2003; Goldsmith & Li, 2005). These studies hint of similar kinematics between the HI and CO on sub-cloud scales (e.g., line width; Krčo et al., 2008; Krčo & Goldsmith, 2010; Wang et al., 2020b), with variations that may relate to the formation or destruction of the cloud (Soler et al., 2019).

Kinematic comparisons in extragalactic studies are limited compared to Milky Way studies due to the coarse resolution in 21-cm observations beyond the Local Group. On > 100 pc scales, where individual molecular clouds are unresolved, there is a strong correlation between the HI and CO velocities at peak intensity (Schruba et al., 2011), suggesting both tracers follow a similar rotation curve in galaxies. By fitting individual HI and CO spectra with single Gaussians in nearby galaxies, Mogotsi et al. (2016) find that CO line widths ($\sigma_{\rm CO} = 7.3 \pm 1.7$ km s^{-1}) are roughly 60% of HI line widths ($\sigma_{\rm HI} = 11.7 \pm 2.3$ km s^{-1}) at matched resolution from $\sim$ 200–700 pc.

Extragalactic studies of HI and CO in individual spectra on smaller scales are otherwise limited to Local Group galaxies. In M33, Imara et al. (2011) compare the HI and CO centroid velocity gradients across 45 molecular clouds on $\sim$ 20 pc scales. They do not find a strong correlation in the gradients across these clouds, though the HI spectra have line structure that may not be captured from the centroid velocity (e.g., their Figures 4 & 5; also see Gratier et al., 2010). In the LMC, Wong et al. (2009) and Fukui et al. (2009) find that the correlation of velocities at peak intensity in HI and CO continues to hold on 40 pc scales, and they further find a suggestive correlation between the HI and CO line widths. They find that CO line widths are just 30% of the HI line widths, far smaller than the ratio from Mogotsi et al. (2016). A second clear result from these LMC studies is the weak relationships between CO detection with peak HI temperature and HI column density. Given the HI line complexity and its high detection fraction relative to CO, these weak relationships suggest that a significant amount of HI emission arises from atomic gas unassociated with the molecular gas. To distinguish HI emission from the

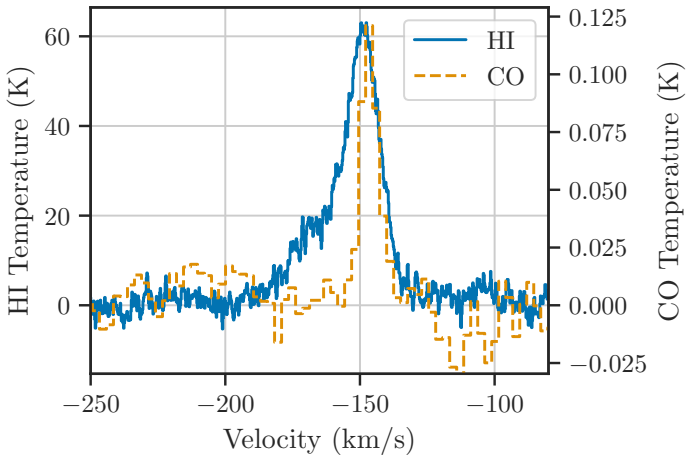

Figure 1.5: HI (Chapter 2) and CO (tracing H_2; Druard et al., 2014) toward one line-of-sight in M33. The CO emission has a single component, with a velocity at peak intensity similar to the HI. The HI has an additional blue-shifted shoulder, in addition to faint emission. These spectra demonstrate the need to separate HI that is associated with H_2 to distinguish the role of HI in molecular clouds.

molecular cloud versus the surrounding environment, molecular clouds should be moderately-resolved.

Due to the low beam filling factor of CO at > 100 pc resolution, CO detections in dwarf galaxies and the outer regions spiral galaxies often rely on spectral stacking (Combes & Becquaert, 1997; Schruba et al., 2011; Caldu-Primo et al., 2013; Caldú-Primo & Schruba, 2016a), following the technique described in §1.2.1. These studies find that HI and CO stacked line widths are *consistent*, unlike the smaller CO line widths from individual spectral fitting above. While some line broadening is expected from the stacking method (Chapter 3; Koch et al., 2018c), the consistent widths are unlikely to be result from systematics alone. Instead, Combes & Becquaert (1997) suggest that the increased sensitivity from stacking allows diffuse CO to be detected. This diffuse CO traces a similar kinematic component to the HI, thereby tracing a similar volume and disc scale height[17] larger than that of the bright CO that

[17]However, CO from low-mass molecular clouds are likely to contribute some fraction of the missing CO flux (Rosolowsky et al., 2003).

arises from GMCs. This offers an explanation for the different CO line widths measured in studies that use the same data but apply stacking (Caldu-Primo et al., 2013) versus fitting individual spectra (Mogotsi et al., 2016).

Diffuse CO has been measured or inferred in different ways, and due to its faintness, there remains many unanswered questions about its properties. Garcia-Burillo et al. (1992) detect CO $\sim$ 1 kpc from the plane of the disk in the edge-on galaxy NGC 891, providing direct evidence for a "molecular halo." Indirectly, Pety et al. (2013) find that $\sim$ 50% of the CO flux in M51 based on single-dish data is filtered out in interferometric data, suggesting the missing flux arises from large-scales. Caldú-Primo et al. (2015) and Caldú-Primo & Schruba (2016a) identify a wide CO velocity component from line stacking is only recovered in single-dish observations on scales $>$ 500 pc. Similarly, Dame & Thaddeus (1994) find excess CO line wings in Milky Way observations and Roman-Duval et al. (2016) detect diffuse CO extended off the Galactic plane, which could account for $\sim$ 25% of the total CO emission.

The presence of diffuse CO complicates the separation of HI and CO associated with molecular clouds. Though faint, if the diffuse CO broadly traces HI, it is necessary to marginally resolve molecular clouds to cleanly separate the cloud kinematics in HI and CO from the diffuse background.

These previous studies highlight suggestive correlations linking the HI and CO kinematics and that diffuse HI and CO emission is a source of confusion when measuring molecular cloud kinematics. Accounting for this confusion from diffuse HI and CO emission *should* provide observational constraints on key unknowns about cloud formation and destruction to test recent predictions (e.g., Jeffreson & Kruijssen, 2018).

Chapter 5 explores the relations between HI and CO kinematics in M33 on 80 pc scales. In this work, I account for line stacking systematics to demonstrate that M33 lacks the diffuse molecular gas suggested in more massive spirals. Further, I show a strong correlation between HI and CO line widths that is only apparent when accounting for HI not associated with CO emission.

1.2.3 Observational signatures of turbulent driving on galactic scales

Turbulence plays a key role in modern star formation theories (McKee & Ostriker, 2007) and is an integral part of ISM dynamics (Elmegreen & Scalo, 2004). Properties of interstellar turbulence[18] are measured using statistical techniques such as the power spectrum. In particular, the energy power spectrum $\mathbf{E}(\mathbf{k})$ can be compared to different turbulence prescriptions (e.g., Kolmogorov, 1941)

Figure 1.6 shows an idealized example of $\mathbf{E}(\mathbf{k})$. On large scales, some source injects energy into the medium driving turbulent motion. Below this scale, the energy power spectrum has a self-similar structure where (in this case) energy is conserved but is transferred to successively smaller scales in turbulent eddies. This is the inertial range, and the transfer of energy to smaller scales gives a power-law shape to $\mathbf{E}(\mathbf{k})$. On small scales turbulence is dampened where viscosity becomes important. The energy is lost from the turbulent motion, and $\mathbf{E}(\mathbf{k})$ drops steeply.

There are three parameters of interest in $\mathbf{E}(\mathbf{k})$: (i) the driving scale of turbulence, (ii) the dissipation scale, and (iii) the index of the power law in the inertial range. The power spectrum index is varies with the type of turbulence for incompressible (e.g., Kolmogorov, 1941) and compressible gas (e.g., Burgers, 1948; Fleck, 1996; Galtier & Banerjee, 2011; Federrath, 2013).

Recovering $\mathbf{E}(\mathbf{k})$ in observations is difficult since, for kinetic energy, $\mathbf{E}(\mathbf{k})$ depends on the density ($\rho(\mathbf{k})$) and velocity ($\mathbf{v}(\mathbf{k})$) fields of the medium: $\mathbf{E}(\mathbf{k}) \propto \rho(\mathbf{k})\mathbf{v}(\mathbf{k})^2$. Spectral lines, like the 21-cm HI line, provide constraints on the 3D turbulent velocity and density fluctuations using HI spectral-line cubes with the projected sky dimensions and the line-of-sight velocity (Lazarian & Pogosyan, 2000, 2006). However, tracers of column density alone also constrain the 3D density turbulent field (Elmegreen & Scalo, 2004). Though this removes the velocity information, the column density to volume density connection

[18]The power spectrum is used more widely for turbulence studies, including for terrestrial sources.

enables a wider range of tracers that can be used to measure ISM turbulence. In particular, dust emission and extinction is extremely useful for this purpose as observations combine high-sensitivity with a large dynamic range, allowing faint emission to be distinguished below the typical sensitivity of spectral-line observations.

In the remainder of the section, I refer to the column density power spectrum as the "power spectrum." Though not fully descriptive, this use is fairly ubituitous in the literature. Due to its connection to the underlying turbulent density field, the column density power spectrum is closely related to $\mathbf{E}(\mathbf{k})$, though is not constraining without velocity information.

While the power spectrum index is measured in many studies (Elmegreen & Scalo, 2004), the source of turbulent driving remains unclear. Because of the central role of turbulence in star formation, however, the source of turbulence has important consequences for the onset of star formation in nearby galaxies (e.g., McKee & Ostriker, 2007; Krumholz, 2014). Turbulent driving is typically assumed to be dominated by a driving mechanism operating on scales near or larger than the disc scale height, for which there are two likely candidates: gravitational instability and stellar feedback (mostly SNe; Krumholz & Burkhart, 2016). Further, the decay timescale of turbulence is short ($\sim$ 10 Myr on molecular cloud scales; Mac Low & Klessen, 2004b), suggesting that this mechanism must act near-continuously and may therefore be detectable in observations. However, power spectra from ISM observations have power-law distributions over a range of scales (e.g., Miesch & Bally, 1994; Stutzki et al., 1998) with no obvious turn-over at large scales. A turn-over on large scales (Figure 1.6) in the power spectrum would distinguish the dominant driving scale and constrain the mechanism responsible. This measurement, however, remains elusive.

A second property predicted from theoretical descriptions of turbulence is the relation between the number of free spatial dimensions in a system and its turbulent power-spectrum index. For example, a system that transitions from three- to two-dimensions is expected to cause a power-spectrum "break,"

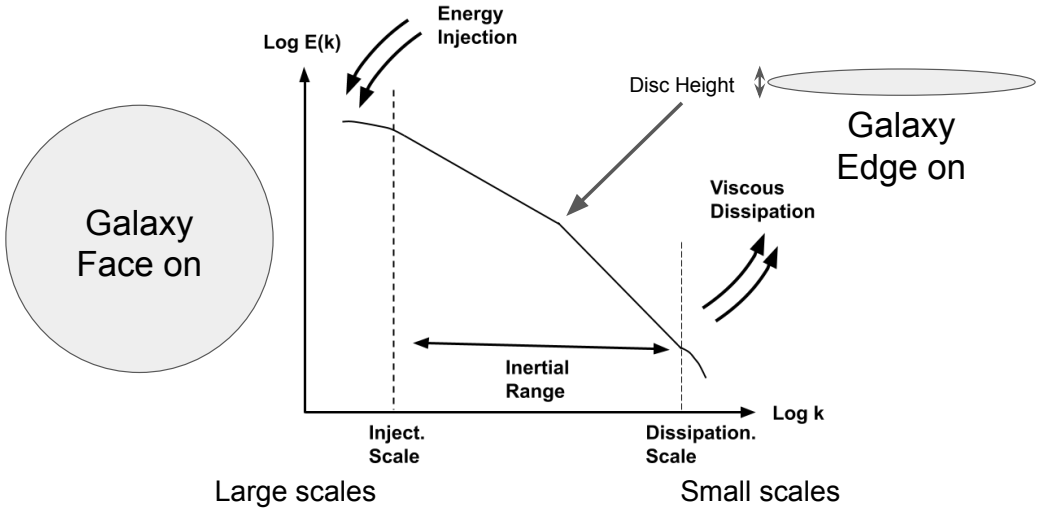

Figure 1.6: Schematic of a turbulent energy cascade using the energy power spectrum $\mathbf{E}(\mathbf{k})$. This schematic highlights the expected transition where motions become primarily three-dimensional below the galaxy disc scale height, altering the power law index of $\mathbf{E}(\mathbf{k})$. Above the disc scale height, motions are primarily confined to two dimensions in the galaxy plane, producing a power spectrum index that is shallower by -1.

where the index changes by -1 (e.g., Lazarian & Pogosyan, 2000; Elmegreen & Scalo, 2004). This is potentially important for turbulence statistics of nearby galaxy observations that resolve below the ~ 100 pc disc scale height (Ferrière, 2001), as the scale of the break would measure the disc scale height, with two-dimensional motion on larger scales and three-dimensional motion on smaller scales. Figure 1.6 shows a schematic of this prediction for a turbulent energy power spectrum. For face-on galaxies, the scale height is difficult to constrain in observations. Whether the break is measured, and if the turbulent properties are consistent with a change in the number of dimensions across the break, can further constrain potential driving mechanisms for turbulence.

Recent numerical simulations from molecular cloud to galaxy scales find that signatures of turbulent driving *should* be measurable in observations. On molecular cloud scales, Boyden et al. (2016) demonstrate that several common turbulence statistics are sensitive to stellar winds in synthetic observations.

Galaxy-scale simulations also find changes in turbulence statistics with stellar feedback, particularly in recent simulations which resolve down to $\sim$ 10 pc scales (Grisdale et al., 2017; Garrison-Kimmel et al., 2019). Several simulations show a break in the power spectrum, as predicted for a change in the number of free spatial dimensions. However, different levels of stellar feedback change the scale of the break (Bournaud et al., 2010; Pilkington et al., 2011; Combes et al., 2012; Grisdale et al., 2017), as does strong spiral arm structure (Renaud et al., 2013).

Milky Way studies of turbulence in the neutral ISM demonstrate a wide range of properties. Within molecular clouds ($<$ 20 pc), power spectra are well-characterized by power law relations and do not provide strong constraints on turbulent driving. For example, in the Perseus molecular clouds, stellar feedback from within the cloud is likely able to drive turbulence at the observed level (Padoan et al., 2009; Arce et al., 2011; Pingel et al., 2018) despite the lack of a characteristic break scale in the power spectrum. However, other turbulence statistics, such as the probability distribution function of column density, do show properties expected for turbulent driving on sub-cloud scales (Bialy et al., 2017).

Several HI studies in the Milky Way, both towards molecular clouds (e.g., Pingel et al., 2018) and in more diffuse regions (e.g., Martin et al., 2015; Blagrave et al., 2017), do not find a clear driving scale from the power spectrum. The key difficulty in distinguishing a driving scale from Milky Way HI is the line-of-sight confusion, with thermodynamic and radiative effects altering HI in the Galactic plane and distance ambiguities in the high-latitude HI emission. Both of these effects may act to "hide" a driving scale by averaging over different spatial scales. Alternatively, turbulent driving may occur on larger scales, making this measurement easier with an external view of nearby galaxies.

Most studies of turbulence in nearby galaxies find either a power spectrum that follow a single power law (e.g., Stanimirovic et al., 2000; Dutta et al., 2013; Zhang et al., 2012) or a broken power law (e.g., Elmegreen et al., 2001;

Dutta et al., 2009a; Combes et al., 2012). In the latter category, the break is interpreted as tracing the disc scale height, as explained above, however the change in the power law index across the break is often much larger than 1 expected solely from the change in spatial dimensions. Further, there is little consensus in the recovered power spectrum indices, even when comparing results from a common tracer (e.g., Dutta et al., 2013). While this range may result from a differences in the ISM environments across different galaxies, how the power spectrum is calculated also changes across these studies, and the power spectrum properties are known to vary with the technique used (e.g., Dutta & Bharadwaj, 2013).

The strongest turbulence constraints of an external galaxy are studies of the SMC due to its proximity to the Milky Way. For example, the 21-cm HI observations from Stanimirovic et al. (1999a) have a physical resolution of $\sim$ 15–20 pc, well below the disc scale height. Stanimirovic et al. (1999a, 2000) find similar HI and dust power spectra in the SMC, with no strong evidence for a power spectrum break. Further work by Stanimirović & Lazarian (2001) use the same 21-cm HI observations with the velocity channel analysis (VCA) technique from Lazarian & Pogosyan (2000) to separate the velocity and density power spectra. They find shallower power spectrum indices than would be expected for a highly super-sonic shock-dominated ISM. Burkhart et al. (2010) explore turbulence in the SMC with several turbulence statistics, including the bispectrum—a three-point statistics that, unlike the power spectrum, retains phase information[19]. They find breaks in the bispectrum phase structure at $\sim$ 160 pc scales, possibly corresponding to the scale of feedback-driven shells in the SMC (e.g., Stanimirovic et al., 1999a). Finally, Chepurnov et al. (2015) use the velocity coordinate spectrum (VCS; Lazarian & Pogosyan, 2006), which uses spatially-averaged power spectrum to map changes in the spectral structure, to constrain the driving scale of turbulence in the SMC. They find a driving scale around $\sim$ 2 kpc, essentially the size of the SMC, and suggest that this may reflect driving from interactions with the LMC and the

[19]See Koch et al. (2019a) for an overview of these turbulence statistics.

Milky Way halo. If this is the case, the dominant turbulent driving scale may indeed change significantly depending on the galaxy environment.

One restriction of many turbulence studies in nearby galaxies is how galactic structure not dominated by turbulence, such as spiral arms, affects turbulence statistics. For example, in Chapter 5 (Koch et al., 2019b), I show that the CO spatial power spectrum in M33 has an excess of power on ~ 2 kpc scales (about the disc scale length Druard et al., 2014) due to the clustering of molecular clouds in the inner disc. Other affects, such as feedback-driven shells, clearly alter the power spectrum in numerical simulations and the scale of the power spectrum break (Grisdale et al., 2017). These variations strongly suggest that informative studies of large-scale turbulent driving may require tracing how turbulence varies *across* galaxies, where observations permit.

Similar to the studies listed above, the SMC is a prime target to study spatial variations in turbulence. Nestingen-Palm et al. (2017) split the SMC into an inner and outer region based on the star formation rate density. They find no difference in the turbulent properties measured by the spatial power spectrum and VCA of HI within these regions. Given the hints of a ~ 160 pc scale from Burkhart et al. (2010), these regions may simply be too large and average over any spatial variations on the scale of feedback-driven shells, or HI holes.

Szotkowski et al. (2019) explore this effect in greater detail by introducing a "rolling" structure function[20] method, which measures an equivalent power spectrum within a $\sim$ kpc sliding window across the galaxy. They similarly find no difference in the HI structure function of the SMC over different regions, which they suggest could be due to the large depth along the line-of-sight through the SMC (e.g., Di Teodoro et al., 2019). However, Szotkowski et al. (2019) *do* find significant variations across the LMC. In particular, they find a characteristic break scale near regions with strong stellar feedback (e.g., 30 Doradus), along with variations in the power spectrum index. Observations of

[20]The structure function is the real-space equivalent to the power spectrum for the range of power spectrum indices relevant in the ISM (e.g., Simonetti et al., 1984).

most Local Group galaxies have sufficient resolution to perform similar analyses.

Spatially-resolved measurements of how turbulence varies across galaxies are a key constraint on how and where turbulence is driven. In particular, combining turbulence statistics across tracers of different ISM densities measures how different phases of the ISM are affected by driving mechanisms such as stellar feedback (e.g., do HI, dust, and CO all show a characteristic break, like those found by Szotkowski et al., 2019, ?). These measurements bridge between studies of the large-scale turbulent energy injection rate with assumed mechanisms (Chapter 3; Tamburro et al., 2009; Stilp et al., 2013a; Utomo et al., 2019a) to test recent star formation theories (Krumholz & Burkhart, 2016; Krumholz et al., 2018). Further, it informs assumptions made about efficient molecular cloud destruction, which plays an important role in observational estimates of molecular cloud lifetimes (Kawamura et al., 2009; Corbelli et al., 2017; Kruijssen et al., 2019; Chevance et al., 2020b).

Chapter 6 explores the question of turbulent driving across the Local Group galaxies using observations of dust, HI, and CO. I show that power spectrum breaks in M31, M33, LMC, and SMC are nearly all due to the influence of the telescope's point-spread-function (PSF), though the extremely bright 30 Doradus, a giant HII region, in the LMC produces a break in 24 μm emission. I also show that the power spectrum index of the dust surface density varies across the LMC and SMC, though we do not find power spectrum breaks like those in the HI from Szotkowski et al. (2019).

1.3 Outline & questions explored in this thesis

The remainder of the thesis is arranged into one technical chapter that introduces a new Local Group 21-cm HI survey from the Karl G. Jansky Very Large Array (Chapter 2), followed by four science chapters (Chapters 3–6). These chapters present my work exploring three questions about the neutral ISM which link galactic to molecular cloud scales:

1. How do models of the 21-cm HI line shape affect our interpretation of the neutral ISM? (Chapters 3 & 4)

2. Are atomic (HI) and molecular (H_2) ISM kinematics coupled on 80 pc scales? (Chapter 5)

3. Is the source of large-scale turbulent driving imprinted on the ISM density structure? (Chapter 6)

Finally, Chapter 7 concludes with a summary of my findings and future work enabled by our new 21-cm HI observations.

Chapter 2

A Modern 21-cm HI Survey of the Local Group

My work in this thesis relies on new HI observations taken with the Karl G. Jansky Very Large Array (VLA) in New Mexico. In this chapter, I briefly describe radio interferometry (§2.1) and present an overview of historical 21-cm HI observations in M31 and M33 (§2.2) to demonstrate the remarkable improvements in instrumentation in the past 7 decades. I then summarize the M31 and M33 VLA observations included in the campaign that I use for this thesis (§2.3). Within this section, I describe observational setup, the reduction and calibration used, and the technique I developed for handling the imaging of massive inteferometric data. Imaging large interferometric data sets remains a challenge for current software and techniques, yet is vital for handling the data from the next generation of radio interferometers.

2.1 A basic overview of radio interferometry

Radio interferometers are collections of antennas that work together to produce high-resolution images at radio and mm-wavelengths[1] The antennas observe the same region of the sky, and by precisely knowing the distance between each antenna, software can be used to measure the light delay time for each

[1]This technique can, in principle, be used at other wavelengths. However, it relies on knowing the distance between antennas to a fraction of the wavelength observed ($\sim \lambda/20$ for wavelength λ). This becomes very difficult at, for example, optical wavelengths (~ 700 nm) but has been achieved. One example is the Very Large Telescope Interferometer.

antenna. By correcting for the delay time, the observed light can be interfered coherently.

Measuring the coherent interference is the key to the aperture synthesis technique that all modern interferometers use. Each pair of antennas samples the interference pattern produced by the observed sky, or equivalently one measurement of the Fourier transform of the observed sky. This is referred to as the uv-plane, where u and v are the spatial frequency axes in the Fourier plane. Formally, each measurement is described as:

$$V(u,v) = \int \int T(l,m) \exp\left(-2\pi i(ul+vm)\right) \, dldm, \tag{2.1}$$

where l and m are the spatial angles on the sky and $T(l,m)$ is the image on the sky (Thompson et al., 2017). By combining samples from all pairs of antennas in the array and over time, frequency, and polarization, $V(u,v)$ can be used to reconstruct the image $T(l,m)$.

Because these samples come from *pairs* of antennas, the image recovered from the interferometer has a resolution equivalent to the longest separation between the baselines. This is the "synthesis" in aperture synthesis: creating an effective image as if a single-dish the diameter of the whole array had been used. This technique is immensely powerful because it removes the extreme engineering issues with constructing radio dishes $\gg 100$ m in diameter. For comparison, the Green Bank and Effelsburg telescopes are the largest steerable telescopes in the world and have diameters of ~ 100 m. At the 21-cm HI line, this corresponds to a resolution of $\sim 9'$. The observations presented in this Chapter recover a resolution of $16''$, 34 times finer than the single-dish resolution. The science in this thesis would not be possible without an interferometer.

The recovery of images from the interferometer requires two steps. First, the inverse Fourier transform of $V(u,v)$ is used to construct the sky image $T(l,m)$:

$$T(l,m) = \int \int V(u,v) \exp\left(2\pi i(ul+vm)\right) \, dudv. \tag{2.2}$$

Practically, this is done by gridding $V(u,v)$ so the Fast Fourier transform

can be used[2] However, the resulting image combines two effects: the actual sky image, and the interferometer's response to a point-source ($s(l, m)$; point-spread function; PSF). Thus the image reconstruction is a corrupted view of the sky, typically called the "dirty map," described as:

$$T^{\mathrm{D}}(l, m) = T(l, m) * s(l, m), \tag{2.3}$$

where $*$ denotes convolution (Thompson et al., 2017). The PSF $s(l, m)$ depends on the distribution of $s(l, m)$ and the array configuration of the inteferometer. For example, the VLA is arranged into a "Y"-shape, and its Fourier transform has a distinct 6-armed star pattern. Bright emission in the dirty map is noticeably affected by this star pattern, and without removing the effect of $s(l, m)$, faint emission cannot be distingiuished from instrumental effects.

The dirty map $T_{\mathrm{D}}(l, m)$ introduces the second step of removing the effect of $s(l, m)$ to recover the true sky image $T(l, m)$. This procedure is formally an inverse problem, which arises in various scientific fields[3]. In radio interferometry, this is referred to as deconvolution.

There are various deconvolution methods, however, most radio interferometry studies use variants of the CLEAN[4] algorithm (Högbom, 1974). The original CLEAN algorithm assumes an image can be modelled as a set of point sources. The emission in $T_{\mathrm{D}}(l, m)$ is iteratively removed starting from the pixel with the brightest feature in the map. The subtracted emission is $g \times s(l_0, m_0)$, where l_0, m_0 is the brightest pixel and g is a small gain factor (typically ~ 0.05). This emission is then added to a model image, $M(l, m)$, at l_0, m_0 *without the PSF*. This process is continued until, ideally, only noise remains in the original image, often defined by when the brightest pixel is $T_{\mathrm{D}}(l_0, m_0) \sim 2\sigma_{\mathrm{rms}}$. To produce $T(l, m)$, or the best approximation of the sky given the data, the model $M(l, m)$ is convolved with the idealized telescope (usually a two-dimensional Gaussian) and the remaining noise residuals

[2]A direct Fourier transform does not require gridding but is prohibitively slow.

[3]For example, Tikhonov regularization which forms of the basis of the ridge regression in linear regression (Tikhonov et al., 1995).

[4]Note that CLEAN is not an acronym. This is the historic use of the algorithm name.

in $T_{\rm D}(l,m)$ are added.

One limitation of radio interferometry is that the interferometer is only sensitive to scales between the minimum and maximum baselines. For the minimum baseline ($b_{\rm min}$), this corresponds to the largest angular scale of $\sim \lambda/b_{\rm min}$ that the inteferometer measures. The sky on larger scales is filtered out. Single-dish observations have a low resolution at cm-wavelengths, but will measure the total flux from the sky. To account for missing flux from the interferometer, interferometric and single-dish observations can be combined through the "feathering" process.

Feathering combines two data-sets, typically at low and high angular resolution, by adding the Fourier transforms of the images weighted by their PSFs (Stanimirovic, 1999). Using the uv-plane notation, the observed single-dish image is

$$I^{\rm SD}_{\rm obs}(l,m) = I(l,m) * s^{\rm SD}(l,m), \tag{2.4}$$

where $I^{\rm SD}_{\rm obs}(l,m)$ is the observed sky image of the true sky convolved with the ($\sim$Gaussian) single-dish PSF $s^{\rm SD}(l,m)$. The Fourier transforms of the interferometric ($V_{\rm int}(u,v)$) and single-dish images ($V_{\rm SD}(u,v)$) are combined:

$$V_{\rm comb}(u,v) = \left[1 - \mathcal{S}^{\rm SD}(u,v)\right] V_{\rm int}(u,v) + f_{\rm cal}\mathcal{S}^{\rm SD}(u,v)V_{\rm SD}(u,v), \tag{2.5}$$

where $\mathcal{S}^{\rm SD}(u,v)$ is the Fourier transform of $s^{\rm SD}(l,m)$ normalized to unity and $f_{\rm cal}$ is a calibration correction factor (see §2.3.4; Stanimirovic, 2002). The inverse Fourier transform produces a combined image $I_{\rm comb}(l,m)$ with the high angular resolution from an interferometer and total flux recovery of single-dish image.

Radio interferometry is complex and this section is a minimal overview of the procedures required to produce inteferometric images. See Thompson et al. (2017) for a detailed explanation of all aspects of radio interferometry. See Stanimirovic (2002) for an overview of combining interferometer and single-dish observations (also see Koda et al., 2019).

2.2 Historical ISM observations in the Local Group

As the nearest galaxies, the Local Group galaxies are historically well-studied at nearly every accessible wavelength across the electromagnetic spectrum. This includes many observations of the 21-cm HI line over the last 63 years. Tables 2.1 & 2.2 show basics properties of 21-cm HI observations in M31 and M33, respectively. These observations are a mix of different telescopes (including single-dish and interferometers).

Two major advances are evident from the evolution of these 21-cm HI observations. The first is the evolution of spectrometers for high resolution observations of spectral lines. Early observations of both galaxies had spectral resolutions of $\gg 10$ km s^{-1}, useful for (a) detection-only studies (van de Hulst et al., 1957; Volders, 1959) or (b) the earliest rotation curve measurements, particularly for M31 (Roberts & Whitehurst, 1975). Since little was known about the HI distribution in other galaxies at this time, these studies identified how ubiquitous HI is across nearby spiral galaxies.

The second advance evident from Tables 2.1 & 2.2 is the introduction of long-baseline (> 1 km) interferometers over the past ~ 3 decades. The development of the VLA, Westerbork Synthesis Radio Telescope (WSRT), and, in the southern hemispheres, the Australian Telescope Compact Array (ATCA), allowed for sub-arcminute resolution with high surface brighness sensitivity[5]. This improvement in resolution continues to be a major advance for high-resolution HI studies, as is clear from this thesis.

The new observations I present in this thesis are the next step forward in HI observations of the Local Group. Compared to previous observations, our new observations combine high angular and spectral resolution with comparable sensitivity to previous observations over the 2000–2010 decade. The key difference with our HI observations is the < 1 km s^{-1} spectral resolution,

[5]I note that the VLA, in its most extended A-configuration, can resolve $\sim 1''$ scales at the 21-cm HI line. However, due to the sparse uv-sampling, the surface brightness sensitivity is poor and therefore is not well-suited for mapping extended emission.

making our observations one of the first extragalactic HI data sets capable of resolving thermal line widths in the CNM (§1.2.1).

At this time, we are now reaching the limitations of cm-interferometers for tracing HI emission. The observations I describe in the following sections are large time investments from the VLA, and in §7.2, I mention our continuation of this survey. If fully observed, the complete survey will use $\sim$ 2200 hr of observing time ($\sim$ 400 hr have already been observed). To reach a higher angular resolution, observations of HI require an expanded interferometer with (1) long (> 25 km s^{-1}) baselines, and (2) an increased number of antennas (the VLA typically observes with 27 antennas). The first provides the high angular resolution and the second provides high-fidelity imaging for extended emission combined with a higher sensitivity from a larger collecting area. These improvements are planned for new/expanded cm-interferometers (e.g., the SKA and ngVLA) that will come online in roughly the next 10 years.

2.3 An L-band Local Group VLA survey

As part of a larger and on-going Local Group VLA survey, I used L-band observations of M31 and M33 taken in the VLA's B, C, and D configurations[6]. Table 2.3 which projects these observations were taken with.

The angular size of both M31 and M33 (see Table 1.1) exceed the VLA's primary beam of $\sim 32'$ at 1.4 GHz. Because of this, these projects observed multiple pointings of each galaxy, which will later be mosaiced together to create larger maps (§2.3.3). M31 is observed with 49 pointings in D-configuration to create a full galaxy map out to $\sim$ 30 kpc in galactocentric radius[7]. The B- and C-configuration observations of M31 are of a smaller 7-pointing region focused on the northern half of M31, a region covered extensively by other larger observational projects at different wavelengths (e.g., the Panchromatic

[6]The D-configuration is the most compact configuration, with maximum baselines of 1 km. The C- and B-configurations are more extended, with maximum baselines of 3.4 and 11.1 km, respectively.

[7]Two additional pointings were used to map the dwarf galaxies NGC 185 and NGC 205, which are satellites of M31. These data are not presented in this thesis.

Reference	Instrument	Res. ($''$)	Channel Width (km s^{-1})	Noise (K)
van de Hulst et al. (1957)	Dwingeloo 25-m	2160	33	0.2
Burke et al. (1963)	Green Bank 90-m	600	2.1	–
Argyle (1965)	DRAO 25-m	2160	35	0.1
Roberts (1966)	Green Bank 90-m	600	21	–
Gottesman et al. (1966)	Jodrell Bank 250-ft	1080×840	42	1.0
Brundage & Kraus (1966)	Ohio State 80/24-m	2280×600	21	–
Davies & Gottesman (1970)	Jodrell Bank 250-ft	1080×840	8.4	–
Whitehurst & Roberts (1972)	Green Bank 140-ft	1260	5.5	–
Guibert (1973)	Nancay	1440×240	12.7	0.16
Emerson (1974)	Cambridge Half Mile	132×90	39	1.5
Roberts & Whitehurst (1975)	Green Bank 90-m	600	6.6	–
Newton & Emerson (1977)	Cambridge Half Mile	330×216	16	0.3
Cram et al. (1980)	Effelsburg 100-m	528	5.5	0.1
Unwin (1980a,b)	Cambridge Half Mile	72×48	16	3.6
Bajaja & Shane (1982)	WSRT	36×24	27	1.4
Brinks & Shane (1984)	WSRT & Effelsburg	36×24	8.2	0.6
Braun (1990)	VLA	11×9	5	9
Thilker et al. (2004)	GBT 100-m	546	1.3	0.1
Carignan et al. (2006)	Effelsburg	553	0.64	–
Braun et al. (2009)	WSRT & GBT	18×15	2.3	2.7
Chemin et al. (2009)	DRAO-ST	90×60	5.3	~ 1
This Work	VLA & Effelsburg	56×56	0.42	0.7
This Work	VLA & Effelsburg	18×16	0.42	2.8

Table 2.1: 21-cm observations of M31. Where available, the noise values correspond to the per-channel rms noise.

Reference	Instrument	Res. ($''$)	Channel Width (km s^{-1})	Noise (K)
Volders (1959)	Dwingeloo 25-m	2040	33	0.2
Dieter (1962)	Harvard 18-m	3180	17	0.2
Burke et al. (1963)	Green Bank 90-m	600	2.1	–
Meng & Kraus (1966)	Ohio State 80/24-m	2280 × 600	100	–
de Jager & Davies (1971)	Jodrell Bank 75-m	1080 × 840	8.4	–
Gordon (1971)	Green Bank 90-m	684 × 600	21	–
Wright et al. (1972)	Cambridge 2 12-m	180 × 90	39	3.2
Huchtmeier (1972)	Nancay 4 200-m	1440 × 240	12.7	0.1
Rogstad et al. (1976)	Owens Valley 2 27.4-m	120	10	0.5
Huchtmeier (1978)	Effelsburg	720	6.5	–
Reakes & Newton (1978)	Cambridge Half Mile	900 × 420	16	0.1
Newton (1980)	Cambridge Half Mile	93 × 47	16	–
Deul & van der Hulst (1987)	WRST & Effelsberg	24 × 12	8.2	1.2
Corbelli & Schneider (1997)	Arecibo	270	4.1	5.5$^{\blacklozenge}$
Putman et al. (2009)	Arecibo	204	5.2	0.3
Gratier et al. (2010)*	VLA	6	1.3	24
Lockman et al. (2012)	GBT 100-m	546	2.9*	0.2
Kam et al. (2017)	DRAO-ST	114 × 58	2.6	1.1
Koch et al. (2018c)	VLA	19 × 18	0.2	2.7

$^{\blacklozenge}$ Sensitivity presented in terms of integrated intensity (K km s^{-1}).
* Though the data was taken with 0.17 km s^{-1} channels, as used here.
* Also presented in Thilker et al. (2002).

Table 2.2: 21-cm observations of M33.

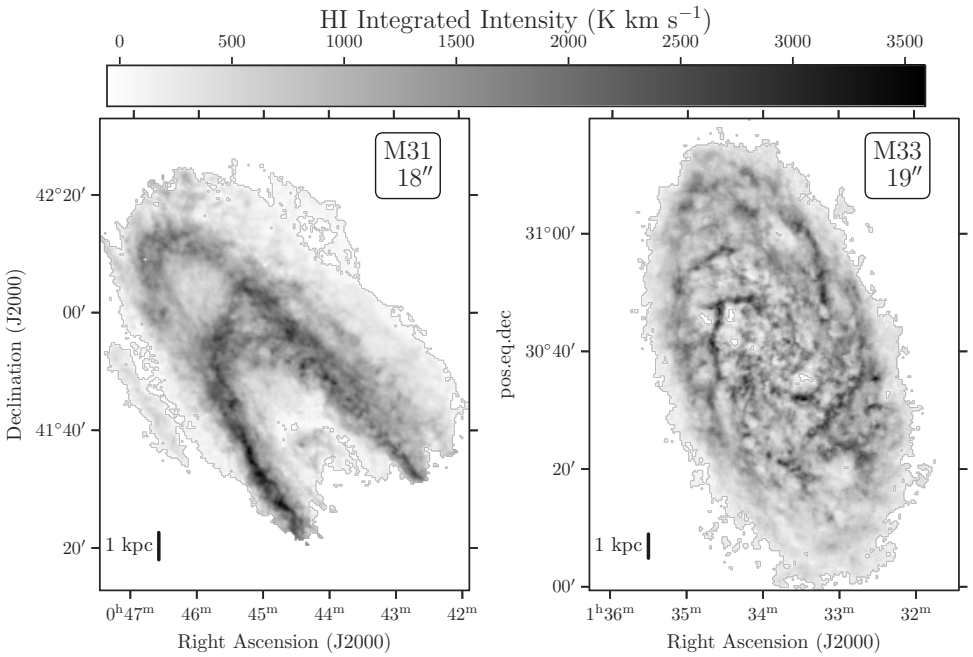

Figure 2.1: HI integrated intensity maps of M31 (left) and M33 (right) on $\sim$ 60–80 pc scales. This M31 mosaic has 7-pointings covering much of M31's northern half. The M33 mosaic is a 13-pointing mosaic which covers M33's star-forming disc.

Hubble Andromeda Treasury; Dalcanton et al., 2012). Figures 2.1 & 2.2 shows the final integrated intensity maps of these mosaics in HI, which are described in the remainder of this chapter.

Our M33 observations used a 13-pointing mosaic to cover the entirety of M33 out to $\sim$ 12 kpc in galactocentric radius. Due to an error in transcribing the pointing centres, the M33 observations actually use 14-pointings. The incorrect pointing centre is within a central region of the galaxy, leading to a moderately higher sensitivity in one part of the mosaic. The variations in the sensitivity are accounted for in the standard imaging process. Figure 2.1 shows the HI integrated intensity map for these observations.

One of the VLA's greatest strengths is the flexibility of its correlator, allowing for multiple windows in frequency to be simultaneously observed at different frequency resolution. Our M33 observations include coverage of the

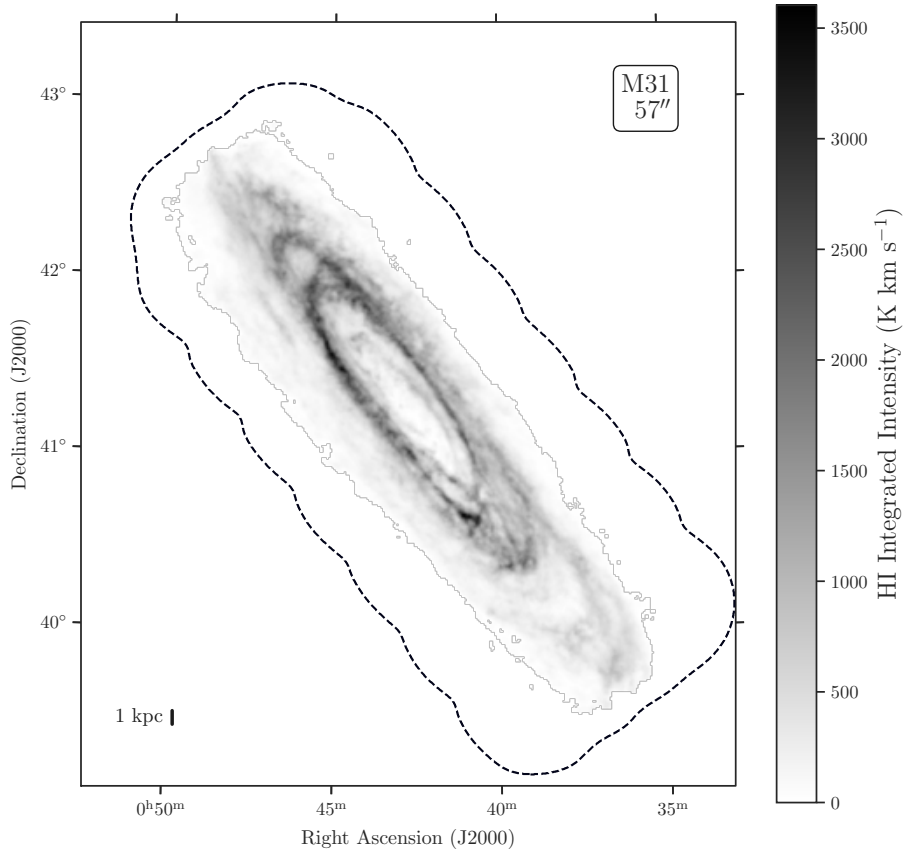

Figure 2.2: HI integrated map covering most of M31 with the VLA's compact D-configuration. The 57″ resolution corresponds to a physical resolution of 200 pc. The dashed contour shows the edges of the mosaic which includes 49 pointings.

entire 1–2 GHz L-band polarized continuum, the HI line, four hydroxyl (OH) lines[8], and six hydrogen radio recombination lines (RRLs)[9]. The M31 observations are similar, though no hydrogen RRLs are included in the spectral setup.

This additional spectral coverage enables a variety of science cases apart from our primary HI-drive goals. While the OH M31 observations have not been explored, I have, used the other spectral line observations of M33 in two places:

- No hydrogen RRL emission is detected from our M33 observations, but an upper limit is provided in Koch et al. (2018c, Chapter 3). The non-detection is consistent with more sensitive single-dish observations (Araya et al., 2004).

- I discovered the first OH maser in the 1665 MHz transition in M33 (Koch et al., 2018a). These masers arise from the regions surrounding late-type stars and young stellar objects. This detection is consistent with the brightest known OH-1665 masers in the Milky Way and two Hα sources are within the position we find, limited by the 10″ resolution of our observations.

The polarized continuum emission has not yet been explored.

The M31 and M33 HI spectral setups are different in their respective observations. All M31 HI observations were observed with a resolution of 1.95 kHz,

[8]The 4 OH lines at 1612, 1665, 1667, and 1720 MHz are probes of thermal and stimulated emission in the ISM. Thermal emission is faint, but stimulated emission is extremely bright. This stimulated emission is referred to as a maser, and OH masers are known to be associated with regions surrounding late-type stars, young stellar objects, and supernova remnants (Gray, 2012).

[9]RRLs are electronic transitions in atoms or molecules, similar to common optical transitions like the $n = 3 \rightarrow 2$ 656 nm Hα transition in the Balmer series of hydrogen. RRLs, however, are the electronic transitions at high n (~ 100), where the separation between energy levels is small and so transitions of $\Delta n = 1$ produce spectral-lines at radio wavelengths. RRLs are faint but do not suffer from dust extinction like lines in the UV or optical do. Because of this, RRLs are excellent tracers of recombination from HII regions, and therefore measure the ionizing photon output from young massive stars (Murray & Rahman, 2010; Murphy et al., 2018).

Galaxy	Config	Time on source (hr)	Program ID	PI
M31	D	39	14A-235	Leroy
M31	C	25	15A-175	Leroy
M31	B	18	15A-175	Leroy
M33	C	39	14B-088	Rosolowsky

Table 2.3: Local Group VLA projects used in this thesis.

corresponding to a velocity resolution of 0.42 km s^{-1} using the radio Doppler formula:

$$v = c\left(\frac{f_0 - f}{f_0}\right), \tag{2.6}$$

where v is the velocity, c is the speed-of-light, f_0 is the rest frequency of the spectral line, and f is the measured frequency[10]. The M33 observations instead have a finer resolution of 977 Hz, approximately half that of the M31 observations. This corresponds to a velocity resolution of 0.21 km s^{-1}. The difference is minimal, and when direct comparisons are made for the velocity information between M31 and M33 (see Chapter 4), we down-sample the M33 HI observations to the 0.42 km s^{-1} resolution.

These observations are the first part of a larger Local Group VLA survey that includes additional observations of M31, M33, and the Local Group dwarf galaxies NGC 6822, IC10, IC1613, and Wolf-Lundmark-Melotte (WLM). Our on-going survey will provide uniform mapping of HI, OH, and the 1–2 GHz continuum.

2.3.1 Single-dish observations

As described in §2.1, radio interferometers filter out emission on scales larger than what the smallest antenna baseline is in the interferometer. For both C- and D-configurations of the VLA, the shortest baselines is $b_{\rm min} = 35$ m,

[10]The radio doppler formula is an approximations valid only for velocities close to 0 km s^{-1} (within $\pm$1000 km s^{-1}) and is primarily used for Milky Way studies. However, the Local Group galaxies have low systemic velocities and so this approximation is valid given our spectral resolution.

though some of our observations have $b_{\rm min} = 45$ m. Since the antennas are 25-m in diameter, the shortest baselines cannot be shortened below 35 m or the antennas may block each other. Because of this, our VLA observations are not sensitive to scales larger than

$$\theta \sim \lambda / b_{\rm min} \sim 16,\, 20', \tag{2.7}$$

with $\lambda = 21$ cm and $b_{\rm min} = 45$ or 35 m. Large single-dish telescopes, including the 100-m GBT and Effelsburg telescopes resolve, scales of $\sim 9'$. Because the angular scales overlap between the VLA and single-dish observations, HI observations from the GBT and Effelsburg can be combined with the VLA data to ensure the full HI flux is recovered. I describe this combination in detail in §2.3.4.

Chapter 3 includes a description of the reprocessed HI GBT observations from Lockman et al. (2012). I briefly describe the key properties here. These data do not optimally overlap in the centre of M33 as these observations were taken to detect faint HI emission surrounding M33 and the HI bridge towards M31. Because of this, we used a Gaussian-Bessel kernel which moderately lowers the effective angular to $9.'8$, or ~ 2.3 kpc physical resolution for M33. We gridded the data from the calibrated scans directly to match the spectral channels of our VLA HI data cube. The highest rms noise level in these data is 120 mK.

For M31, I used the publicly-available Effelsburg HI data from the Effelsburg-Bonn HI Survey (EBHIS; Winkel et al., 2016). EBHIS is a whole northern sky HI survey with spectral coverage that includes the Local Group galaxies. I used the region containing M31 and regridded the data to match our VLA mosaics.

The EBHIS data originally have a spectral resolution of 0.47 km s^{-1}, slightly coarser than the VLA data. However, with an effective resolution of $10.'8$, the HI spectral lines are signficantly smoother from averaging over large regions in M31. Because of this, moderately up-sampling the EBHIS data to 0.42 km s^{-1} did not alter the overall line shape. The regridded data have an rms noise

level of 70 mK.

Both single-dish data sets have a sensitivity below our VLA observations. Because of this, combining the single-dish and VLA observations (§2.3.4) will not appreciable add increase the noise in the final HI images.

2.3.2 VLA data reduction

I calibrated the data sets using the VLA reduction pipeline implemented in the Common Astronomy Software Applications (CASA)[11] package. Because the reduction of the data took place at different times, I used different versions of the pipeline. For the M33 spectral line data, I calibrated the data with a modified version of the VLA scripted pipeline (version 1.3.0)[12] running on CASA 4.2.2. For the M31 spectral-line data, I used the VLA pipeline packaged with CASA 5.4.1. The basic steps in both reduction pipelines are the same, with improvements made in more recent versions to improve calibration solutions and automatic flagging of radio frequency interference (RFI) from terrestrial sources. RFI drastically impacts the 1–2 GHz continuum, however, the 1.42 GHz HI and 1.665 and 1.667 GHz OH lines are within or near to protected frequency bands. Poorly performing antennas or other systematics affect the HI and OH data more than RFI.

Both versions of the pipeline were modified to better handle spectral-line observations. The modification steps are:

- I split the spectral windows used for continuum versus spectral-lines before running the reduction. The continuum data is not used to find calibration solutions and can therefore be separated.

- I disabled Hanning smoothing for the reduction. Hanning smoothing convolves the spectral channels with a $[0.25, 0.5, 0.25]$ kernel, which avoids spectral ringing due to the Gibbs phenomenon. The ringing is severe

[11] casa.nrao.edu/casadocs

[12] available in 2015/2016; https://science.nrao.edu/facilities/vla/data-processing/pipeline/scripted-pipeline

when bright RFI is observed. However, this introduces a strong correlation between adjacent channels that is not ideal for our high spectral resolution imaging of HI. And since HI falls within a protected band, we do not expect to encounter strong RFI.

- Calibration sources are used to (1) set the absolute flux to convert the observations to physical units, (2) correct for systematic imperfections, and (3) atmospheric corrections. These corrections require the calibration source to have a known flux, structure (ideally, a point source), and spectral shape. The spectral shape is the key to correct for the imperfect sensitivity of the telescope's electronics at different frequencies, i.e. the bandpass. To do this, the calibration sources are assumed to have no spectral line features and their emission is modelled as a power-law in frequency. However, our HI spectral coverage overlaps with Milky Way velocities and some calibration sources have HI absorption features. To correct for this, I flagged frequencies affected by HI absorption from the Milky Way in the calibrator sources. To recover the frequencies affected by this range in the observations, the bandpass is interpolated over the missing range.

- The standard pipeline includes automated RFI flagging that identifies bright narrow spectral features. Since the HI and OH may be detected near the spectral resolution of the data, I disabled running the automated flagging on all science targets. It was, however, kept for the calibration targets used to derive calibration tables for the science targets.

- The standard pipeline includes a re-weighting routine (`statwt`) that downweights integrations with a large variance. This re-weighting minimizes the effect of remaining RFI or certain systematics when imaging the data. However, a spectral-line may also increase the variance and therefore be downweighted. To avoid this, I excluded the frequency range of M31 and M33 from the weighting routine.

To achieve an optimal calibration solution, I iteratively ran the pipeline procedure. After each run, I reviewed the diagnostic plots made by the pipeline (the "weblog") and a custom set of figures made per scan in the observations to rapidly identify regions with poor data. From this, I investigated data that required additional flagging, identified the source of the issue, and added a custom flagging command that is applied to the data when the pipeline is re-run. Because most of the spectral lines fall in or near a protected band, the majority of the additional flagging was due to poorly performing antennas or likely systematics from the telescope. The number of pipeline runs for each observation track ranged from 1–5 with an average of 2 runs.

From the calibrated data, I applied two final steps before imaging the data:

1. I combined all observations together. Individual observation "tracks" are typically $\sim$ 4 hr long for the VLA. To achieve high sensitivity, that data are combined together for imaging.

2. I subtracted continuum emission from the spectral-line band. The continuum emission is typically from extremely bright background sources (mostly AGN, though some radio jets are detected in narrow spectral channels, as well) and bright HII regions within the galaxy.

I note that the continuum subtraction step is optional, and including the continuum emission can be used for HI absorption studies . Our observations enable HI absorption studies, similar to the recent SMC study by Jameson et al. (2019). The HI absorption is not studied in this thesis but will be explored in future work.

2.3.3 Techniques for imaging massive interferometric data

The upcoming generation of radio and mm-interferometers will produce petabytes of data, firmly converting radio astronomy into a "big data" field. One of the primary bottlenecks is the computing power required to convert the *uv*-data into producing science-ready images. The $\sim$ 100 hr of VLA data described

here, and the additional $\sim$ 300 hr already observed in our survey, are an example of this deluge of data that needs a new imaging approach to create science-ready products in less than a thesis timescale. For reference, the raw data from the VLA has a total footprint of $\sim$ 2.5 TB, which is reduced to $\sim$ 500 GB for only the HI after calibration and removal of flagged (unusable) data. Our final imaging products are around 30 GB in size.

The imaging products are position-position-velocity cubes, which I will refer to as "data cubes" throughout this section. These are collections of spectral channels in the data arranged into a three-dimensional data product. Each spectral channel is the emission from the galaxy at that frequency, corresponding to some Doppler-shifted velocity of the atomic gas.

Standard imaging software, like CASA, can handle the input $\sim$ TB data sizes but struggles to produce the $\sim$ 30 GB imaging products as it needs to be kept in memory. Furthermore, the imaging process using the CLEAN algorithm is fairly inefficient as the number of iterations to deconvolve the signal in a $\sim$ 30 GB data cube for our targets is $\sim 3 \times 10^7$.

The process I describe here was used to image the three HI data cubes—one for M33 and two for M31 (see Figures 2.1 & 2.2)—highlighted in this thesis (Table 2.4). The M33 cube is imaged from the C-configuration observations (see Table 2.3). For the M31 imaging, I used the D-configuration observations to produce the lower-resolution 57″ cube covering the whole of M31. The smaller 7-point mosaic observed in B- and C-configurations was combined with the matching pointings in D-configuration. All three configurations were imaged together, with a *uv*-taper applied, to produce the 18″ M31 cube. The *uv*-taper was introduced to produce an equivalent C-configuration resolution cube to match the M33 observations. The sensitivity is boosted by using multiple configurations of the VLA as there is a large overlap in baseline lengths in configurations B, C, and D.

These data were imaged at different times: the M33 data in 2017 and the M31 in 2019. Because of this, CASA versions 4.4 and 5.4.1 were used to image the respective data cubes.

To handle the large computational time needed for imaging the HI data, we imaged and deconvolved each spectral channel separately (additionally described in Chapter 3; Koch et al., 2018c). This approach minimizes the size of the *uv*-data and the time required for input/output operations. Furthermore, each channel can be imaged independently and can therefore be parallelized by deploying the imaging on a cluster. I split the HI *uv*-data into individual channels at the target spectral resolution, either 0.21 or 0.42 km s^{-1} for the M33 and M31 data, respectively.

I produced the final deconvolved images for each spectral channel in two stages. First, I imaged the data and deconvolved using the multi-scale CLEAN algorithm. For imaging, I used natural weighting[13] for all HI imaging to maximize point source sensitivity and recovery of extended emission. The multi-scale CLEAN algorithm extends the original CLEAN algorithm by subtracting components of extended emission instead of only a point-source (Cornwell, 2008). To capture the *uv*-range sampled by the observations, I used the scales of $[0, 18, 36, 72, 180, 360, 720, 1440]''$ for deconvolution[14]. The deconvolution was considered complete once it reached the 2σ noise level in a channel, where σ was estimated from signal-free spectral channels[15]. Figure 2.3 shows an example of the imaging process for one spectral channel in M31.

Due to the complex extended emission, I found that a signal mask was necessary to restrict the deconvolution. Without a signal mask, the deconvolution often became unstable, leading to a diverging model solution. This was particularly important for M31 because its high inclination angle produces bright elongated emission in individual spectral channels, and due to the missing short-spacing from the VLA, the dirty map has large negative "bowls" in the dirty image. The multi-scale CLEAN algorithm is not well-suited for elongated structure as the model components at large scales are produced by an

[13]Natural weighting uses the inverse variance $1/\sigma^2_{\rm rms}$ when gridding the *uv*-data.

[14]I tested difference choices of scales, distributed over a similar range, and found that they did not affect the final images.

[15]I note that the M33 cube, as is described in Chapter 3, some channels did not converge to 2σ and instead reached 2.0–2.5σ. Because of this, we reported the upper limit. This issue was not encountered using the newer CASA 5.4.1 for the M31 imaging.

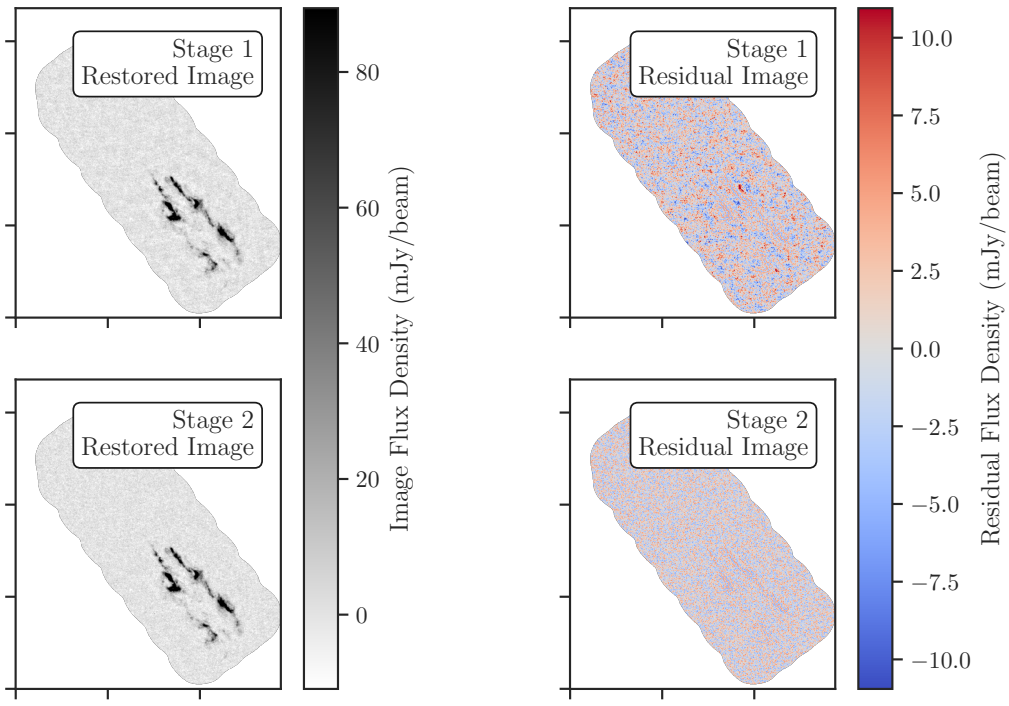

Figure 2.3: An example of one HI spectral channel from the M31 D-configuration mosaic using the two stage imaging process. The two columns show the restored images (left) and residual images (right) for each stage. The residual colourbar shows $\pm 3\sigma$ the rms noise. After the first stage, which uses the `tclean` auto-masking procedure, the residual image still has small regions where significant emission remains above the 3σ level. Furthermore, there are low-level residuals across other parts of the map that appear "patchier" than simply white-noise correlated over the beam size. The second stage accounts for this by deconvolving all emission above 2σ to produce the stage 2 restored and residual images. The patchy residuals are no longer evident after the second round of deconvolution without a signal mask applied. I note that the bright elongated emission remains an issue for the multi-scale CLEAN algorithm that deconvolves isotropic components. Thus the stage 2 residual is left with low-level anisotropic noise residuals near the brightest emission in the restored image.

isotropic two-dimensional Gaussian.

I produced signal masks for deconvolution in two ways. For the M31 D-configuration mosaic, I used the auto-masking routine built into the `tclean` routine in CASA (Kepley et al., 2020). This auto-masking routine identifies peaks in the dirty image and includes regions containing those peaks out to a lower level emission set by a signal-to-noise ratio or sidelobe level from the PSF. This mask is expanded during the deconvolution process with updates when minor cycles complete and the residual maps are recomputed from the *uv*-data (Thompson et al., 2017). As the bright sources are deconvolved and removed, the sidelobe threshold decreases. Due to the bright HI emission in M31, I found that the auto-masking was improved by triggering rapid minor cycles in the deconvolution.

For the other two data sets, I created signal masks from lower resolution HI data cubes. For M33, I made a signal mask based on emission brighter than 3σ in the GBT data cube. This created a broad signal mask that successfully included all detected emission in the VLA observations. This worked well for M33 because it has a lower disc inclination and the blue- and red-shifted extents of its emission are not nearly as elongated as M31's emission.

For the $18''$ BCD M31 imaging, I created a signal mask from the D-configuration M31 map, following the signal identification algorithm described in Chapter 3. Because the resolution between the images is only different by a factor of ~ 3, the D-configuration map provides excellent prior information for emission at higher resolutions in M31.

This first imaging step effectively accounts for the vast majority of the emission in each spectral channel. However, noticeable low-lying emission or mild artefacts remains in the residual and restored images (Figure 2.3). Because of this, I added a second deconvolution step where the signal mask was removed and the entire image was deconvolved to the 2σ level. Because most of the bright emission is deconvolved in the first step, the stability issues were not an issue without the mask. The final image for each spectral channel is used following this second step.

A recent study by Ianjamasimanana et al. (2017) recommend cleaning down to the 1.5σ level to fully recover the HI line shapes. I found that the CASA cleaning algorithm struggles to get to this noise level and drastically increases the computing time relative to a stopping threshold at 2σ without a noticeable change in the recovered image properties. For some test channels, the deconvolution converged to an apparent "noise floor" before reaching 1.5σ. However, I used a different software package than Ianjamasimanana et al. (2017), suggesting there may be small variations resulting from different implementations of the CLEAN algorithm. If the resulting residual images are indistinguishable from noise, it is unlikely that the HI line shapes are strongly biased from the remaining residuals.

After all spectral channels had been imaged, I concatenated the channels together to create a final HI data cube. As mentioned above, each of our 3 data cubes is ~ 30 GB in size.

I used the Compute Canada Cedar cluster for the M31 imaging and the now defunded Westgrid Jasper cluster for the M33 imaging. Each spectral channel was imaged in parallel using 8 cores and at least 16 GB of memory. Stage 1 and 2 imaging each required 12 hr for the M31 D-configuration and 30 hr for the higher resolution M33 and M31 data, on average, respectively. The ability to deploy many imaging jobs drastically sped up the imaging versus imaging the cube together. While CASA does have the ability to automatically split spectral chunks of the data, the input/output operations continued to be a bottleneck from tests on the Cedar cluster.

2.3.4 Combining single-dish and interferometric observations

To account for missing flux in the VLA observations, I combined our VLA HI data cubes with the single-dish HI observations presented in §2.3.4. To combine the data, I used the feathering technique implemented in the `uvcombine` python package,[16] which I have made large contributions to. The details of

[16]https://github.com/radio-astro-tools/uvcombine

this process are described in §3.8.1.1. I used the same procedure for all of the HI cubes and so only outline the steps here.

Before the feathering step, the VLA and single-dish data sets were spatially-registered using the `image-registration` package[17]. All offsets were within the resolution of the single-dish observations.

Feathering is strongly dependent on the weighting functions used for the inteferometric and single-dish images, respectively (Equation 2.5). However, the weighting functions depend on both the beam size and the assumption of equivalent absolute flux calibration between the two images. If either of these quantities are incorrect for the images, the combined image will not be added correctly and may bias derived properties.

Following Stanimirovic (1999), I used two tests to ensure the correct single-dish beam size was used for the weighting function and that the absolute flux scale matched in the VLA and single-dish data. Both of these tests require matching the scales where both images are sensitive to. If there are no systematic issues, the flux properties on those scales should be equivalent. For the M33 data, the overlapping scales from from $9.'8$ to $16'$ (set by the GBT resolution and the VLA shortest baselines, respectively). For the M31 data, the overlapping scales are from $10.'8$ to $16'$. Ideally, there should be a factor of at least 1.7 in the spatial overlapping scales (Kurono et al., 2009). The M33 scale overlap satisfies this criterion, though the M31 scale overlap is just below this factor due to the moderately larger beam size in the EBHIS data.

For each spectral channel, I calculated the Fourier transform of both images and considered each pixel that falls within the overlap region. This provided $\sim$ 200 pixels per spectral channel. Each of the tests considers the interferometer-to-single-dish intensity ratio, defined as :

$$f_{\rm cal} = I_{\rm inter}/I_{\rm SD}. \tag{2.8}$$

The ratio $f_{\rm cal} = 1$ corresponds to an ideal overlap between two noise-less images at different resolutions.

[17] image-registration.readthedocs.io

The first test ensures the correct single-dish weighting is applied. To do this, I used the approximate relation between scale factor and uv-distance k from Stanimirovic (1999):

$$f_{\rm cal} = \left[1 + \frac{\Delta\theta(2\theta_0 + \Delta\theta)}{4{\rm ln}2} k^2\right], \tag{2.9}$$

where θ_0 is the true FWHM single-dish beam size and $\Delta\theta$ is the deviation from the true beam size and assumed to be small. When the incorrect θ_0 is used, there will be a bias with uv-distance between $f_{\rm cal}$ and k^2. The true beam size recover a slope of 0 in this relation. This test was used to ensure the $9.'8$ GBT resolution for the M33 data is correct. The EBHIS data showed no bias at the $10.'8$ beam size reported by Winkel et al. (2016).

The second test measures the flux factor between the two images. If the flux ratio is not 1, the absolute calibration between the data sets differs. Since single-dish observations are more difficult to flux calibrate compared to inteferometers, this correction factor is typically applied to the single-dish image.

The issue in determining the scale factor from the $f_{\rm cal}$ is the presence of noise in $I_{\rm inter}$ and $I_{\rm SD}$. In §3.8.1.1, I compared different methods of calculating the scale factor and found that the distribution of $\log_{10} f_{\rm cal}$ values typically follows a Cauchy (or Lorentzian) distribution. I then adopted the centre of the fitted distribution as the scaling factor.

For the M33 observations, I calculated a scaling factor of 1.02 ± 0.06. Since this values is consistent with 1, no scaling factor is applied for feathering.

For the M31 observations, I used these overlap tests with our D-configuration mosaic since it includes most of the galaxy and therefore has more overlapping area with the EBHIS image. Using the Cauchy distribution fit, I calculated a scale factor of 1.10 ± 0.01 and applied this scaling to the EBHIS data. Because the scale factor is > 1, this effectively increases the total flux in the combined data. However, a similar and independent test by Blagrave et al. (2017) also found a similar scaling factor of 1.10 ± 0.01 when combining EBHIS and interferometric data from the Dominion Radio Astrophysical Observatory Synthesis Telescope. There is no flux calibration discrepancy between the dif-

ferent M31 VLA configuration observations, so the 1.1 scale factor is applied to the 18″ M31 cube, as well.

These feathering tests I developed are publicly available in the `uvcombine` package. Since the relative flux scale is set by the single-dish data alone, these tests are crucial to ensure correct flux recovery and relative weighting between data sets. This is particularly important when combining data products from different observatories or archival products initially intended for another purpose. For example, Chapter 3 describes subtle effects in the GALFA data from the Arecibo Observatory of M33 (Putman et al., 2009; Peek et al., 2011) which produced a large-scale "ripple" in the feathered data.

The combined HI data cubes are the products used throughout this thesis. Table 2.4 summarizes the properties of the HI data cubes. Figures 2.1 & 2.2 show the integrated intensity maps from the final combined data cubes.

Target	VLA Config.	Angular Res. (″)	Physical Res. (pc)	Channel Width (km s^{-1})	Per-Channel Sensitivity (K)	5σ HI Column Density (cm^{-2})
M31	D	57×56	200	0.4	0.7	2.6×10^{18}
M31	B, C, D	18×16	65	0.4	2.8	9.8×10^{18}
M33	C	19×18	80	0.2	2.8	5.1×10^{18}

Table 2.4: Summary of the HI VLA data cube properties used in this thesis. The HI column densities assume optically-thin emission (§1.2.1).

2.4 Summary

In this Chapter, I presented the new HI VLA observations that enabled the following research in this thesis. The M33 HI data is used in all of the science chapters (Chapter 3–6). The M31 18″ HI data is used in Chapter 4. And the M31 57″ HI data is used in Chapter 6.

Chapter 3

Kinematics of the Atomic ISM in M33 on 80 pc scales

Koch, E.W et al. (2018). "Kinematics of the atomic ISM in M33 on 80 pc scales," *MNRAS*, 479, 2505–2533.

Abstract

We present new L-band (1–2 GHz) observations of the nearby spiral galaxy M33 with 80 pc resolution obtained with the Karl G. Jansky Very Large Array. The HI observations, combined with HI measurements from the Green Bank Telescope, improve the spectral resolution and sensitivity (2.8 K rms noise in a 0.2 km s^{-1} channel) compared to previous observations. We find individual profiles are usually non-Gaussian, harbouring line wings, multiple components, and asymmetries. Given this spectral complexity, we quantify the motions in the atomic ISM through moment analysis of the spectra and fits to aligned, stacked profiles. The measured value of the HI line width depends strongly on the method used, with the velocity stacked profiles aligned to the peak velocity giving the minimum value of $\sigma = 7$ km s^{-1} and all other methods giving higher values ($\sigma \sim 10$ km s^{-1}). All measurements of the line width show a shallow radial trend, with σ decreasing by ~ 2 km s^{-1} from $R_{\rm gal} = 0$ to $R_{\rm gal} = 8$ kpc. We consider a number of energy sources that might maintain the line width against turbulent dissipation, but no single source is adequate.

We find excess emission relative to a Gaussian in the stacked profile line wings, ranging from 9% to 26% depending on how the spectra are aligned. By splitting the line wings into symmetric and asymmetric components, we find that the lagging rotational disk accounts for one-third of the line wing flux. We also find emission far from the rotation-axis of the galaxy in multiple discrete HI clouds, including a filament with a projected length of ~ 8 kpc.

3.1 Introduction

HI 21-cm emission is an ideal tracer of kinematics in the atomic interstellar medium (ISM) because of its ubiquitous distribution that extends beyond the optical extent of galactic disks. On large-scales, HI emission can be used to study disk kinematics and rotation, while small-scale variations trace the kinematic and turbulent nature of the ISM.

Modeling the kinematics of the atomic ISM remains a significant challenge. The HI detected within the Milky Way exhibits kinematic features related to the evolution of clouds on small scales. Our limited perspective within the Galaxy makes it difficult to place these small-scale motions in the context of large-scale motions and the galactic potential. However, since the kinematics on both scales are linked, a consistent understanding of ISM kinematics requires sampling scales from molecular clouds to the entire disk (Dobbs et al., 2014). The external view of nearby galaxies can overcome these limitations in Galactic observations. However, these observations often lack spatial resolution and high sensitivity, yielding studies of 21-cm emission that blend the small-scale motions seen in Milky Way studies.

This trade-off of resolution and scale between Galactic and extragalactic observations is also important for discerning the multiple thermal states in the atomic medium. While superseded by later work, the Field et al. (1969) model predicts the presence of a warm neutral medium (WNM, $n \sim 1$ cm^{-3}, $T \sim 1000$ K) and cool neutral medium (CNM, $n \sim 10$ cm^{-3}, $T \sim 100$ K) over a range of pressure. Subsequent studies of the atomic medium qualitatively support

the two-phase picture (Wolfire et al., 1995; Wolfire et al., 2003). However, in observational studies, it remains difficult to separate these two gas phases. The optically-thin WNM is typically measured in HI emission, while the optically-thick CNM can only be unambiguously separated through absorption studies.

Decomposing HI spectra into thermal components has best been studied in the Milky Way through absorption towards bright extragalactic sources (Heiles & Troland, 2003; Murray et al., 2015), or by tracing HI self-absorption (Gibson et al., 2005). These studies find that the atomic ISM is found in both the WNM and CNM states, though a non-negligible fraction of components are in an unstable intermediate state (Heiles & Troland, 2003; Murray et al., 2015). Extragalactic absorption studies toward $\sim$ 50 lines-of-sight in Local Group galaxies suggest that the fraction of the atomic medium in the CNM varies between local systems (Dickey et al., 1990, 1994, 2000).

Most extragalactic studies focus on interpreting the HI emission line profiles at the telescope resolution (typically 200–1000 pc), though different modelling approaches are used. Fitting individual spectra with a Gaussian is the most straight-forward approach (Boulanger & Viallefond, 1992; Mogotsi et al., 2016), however non-Gaussian line features or multiple components require a more sophisticated model, such as a two-Gaussian model (Young & Lo, 1996; Warren et al., 2012). A simpler approach is to only estimate the line width of a spectrum using the second-moment (Tamburro et al., 2009), though this approach has been found to overestimate the line width in many cases (Mogotsi et al., 2016).

To study kinematics in in faint regions, several studies stack velocity-aligned spectra to increase the signal-to-noise (S/N) in the data. Nearly all studies that utilize this method find a common line shape of a central Gaussian peak with enhanced line wings (Young & Lo, 1996; Braun, 1997; Petric & Rupen, 2007; Ianjamasimanana et al., 2012; Stilp et al., 2013a). To characterize this shape, the stacked spectra are either fit with a two-Gaussian model (Young & Lo, 1996; Ianjamasimanana et al., 2012) or as a Gaussian peak with a non-parametric measure of the enhanced line wings (Stilp et al., 2013a; Stilp et al.,

2013b). The HI line widths from a single Gaussian range from 6–10 km s^{-1}, while the two-Gaussian models have a narrow component width of 3 to 6 km s^{-1}, and 9 to 25 km s^{-1} in the wider component.

The results of stacked profile studies, and the models chosen to explain the profile shape, have led to diverging physical interpretations of the atomic medium. Studies using a two-Gaussian model argue that the narrow and broad components are naturally explained by the CNM and WNM, respectively (Young & Lo, 1996; Ianjamasimanana et al., 2012). On the other hand, Stilp et al. (2013a) argues that the central Gaussian peak represents emission from a turbulent mixture of CNM and WNM, and the enhanced line wings result from stellar feedback. Alternatively, Braun (1997) proposes that the brightest emission arises from a "high-brightness network" (HBN) of narrow filamentary structure across the disk. The HBN then arises from a dense and optically thick CNM component, where the central peak becomes flattened in individual spectra due to the higher optical depth (Braun et al., 2009; Braun, 2012).

The use of different methods and their disparate interpretations makes it difficult to create a clear connection with Milky Way studies. Yet, with sufficient sensitivity and resolution, extragalactic HI observations should recover the small-scale complexity observed in the Milky Way. New observations of nearby galaxies can provide this connection between galactic and extragalactic approaches. Observations of the Local Group provide the best means for determining this connection as they allow for scales of indiviudal molecular clouds to be resolved, similar to studies within the Milky Way and Magellanic Clouds (e.g., Wong et al., 2009; Fukui et al., 2009). In particular, M33 is an ideal target due to its proximity (840 kpc, Freedman et al., 2001), moderate disk inclination (55°), and relatively small angular size. To that end, we have conducted new observations of M33 with the NSF's Karl G. Jansky Very Large Array (VLA) that focus on high sensitivity coupled with a high spectral resolution.

Recent work on M33 has focused on deep, lower resolution HI observations

seeking the origins of HI halo gas and high-velocity clouds in the M31 group (Lockman et al., 2012; Keenan et al., 2016), or the structure of the dark matter halo (Corbelli et al., 2014; Kam et al., 2017). There has been comparatively less attention on the nature of the atomic gas in M33's star-forming disk. The first complete HI map with the resolution to discern small-scale structure was presented by Deul & van der Hulst (1987), which unveiled the filamentary structure of M33's HI morphology. More recently, Gratier et al. (2010) use archival HI VLA observations from Thilker et al. (2002) to compare the HI properties around giant molecular clouds (GMCs). Druard et al. (2014) further perform a stacking analysis of the HI and CO to study the relationship of atomic and molecular gas. Imara et al. (2011) use an independent reduction of these same data to examine the HI environments around 45 GMCs in M33.

In this paper, we present new L-Band (1–2 GHz) observations from the VLA of M33 with a focus on high velocity-resolution observations of the 21-cm HI line. Taken in the VLA's C-configuration, the beam size at the HI line is $18-20''$or a physical size of ~ 80 pc. The VLA's correlator allows for simultaneous observations of the 21-cm HI line, four OH transitions, several hydrogen radio recombination lines (RRLs), and the polarized radio continuum. Here we focus on the HI observations to analyze the HI profile shapes on 80 pc scales and to determine the structure and kinematics of the HI disk. We detect no RRL emission (§3.3.1) and find a single OH maser (Koch et al., 2018a). The polarized radio continuum will be presented in a future paper.

In §3.2 we present the VLA observations and the Green Bank Telescope (GBT) HI data used to provide short-spacing information (Lockman et al., 2012). Aspects of the HI imaging, signal masking, and upper limits on the hydrogen radio recombination line (RRL) emission are given in §3.3. We explore the atomic gas properties of M33 in §3.4 and extra-planar HI structures in §3.5. In §3.6, we critically evaluate the meaning of 21-cm HI line profiles generated in extragalactic observations. §3.8.1 provides a detailed description of the imaging and short-spacing combination.

3.2 Observations

3.2.1 VLA

Using the VLA in C-configuration, we observed a 13-point mosaic covering the disk of M33 out to a radius of 12 kpc (Project 14B-088). The data were taken in 12 tracks, split equally before and after transit, for a total of 52 hours. We used a hexagonal grid of pointings at the HI frequency to ensure uniform sensitivity across the mosaic. To remedy a pointing error that omitted one of the mosaic centres, our final observation used a single $\sim$ 5 hour pointing on that location, yielding nearly equal integration times and noise properties across the mosaic. Due to its close proximity to M33, we use 3C48 as the flux, delay, and gain calibrator. 3C138 is observed as the polarization leakage calibrator for the continuum data.

Using the full capabilities of the VLA correlator, our spectral setup covers the entire 1–2 GHz L-band polarized continuum, the HI line, four OH lines, and six hydrogen radio recombination lines (RRLs), enabling a variety of science cases to be explored. The setup of the line spectral windows (SPWs) are shown in Table 3.1. The HI spectral window has a high spectral resolution of 977 Hz, corresponding to a velocity of 206 m s^{-1}. This high-spectral resolution is required to detect HI self-absorption (HISA), based on prior Milky Way observations (e.g., Gibson et al., 2005).

3.2.2 GBT

We reprocessed the high spectral resolution HI data presented in Lockman et al. (2012, project AGBT09A_17) to provide the short and zero spacing information on HI emission. The angular resolution of the GBT in the 21-cm line is 9.′1 and the spectral resolution is 0.16 km s^{-1}. Spectra were calibrated and corrected for stray radiation as described in Boothroyd et al. (2011), and a first-order polynomial was fit to emission-free regions of the spectra to remove any residual instrumental baselines.

The data were originally collected as four separate maps with minimal spa-

Table 3.1: The VLA SPW setup for spectral lines used in the observations.

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

tial overlap, so we gridded the data into a cube using a Gaussian kernel, rather than the preferred Gaussian-Bessel kernel. This choice eliminates edge effects in the data at the centre of the galaxy, but lowers the effective angular resolution to 9.′8 (see §3.8.1.2), which is a linear resolution of $\sim$ 2.3 kpc at the distance of M33. We used the GBTPIPE package[1] to build the data cube, which performs spectral preprocessing to eliminate bad scans and a kernel based gridding approach developed for on-the-fly data (Mangum et al., 2007). The noise in the final cube varies with position ranging from 50 mK to 120 mK in a 0.16 km s^{-1} channel giving a median column density error of 3.1×10^{17} cm^{-2} for a 20 km s^{-1} FWHM line. The shortest baselines in our VLA data provide information on scales of 16.′5, providing sufficient overlap in *uv*-space with the GBT beam to calculate comparison statistics (§3.8.1.1).

[1] v0.1.2, https://github.com/low-sky/gbtpipe

3.3 Imaging & calibration

We calibrated the data using a modified version of the VLA pipeline (version 1.3.0) with CASA 4.2.2 [2]. Following observatory recommendations, we used a modified version of the pipeline to better handle line spectral windows (SPWs), namely not using Hanning smoothing or automated RFI flagging on the line SPWs to maintain the spectral resolution and avoid flagging of narrow emission features (e.g., OH maser emission). Using 3C48 as the sole calibrator (excluding polarization) gives a high S/N for all calibration scans, and the automated calibration solutions found by the pipeline are excellent. After the initial pipeline run, we manually flagged the data before running the pipeline once more. Most of the line SPWs required little manual flagging beyond the pipeline solutions, though one RRL SPW could not be recovered due to RFI.

Here we present upper limits for RRL emission and detail the imaging and masking approaches adopted for the HI data. A single OH(1665) maser is detected and is presented separately (Koch et al., 2018a). We find no detections in the 1612, 1667, or 1720 MHz OH lines.

3.3.1 Radio recombination lines

We detect no RRL emission from the six observed transitions. One RRL SPW, the H(172)α line, was dominated by RFI and was unrecoverable. Prior to imaging, we subtracted a constant continuum background level using the CASA UVCONTSUB task by fitting the velocity ranges found to be emission-free in the HI imaging (§3.3.2). RRL emission is expected to be very faint and based on a prior RRL detection towards NGC 604 (Araya et al., 2004, see below), we do not expect to detect any signal in a single line. To lower the noise level, we average the individual line cubes together[3]. To use this process, we generated data cubes for each line at a common spectral resolution ($\sim$ 10 km s^{-1}) and corrected these by the primary beam pattern of the mosaic. Each

[2] https://science.nrao.edu/facilities/vla/data-processing/pipeline/scripted-pipeline

[3] https://casaguides.nrao.edu/index.php?title=Stacking_Multiple_Spectral_Lines_at_Same_Position

line was then convolved to a common spatial resolution, set by the lowest resolution line, H(166)α. Because the mosaic pointings were set based on the primary beam size at the HI frequency, the primary beam pattern and sensitivity differs between the RRLs. However, based on the the peak Hα flux from the giant HII region NGC 604, we place limits where the RRL intensity is expected to be brightest. Because of the pointing issue mentioned in §3.2.1, the region around NGC 604 also has the best primary beam coverage for all RRLs observed. Using these regions, we set a 3σ upper limit on the RRL intensity at 3.0 mJy within a $60''$ region that encompasses the optical extent of the HII region. This upper limit in agreement with observations by Araya et al. (2004) with Arecibo in C-band, where they detected the H110α line in NGC 604 with a peak line flux of 1.36 ± 0.19 mJy in a $58''$ beam.

3.3.2 H I imaging

We subtract a constant continuum component from the HI data using the CASA task UVCONTSUB. The continuum level is fit based on identified emission-free channels in both the VLA and GBT data. The VLA data are inspected in the *uv*-plane to determine emission-free channels. We then use the GBT data to determine the velocities affected by Milky Way HI emission. The large bandwidth covered by the HI line gives an ample number of HI-free lines, and the fitted background level is well-constrained.

Examining the extent of Milky Way HI emission in the GBT data is imperative for the red-shifted HI emission at the southern-most tip of M33, as they become spatially coincident near $v_{\rm LSRK} \sim -71$ km s^{-1}. This is indicated by the shaded region in Figure 3.1. This overlap region is clearly shown in the HI Arecibo data from Putman et al. (2009, see their Figure 3). Based on this, we begin imaging the VLA data at a velocity of -73.0 km s^{-1} to best avoid Milky Way emission. We estimate that excluding this spectral region removes ~ 50 Jy km s^{-1} of flux from M33, assuming that the missing emission is similar to the blue-shifted side. This flux corresponds to 0.5% of the total emission from M33.

The resulting data cube is large with 1178 velocity channels, each of which has a grid size of 2560^2 pixels. This presents significant computational barriers for imaging; our imaging process is described in §3.8.1. Prior to imaging, we determine optimal CLEAN settings by running numerous combinations on a single channel, identifying the imaging parameters that yield the lowest peak residual without the algorithm diverging. Only natural weighting is tested as we prioritize maximizing sensitivity to extended structure. A CLEAN mask is defined for each channel based on the 3σ limit in the GBT data, which covers nearly all of the emission in the VLA data. Some channels with emission near the mask edge used an expanded mask to ensure all emission was included.

We use the multi-scale CLEAN algorithm (Cornwell, 2008) for deconvolution with six scales, ranging from a point response to a quarter the largest recoverable scale ($\sim 970''$). Changing the range and the specific scales did not have a significant effect on the resulting image. Based on our single channel tests, we deconvolve each channel until reaching 3.8 mJy beam^{-1} (7.1 K). This limit is 2.5σ times the noise level in the final cube.

We then use the GBT data to provide short-spacing information for the deconvolved VLA cube by feathering the VLA and GBT data cubes together with the UVCOMBINE package[4]. A detailed explanation of the combination is provided in §3.8.1.1. The resulting cube fully recovers the total emission in the GBT data.

The final VLA HI cubes, with and without short-spacing data, have a 1σ sensitivity of 2.8 K per 0.2 km s^{-1} channel and a beam size of $19'' \times 17''$. This corresponds to a hydrogen column density of 1.0×10^{18} cm^{-2}, assuming optically thin emission. The mosaic is cut-off at a primary beam coverage of 0.5 to avoid noise-dominated regions near the edge.

3.3.3 Signal masking

We define a multi-step process for creating a reliable signal mask enabled by the high spectral resolution of our data. This masking is critical for our

[4]https://github.com/radio-astro-tools/uvcombine

analysis to ensure that moment-based estimations are not influenced by noise. When using the VLA-only data, the masking also removes the influence of negative bowls in the data from missing short-spacing information.

Assuming a typical Gaussian line width of $\sigma = 6$ km s^{-1} for the broad HI emission across the disk (see §3.4.4), each spectral feature will have a FWHM that spans about 70 channels. For the purpose of characterizing emission features alone (not HISA), we can substantially smooth in the spectral dimension to highlight low surface brightness features.

To determine the mask, we first spectrally smooth the data using a median filter with a width of 31 channels (~ 6 km s^{-1}). The noise in each smoothed spectrum is found using the median absolute deviation (MAD), iteratively rejecting points beyond $2\sigma_{\rm MAD}$ until convergence. We then search for valid spectral components by requiring that each component has a maximum intensity above $5\sigma_{\rm MAD}$ and has 30 consecutive channels above $2\sigma_{\rm MAD}$. To ensure low surface-brightness line wings are included, the edges of the valid components are extended until the intensity reaches $1\sigma_{\rm MAD}$. This procedure is performed on all spectra within the cube.

Next, we consider the spatial connectivity. To remove spurious spectral components found in the first step, we use the morphological opening and closing operators with a top-hat kernel equivalent to the FWHM of the beam (Shih, 2009). The opening operator will erode the mask edges, removing all features smaller than the beam kernel. The closing operator then dilates the remaining components in the mask to restore their original area. The combination of these two masking steps yields a robust signal mask [$\mathcal{M}(x, y, v)$], consistent in both the spectral and spatial dimensions.

This procedure is essential for high spectral-resolution data. Throughout the map, we find many lines-of-sight with either multiple, resolved components, or significantly skewed profiles. Not recovering the line wings in the masking procedure may hide these asymmetries at low-surface brightness. This also ensures negative bowls in the VLA-only cube are rejected from the signal mask and will not bias measurements of the asymmetry of a spectrum.

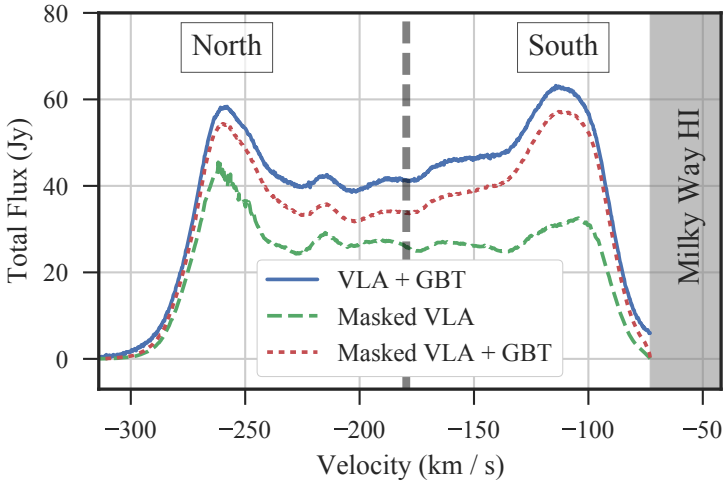

Figure 3.1: Total emission profiles of the VLA-only and the combined VLA+GBT data before and after masking. The curve for the unmasked combined data set is indistinguishable from an umasked GBT-only spectrum. The thick dashed vertical line indicates the systemic velocity of M33 at −179.2 km s^{-1}. The flux density in the masked VLA map is well recovered by the VLA data in the northern half of M33, but the southern half lacks a significant portion of the emission without short-spacing information.

In Figure 3.1, we compare the total GBT emission measured within the VLA mosaic to the emission retained after applying the masking procedure. For the VLA-only data, 60% of the total HI emission is contained within the mask. The combined VLA+GBT data contains 87% of the total emission in the mask. The missing flux density in the masked version arises from low surface-brightness features that do not satisfy the masking criteria. We discuss these low surface-brightness features in §3.5. The ratio of flux density retained in the mask is roughly constant across all channels for the combined VLA+GBT data, indicating that only low surface brightness HI is excluded by the masking procedure.

3.4 Properties of the atomic medium

In this section, we examine the properties of the HI emission derived from the new data in the context of previous HI studies (Corbelli et al., 2014; Druard

et al., 2014; Kam et al., 2017).

We calculate an atomic mass of $9 \pm 2 \times 10^8$ $M_\odot$ for the VLA data and $1.3 \pm 0.3 \times 10^9$ $M_\odot$ for the combined data, including a factor of 1.4 for He and heavier elements. When including regions outside of the VLA mosaic, the GBT data give a total atomic mass of $1.8 \pm 0.2 \times 10^9$ $M_\odot$, consistent with other independent HI mass measurements (e.g., Putman et al., 2009).

Figure 3.2 shows the column density map of the VLA data, highlighting M33's flocculent structure. The inner region of the disk ($R_{\rm gal} < 2$ kpc; see Figure 3.3 for context) is dominated by small-scale HI shells and lacks large ($\sim$ kpc) scale structures. The "mid-disk" ($R_{\rm gal} = 2$–4 kpc) is dominated by spiral arms prominent in the optical (e.g., Figure 2 in Corbelli et al., 2014). The HI shows similar spiral arm structure in the Northern half, while the Southern arm is dominated by one of the brightest HI clumps in the galaxy (Rosolowsky et al., 2007). The HI in the outer disk ($R_{\rm gal} > 4$ kpc) is nearly ubiquitous, with bright spiral-arm segments throughout. The outer edge of our signal-masked HI map is near the radius where the warp in M33's disk becomes significant (Corbelli et al., 2014).

Figure 3.3 shows four spectra extracted from different positions in the HI data cube. Each of the spectra shows a non-Gaussian profile with either multiple velocity components, extended line wings, or both. To describe these features, we use moment-based descriptions of the line profiles and velocity-aligned spectral stacking. Given the lines are clearly non-Gaussian, direct Gaussian fits to the lines do not produce a complete description of the emission.

We note that all uses of the line width in this paper are defined as the Gaussian standard deviation, not the full-width-half-max (FWHM).

3.4.1 Rotation curve and disc parameters

Using the masked cube from §3.3.3, we calculate the velocity at the peak intensity, shown in the left panel of Figure 3.4, to derive a rotation model. We define this velocity at peak intensity as $v_{\rm peak} \equiv {\rm argmax}_v\ I(v)$, which is

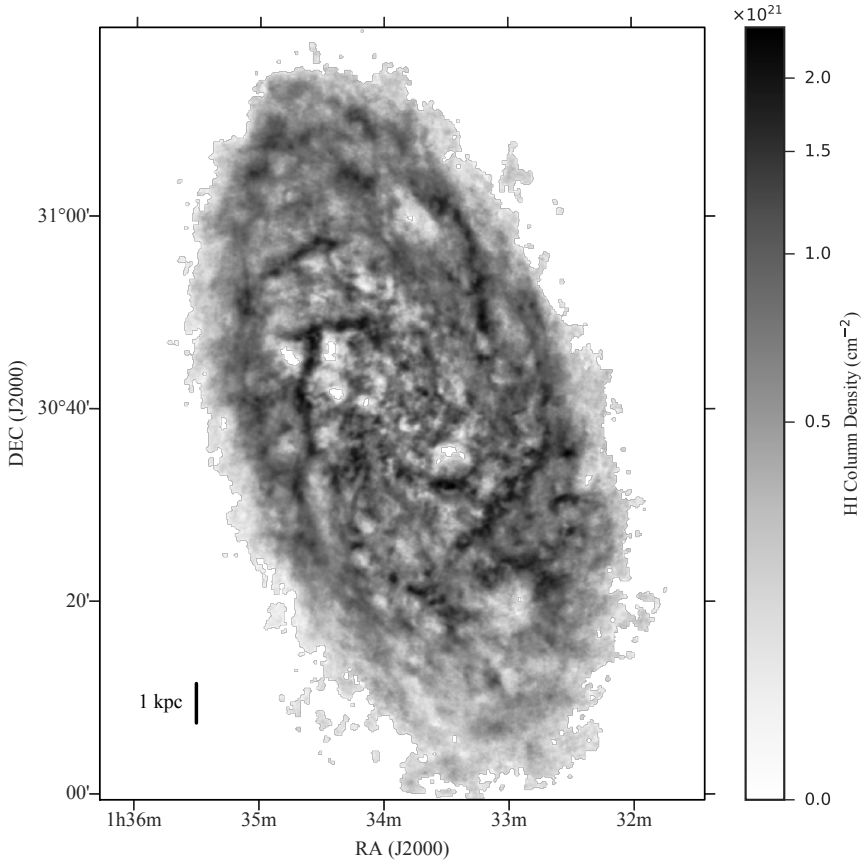

Figure 3.2: HI column density map of the VLA+GBT mosaic, assuming optically thin emission. The map is masked using the technique described in §3.3.3, and shown using an arcsinh stretch. The mosaic recovers emission well out to a galactic radius of 10 kpc.

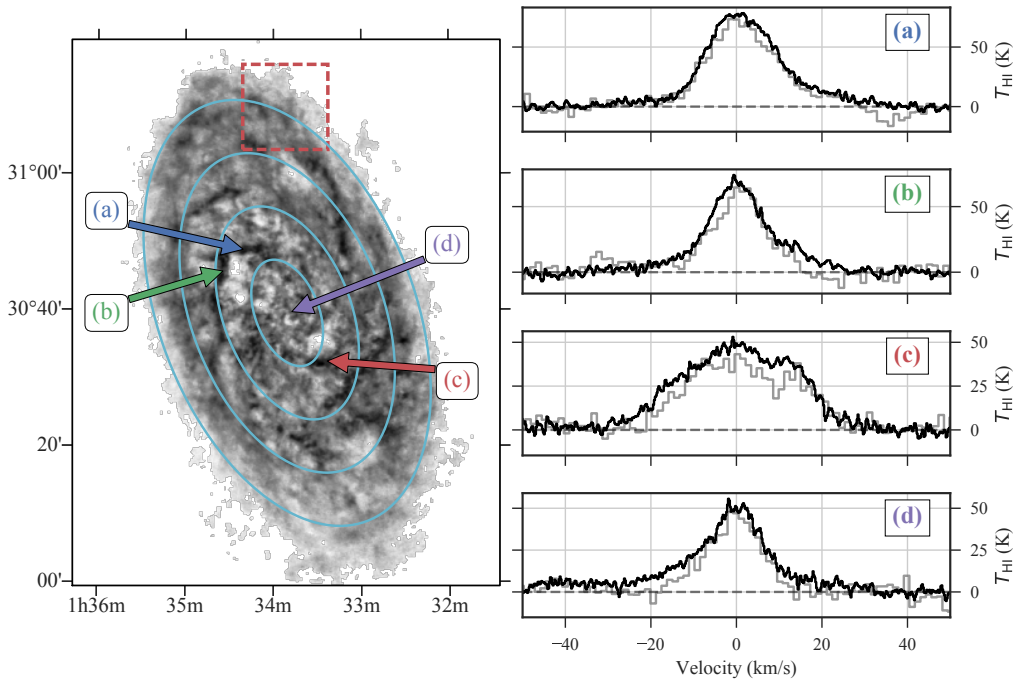

Figure 3.3: Example HI spectra compared between the archival VLA (gray) and the new VLA+GBT HI data (black). The spectra are centred at the velocity of peak intensity, calculated by smoothing with a 2 km s^{-1} Gaussian kernel. The location of the spectra are indicated by the arrows on the column density map in the left panel. The new data highlight the complex line shapes near the peaks and demonstrate that signal seen in the extended line wings is real. The red-dashed box on the left panel shows the position where a high-velocity cloud is found (§3.5.3) at a velocity of -30 km s^{-1} from the galaxy's rotation. The cyan contours on the map indicate galactocentric radii of 2, 4, 6 and 8 kpc, respectively, in the plane of the galaxy.

commonly referred to as the 'peak velocity'. To find $v_{\rm peak}$, we smooth the data with a 1 km s^{-1} ($\sim$ 5 channel) Gaussian kernel to minimize noise before identifying the velocity of the peak. In §3.8.2, we discuss the difference in using $v_{\rm peak}$ versus the centroid (first moment) velocity, $v_{\rm cent}$, in deriving the rotation curve. Briefly, $v_{\rm cent}$ becomes biased by asymmetric line wings (§3.4.3); whereas $v_{\rm peak}$ is not biased by the line shape and should provide a more accurate representation of the rotation velocity.

We fit the velocity surface using DISKFIT (Spekkens & Sellwood, 2007; Sellwood & Spekkens, 2015) with a circular velocity model. DISKFIT derives a global rotation model by simultaneously fitting the whole velocity surface over a given set of radial bins. This provides better constraints than minimizing a set of individual rings, but requires that the disk be characterized by a common set of disk parameters (e.g., position angle, inclination). The rotation curve is shown in Figure 3.5, and the fit parameters and statistics are given in Table 3.2. The rotation velocities in each bin are given in Table 3.6. The parameter uncertainties are calculated from 200 bootstrap iterations in DISKFIT (Sellwood & Sánchez, 2010). The χ^2 calculation in DISKFIT includes an ISM line width parameter for the expected model dispersion, which we set to 8 km s^{-1} (see §3.4.2). However, since this is a constant factor applied to every position in the fit, it will not affect the resulting disk parameters (Kuzio de Naray et al., 2012). Our rotation model does not include a warp component since a warp only becomes prominent at galactocentric radii beyond $\sim$ 8 kpc (Corbelli et al., 2014) and we do not fit for a bar component. We use radial bins with a constant width of 100 pc in the plane of the galaxy, just larger than the VLA beam size, out to a galactocentric radius of 7.5 kpc. Because the radial bins are larger than the beam, we do not include a correction for beam smearing. Beyond this radius, the uncertainty in the derived moments increases substantially as the integrated intensity decreases. The fitted position of the galactic centre is 7″ from the 2MASS position (Skrutskie et al., 2006), within the 1σ parameter uncertainties.

Figure 3.5 shows the circular rotation velocities, $v_{\rm rot}$, from our rotation

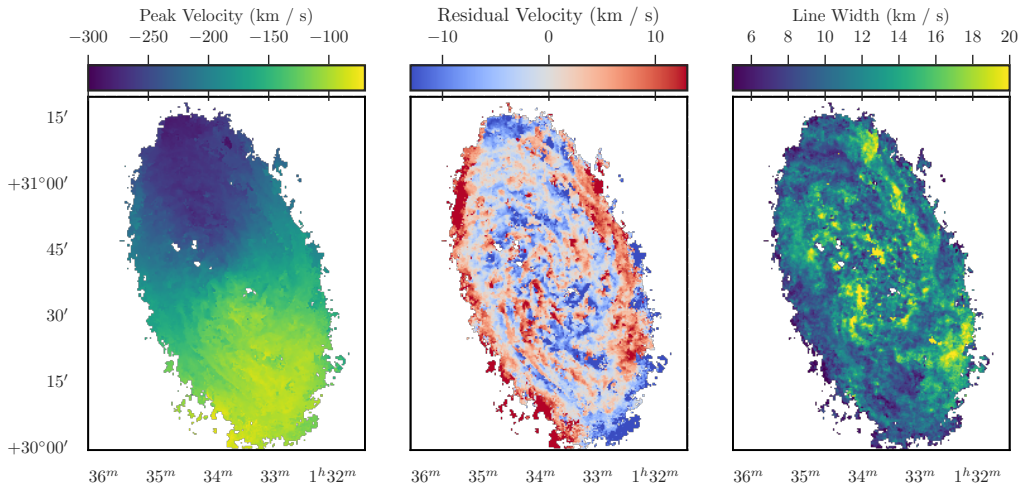

Figure 3.4: Left: HI $v_{\rm peak}$ surface used for fitting the rotation curve. The masked regions from the signal mask (§3.3.3) are shown in white. Centre: Residual velocities from subtracting the DISKFIT rotation model from the peak velocity surface. Right: The line width map derived from the second moment.

Table 3.2: Galactic disk parameters and fit statistics from DISKFIT for the VLA+GBT peak velocity surface. Errors are the 1σ intervals based on the DISKFIT bootstrapping. The n, $v_{\rm max}$, and $R_{\rm max}$ parameters are the fit parameters to the Brandt (1960) model (Equation 3.1).

Points used	39970
Degrees of Freedom (DOF) for error	39663
χ^2/DOF	1.44
Iterations	4
Centre R.A.	$23.4607 \pm 0.0042°$
Centre Dec.	$30.6583 \pm 0.0032°$
Position Angle	$201.12 \pm 0.47°$
Inclination	$55.08 \pm 1.56°$
$v_{\rm sys}$ (LSRK)	-179.18 ± 0.76 km s^{-1}
n	0.56 ± 0.04
$v_{\rm max}$	110.0 ± 1.5 km s^{-1}
$r_{\rm max}$	12.0 ± 1.3 kpc

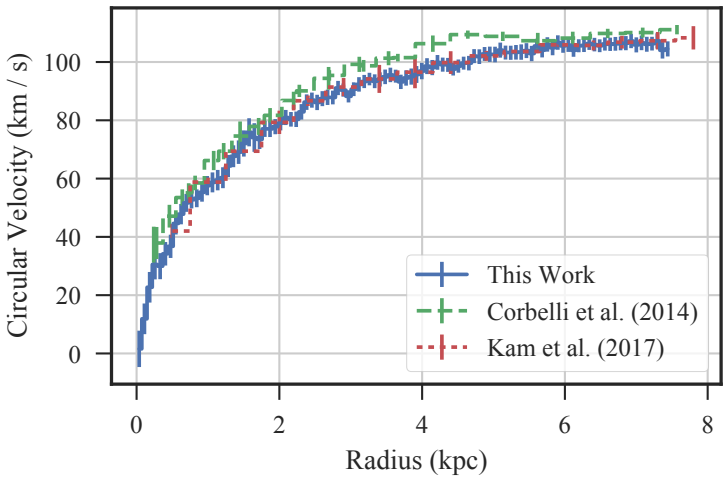

Figure 3.5: Rotation curve fit by DISKFIT from the VLA+GBT $v_{\rm peak}$ surface (blue solid). The rotation curves from Corbelli et al. (2014, green-dashed) and Kam et al. (2017, red-dotted) are also shown. Our model is consistent with the Kam et al. (2017) model, but has a lower velocity than the Corbelli et al. (2014) curve in the inner 6 kpc.

model, along with the recent HI models from Corbelli et al. (2014) and Kam et al. (2017). There is excellent agreement with the Kam et al. (2017) results, though our model has rotation velocities below the Corbelli et al. (2014) rotation curve up to 6 kpc. The difference may be due to how Corbelli et al. (2014) apply finite disk corrections (see their Appendix B), but since both the data and modelling approaches differ for the three rotation curves, it is difficult to be sure of the cause.

The residual velocity surface from the difference between $v_{\rm peak}$ and $v_{\rm rot}$ is shown in Figure 3.4. The residual velocity surface shows only small-scale variations, suggesting that additional velocity components in the rotation model are not needed. However, the outer edges of the mosaic show what is likely the beginning of the warp. Large deviations ($\sim$ 20 km s^{-1}) both above and below the model values are evident along the edges, similar to the deviations found by Kam et al. (2017).

To remove bin-by-bin variations in the fitted rotation curve we follow the prescription used by Meidt et al. (2008) by fitting a Brandt (1960) rotation

curve:

$$v_{\rm rot}(r) = \frac{v_{\rm max}(r/r_{\rm max})}{[1/3 + 2/3(r/r_{\rm max})^n]^{3/2n}}. \quad (3.1)$$

The fit parameters for the Brandt model are given in Table 3.2. Using the analytical approximation, we create a smooth version of the rotation surface for use as a model. This smooth rotation model surface is used for creating rotation-subtracted versions of the HI cube.

3.4.2 Surface density profiles

Using the galactic disk parameters in Table 3.2, we create a radial profile of the HI surface density, corrected for the disk inclination, in 100 pc radial bins out to 10 kpc. Assuming optically thin emission, we use a mass conversion factor of 0.019 $\rm M_\odot\,pc^{-2}\,(K\,km\,s^{-1})^{-1}$ to find the HI surface density, which is the standard HI column density conversion factor with a factor of 1.4 for the mass of heavier elements. Figure 3.6 shows the radial surface density profile of the combined VLA and GBT data. The profiles presented here use smaller bin widths than those in Druard et al. (2014), Corbelli et al. (2014) and Kam et al. (2017), but are consistent with all of these previous works.

The profile has an average value consistent with a surface density of ~ 8 $\rm M_\odot\,pc^{-2}$ out to 7 kpc, where the surface density begins to taper off. The radius where the tapering begins is comparable to where the stellar surface density equals the gas surface density, based on Figure 10 in Corbelli et al. (2014). Within this radius, the average stellar surface density exceeds the gas component.

When averaged over galactocentric radii of < 7 kpc, the atomic gas surface density is nearly constant, consistent with the surface density profile from Kam et al. (2017). However, averaging over smaller radii highlights the deviations driven by the large-scale disk structure. When averaging with the VLA-only data, the average value is ~ 6 $\rm M_\odot\,pc^{-2}$ for the inner 7 kpc, roughly consistent with the fraction of emission recovered without the GBT data.

The profile shows two significant peaks located around 2 and 4.5 kpc; these

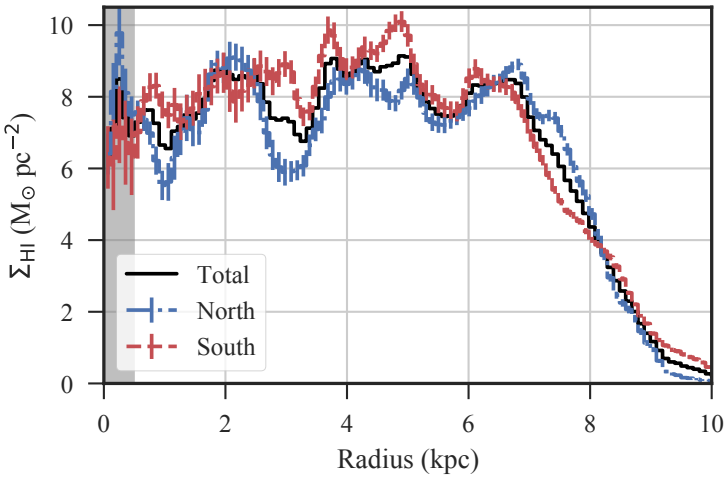

Figure 3.6: The combined GBT and VLA atomic gas surface density profile (black) split into the northern (red dot-dashed) and southern (blue dashed) halves in 100 pc bins and corrected for the disk inclination. The surface density has an average value of 8 $M_\odot$ pc^{-2} in the inner 7 kpc, with substantial variation from galactic structure. The error bars are shown for the northern and southern halves; the errors for the entire disk (not shown) are $\sqrt{2}$ smaller. These uncertainties are the standard deviations in each radial bin, corrected by the number of beams (i.e., the independent samples) within the annulus (e.g., Druard et al., 2014). The shaded region indicates the inner 0.5 kpc where beam smearing and a small number of samples in each bin gives large uncertainties.

occur at radii that enclose the prominent spiral arm structure. Within ~ 2 kpc, the disk lacks spiral structure and is dominated by wind-driven shells (Rosolowsky et al., 2007), making it difficult to connect these large-scale averages to the morphology of the emission. The outer disk also lacks significant coherent large-scale structure, with the emission dominated by multiple spiral arm fragments.

To further examine these variations in the surface density profile, we create radial profiles split into the northern and southern halves. The aforementioned peaks are far more prominent in the northern half than the southern. This is even more dramatic for the dip at 3 kpc, which is entirely driven by variations in the northern half. These differences are driven by the asymmetric main spiral arms: the northern arm, in HI emission, is more prominent and distinct than the southern arm.

The surface densities between the two halves are approximately equal between a radius of 1.5 to 2 kpc and beyond 5.5 kpc. The region from 1.5 to 2 kpc contains the beginning of the main spiral arms. The HI arms begin at ~ 1.8 kpc, are offset by nearly 180°, and are nearly symmetrical, despite tracing different extents at larger radii. The other matching region at 5.5 kpc is beyond the extent of the main spiral arms where the disk becomes dominated by flocculent structure. These matching regions over all angles suggest that this middle region of the disk is dominated by an arm driving mechanism and distinct from the inner and outer regions in the disk. However, the spatial scale of these variations may be different due to the large area averaged over at larger radii.

The outer region of the disk shows less variation between the halves compared to the inner and middle disk. While this region contains variations in the surface density due to the spiral arm fragments, the area averaged over is also larger and will tend to smooth out variations more than the regions at smaller radii.

3.4.3 Exploring spectral complexities

With the high spectral resolution in our data, there are a sufficient number of channels to use higher-order spectral moments for describing line profile shapes. Figure 3.3 demonstrates that simple spectral models (one or two Gaussians) are inappropriate to represent the data at this resolution. We quantify these features by calculating maps of the skewness and kurtosis of the velocity profiles. These higher-order moments require a larger number of channels than the line width (second moment) to reduce uncertainties, which has limited their use in previous extragalactic HI studies with coarser resolution. Instead, previous studies have relied on adding equivalent parameters to Gauss-Hermite polynomials (e.g., Young et al., 2003; Oh et al., 2011). However, this analytic form is limited in how well it can explain extended line wings, making the moments-based description better suited for our data.

The skewness is calculated from the third moment of each spectrum, normalized by the line width cubed:

$$S(x,y) = \frac{\sum_{v\in\mathcal{M}(x,y,v)}[v - M_1(x,y)]^3 I(x,y,v)\ \delta v}{M_0\sigma_v^3}, \tag{3.2}$$

where M_0 is the integrated line intensity (zeroth moment), and M_1 is the line centroid (first moment). The line width here is defined from the second moment: $\sigma_v^2 = M_2(x,y)$. A positive skewness indicates a spectrum with a red-shifted tail, while a spectrum with negative skewness has a blue-shifted tail.

Kurtosis is similarly calculated from the fourth moment. We report the *kurtosis excess* by subtracting the value of 3 expected for a normal distribution:

$$K(x,y) = \frac{\sum_{v\in\mathcal{M}(x,y,v)}[v - M_1(x,y)]^4 I(x,y,v)\ \delta v}{M_0\sigma_v^4} - 3. \tag{3.3}$$

A spectrum with a positive kurtosis will have strong tails and have a peak sharper than a Gaussian.

The skewness maps are shown in Figure 3.7 for the VLA-only and combined data. The addition of the total power component from the GBT data

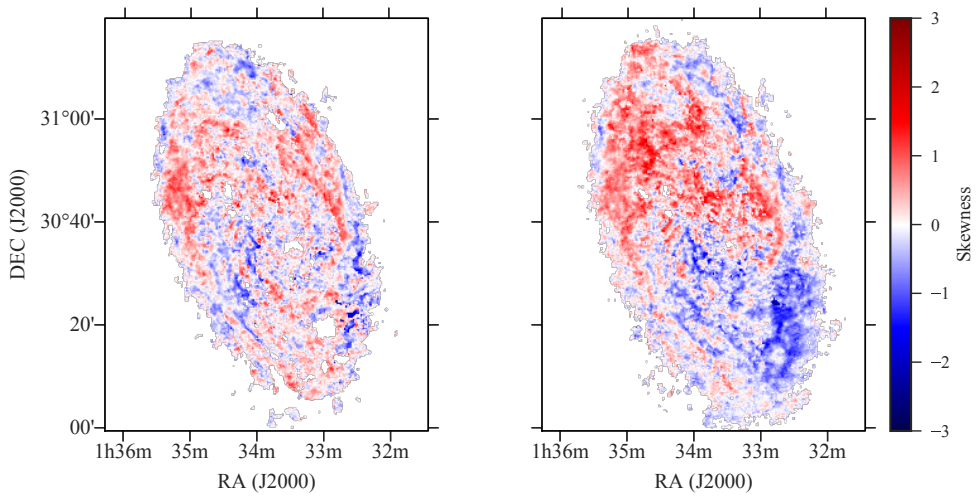

Figure 3.7: The skewness maps of the VLA (left) and combined data (right). Positive skewness indicates a red-shifted tail, while a negative value is a line profile with a blue-shifted tail. The low surface-brightness component, primarily added to the line wings, from the GBT data drives the large-scale skewness properties.

has a significant effect on the skewness. The VLA-only map shows small-scale structure across much of the disk, indicating variations of the spectral shape down to the resolution of the data. This map also shows some larger-scale variations, particularly for the outer disk in the northern half. With the inclusion of the GBT data, these large-scale variations become far more prominent. The skewness structure is split into a postively-skewed northern half and a negatively-skewed southern half. This separation is highlighted in Figure 3.8, where the radial profiles of skewness are shown for the northern and southern halves. The spectral shapes are skewed to the red-shifted side in the northern half and to the blue-shifted side in the southern half. This is consistent with the position-velocity slices in §3.5 that show a lagging rotational HI component across the disk.

The kurtosis maps are less interesting due to the influence of the line wings, and this drives a strong correlation to the peak temperature map. Figure 3.9 shows the close resemblance between the peak temperature and kurtosis maps for the combined data. The excess line wings bias the kurtosis, mak-

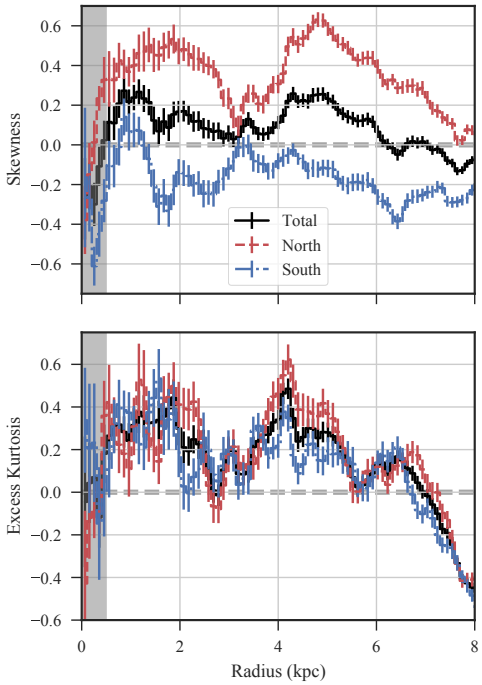

Figure 3.8: Radial profiles of the skewness (top) and kurtosis (bottom) of the VLA+GBT data averaged over the entire disk (black), and the northern (blue) and southern (red) halves. The shaded region indicates the inner 0.5 kpc where beam smearing and a small number of samples in each bin gives large uncertainties. The skewness profile demonstrates the differing skewness properties between the halves of the galaxy. The excess kurtosis indicates that the average HI spectrum has heavier tails relative to a Gaussian, though the bias in the second moment requires caveats for interpreting the kurtosis (§3.4.3.1).

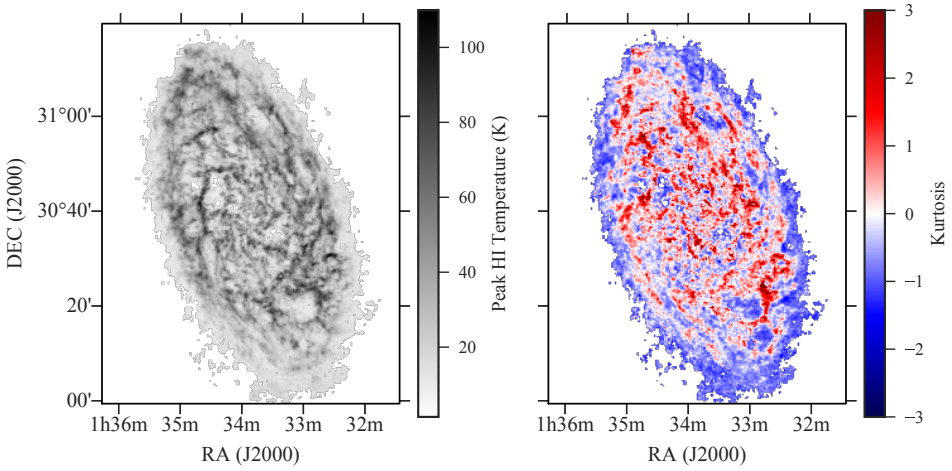

Figure 3.9: The peak HI temperature (left) and the kurtosis (right) maps of the combined data. A positive kurtosis indicates a shape with strong tails and more peaked relative to a Gaussian. There is a strong correlation between the two maps that results from the influence of the line wings on the kurtosis (§3.4.3.1).

ing it a measure of the spectral shape relative to the line wings rather than the bright HI component. A higher peak temperature then leads to a larger positive kurtosis. We discuss this relation further in §3.4.3.1. Since kurtosis is velocity-weighted by the fourth power, the assumption that line shapes are well-described by single-peaked near-Gaussian breaks down and it provides less-useful empirical relations compared to the skewness.

The skewness profiles in Figure 3.8 are nearly mirror images of each other in the northern and southern halves. By averaging over the entire disk, this distinction between the halves is lost, and the average values are near to a skewness of zero over most radii. This hides interesting variations seen separately in the halves, particularly near 3 kpc, where the average skewness in both halves is close to zero. In the surface density profiles (Figure 3.6), this radius corresponds to a dip in surface density, though it is more prominent in the northern half. This region contains a large portion of the northern spiral arm, including some of the more active star-forming regions (e.g., NGC 604). Apart from this, it is unclear why only this region has a skewness close to zero

and equal between the halves.

Though this asymmetry dominates both skewness maps, there are "limbs" in the outer disk that do not follow the same trend. The western side of the northern half in Figure 3.7 has a predominantly blue-shifted component along most of the map edge, while the eastern side of the southern half is mostly red-shifted. These features are nearly symmetric and are roughly pointed towards the position angle of the warped disk's semi-major axis ($\sim 165°$; Corbelli et al., 2014; Kam et al., 2017). The skewness "limbs" may indicate where the warped disk becomes more prominent than the main disk. Indeed, the northern limb has larger line widths relative to the surrounding structure in Figure 3.4. These enhanced line widths are also consistent with the line width map from Kam et al. (2017), which includes the structure at larger radii. Note that this does not contradict the lack of a warped disk component in the rotation model (§3.8.3) since the peak velocities will not be sensitive to extended tails.

The kurtosis profiles in Figure 3.8 do not show variation between the halves and are consistent within the uncertainties with the whole disk average at most radii. The positive kurtosis values indicate that the typical HI spectrum in the inner 6 kpc has strong tails that do not vary significantly in shape. The negative kurtosis beyond 7 kpc occurs where the brightest HI emission tapers off in the surface density profile (Figure 3.6). A negative kurtosis value in these regions indicates a lack of narrow and bright HI components.

We note that the profile values in the inner ~ 0.5 kpc are more uncertain due to the low number of spectra in these bins and may be the cause of the negative skewness slope out to 1 kpc. If the spectra were preferentially red-shifted at the centre, the PV-slices in §3.5.1 would show this asymmetry.

3.4.3.1 Correlations with skewness and kurtosis

In Table 3.3, we show the Spearman correlation coefficients calculated by comparing either the kurtosis or skewness with other HI properties. Most lines-of-sight have HI detected along them, giving a population of ~ 1.2 million

Table 3.3: Correlation between skewness and kurtosis with other HI properties. The uncertainties on each correlation are ± 0.002, using bootstrap resampling of one-third of the data to recompute the correlation 1000 times. We also perform a permutation test with 1000 iterations. All correlations have a p-value of < 0.001. The mask width is the number of pixels in the spectral dimension contained within the signal mask.

	Skewness	Kurtosis
Mask Width	0.019	0.408
Zeroth Moment	0.065	0.682
Peak Temperature	0.088	0.800
Centroid	−0.463	−0.155
Peak Velocity	−0.516	−0.162
Line Width	−0.033	−0.050

spectra to compare properties[5]. Since the distribution of these parameter values need not be Gaussian, we perform a bootstrap to calculate the uncertainty of the correlation and the p-value. In each bootstrap iteration, we randomly choose 30% of the sample and re-calculate the correlations; the standard deviation of the resulting distribution from all iterations gives the uncertainty. The p-value calculation uses a randomly shuffled version of one of the parameters to calculate the correlation with the other parameter. Both steps are run for 1000 iterations, and all correlation values are found to be significant with small uncertainties.

Skewness is strongly correlated with $v_{\rm cent}$ and $v_{\rm peak}$ for the combined VLA and GBT data. This is well-explained by the north-south asymmetry in skewness evident in Figures 3.7 and 3.8.

The kurtosis shows strong correlation with the peak temperature, zeroth moment, and the spectral width of the signal mask. We test whether these correlations are driven only by the shape of the signal mask by recalculating the correlations using points with similar mask widths (e.g., the correlation

[5]We do not correct for small-scale correlations over the beam size since the map extent is significantly larger. This will not affect correlations on large-scales.

between kurtosis and peak temperature only for points with a mask width of 300 to 320 pixels). We find that the correlations with peak temperature and the zeroth moment persist when controlling for the mask width, indicating the correlations are not solely dependent on the signal mask.

The correlations with kurtosis are instead the result of the line width bias due to extended line wings. As discussed in §3.8.5, the presence of line wings in the profile leads to an overestimate of the line width. This results from having a larger number of channels associated with the line wings than the bright component (e.g., spectrum d in Figure 3.3), and the kurtosis suffers from the same issue. The measured kurtosis is then with respect to the shape of the line wings, rather than the shape of the bright component. The positive kurtosis reflects the strong line wings present in most spectra relative to a Gaussian shape. Since peak temperature and zeroth moment are measures of how "peaked" that bright component is, they are positively correlated with the kurtosis. This highlights where the naive moments approach to examining complex line shapes breaks down and that the results should be treated with caution.

Neither skewness nor kurtosis shows a strong correlation with the line width. This is encouraging since both depend on the line width, with increasing power, and a strong correlation may imply the variations in skewness and kurtosis are only driven by changes in the line width. Our findings above also show that correlations with kurtosis are not driven by the mask shape. These results show that our measurement of significant non-zero values in the higher order moments is robust and that single-Gaussian descriptions of the HI line becomes inadequate at 80 pc resolution. However, since moment-based descriptions are subject to biases in the centroid, we further explore HI line profile shapes using stacked profiles below.

3.4.4 Stacking spectra

Stacking spectra is a common procedure to enhance the signal-to-noise in regions of faint emission. By aligning spectra to a common central velocity,

combining the spectra will coherently add emission components, while the noise will add incoherently. This technique has often been used for studying typical HI line properties in nearby galaxies.

We create stacked profiles using three different definitions of the line centre: the rotational velocity ($v_{\rm rot}$; §3.4.1), the centroid velocity ($v_{\rm cent}$), and the peak velocity ($v_{\rm peak}$). From the resulting stacked profiles, we can assess how the reference velocity affects the stacked profile shape. A cursory test with Gauss-Hermite polynomials, as used in several previous works (e.g., Petric & Rupen, 2007; Ianjamasimanana et al., 2012; Stilp et al., 2013a), gave poor results due to spectra with multiple bright HI components and those with extended line wings. The typical S/N in the data used in these studies prohibited directly estimating $v_{\rm peak}$, which is well-defined in our higher S/N data (see also Braun, 1997).

The different definitions of the line centre will highlight different ISM properties. Previous studies have typically stacked spectra based on $v_{\rm peak}$, which will minimize the stacked profile width and recover the typical spectrum properties at the resolution of the data (Caldú-Primo & Schruba, 2016b). This is ideal for attempting to recover the thermal and turbulent properties of the medium. Stacking with $v_{\rm rot}$ yields stacked profiles with widths sensitive to large-scale turbulence and non-circular motions. For example, variations from the circular rotation velocity are clear from the residual velocity map in Figure 3.4. Using $v_{\rm cent}$ as the common velocity gives, for an optically-thin tracer, the mass-weighted velocity average. As discussed in §3.4.3, this can make $v_{\rm cent}$ biased if the line profiles are asymmetrical. This bias in $v_{\rm cent}$ makes a physical interpretation less clear. Nonetheless, we explore these three stacking methods since they are commonly used in spectral-line studies.

We shift the spectra using Fourier transforms before co-adding them to create stacked profiles. Fourier shifting allows shifting by fractions of the channel width without interpolating, minimizing artificial broadening due to a finite-sampled grid. We then add the aligned spectra to create the stacked profile. The uncertainty in each channel of the stacked profile is the sum-

in-quadrature of the uncertainty from each PPV-pixel. Since the noise is approximately constant in our data, this reduces to the noise of one PPV-pixel multiplied by the square root of the number of beam elements in the stacked region. We include all of the spectra in the stacked profiles where $v_{\rm cent}$ and $v_{\rm peak}$ can be calculated after applying the signal mask (i.e., the signal mask contains at least one component along the line-of-sight). This excludes only the small masked regions in Figure 3.4.

To model the stacked spectra, we adopt the half-width-half-max (HWHM) scaling method from Stilp et al. (2013a). We favour this approach rather than other models due to the large variation in the shapes of the stacked profiles we find using different definitions of the line centre. In particular, we found that the two-Gaussian model used in other works (e.g., Young & Lo, 1996; Ianjamasimanana et al., 2012) cannot account for how peaked the $v_{\rm peak}$ stacked profiles are and the ratios between the two components are highly-sensitive to covariances between parameters. We provide a more thorough comparison in §3.8.4.

The HWHM method assumes that the stacked profile within the FWHM of the peak intensity can be modeled as a Gaussian line profile. Since these stacked profiles are created over large areas, the S/N is extremely high and the FWHM is a reliable measure of the width. Additional properties describing the stacked profile shape are defined based on the residual between the stacked profile and the assumed Gaussian profile for the peak. We use the definitions for four of these parameters from Stilp et al. (2013a). The effective Gaussian width, $\sigma_{\rm HWHM}$, is estimated by scaling to a Gaussian shape within the FWHM of the stacked profile. This assumed Gaussian shape is centred at the velocity of the maximum intensity, $v_{\rm centre}$. The line wings are defined as the excess relative to this Gaussian peak beyond the FWHM points. The fraction of excess flux is given as $f_{\rm wing}$:

$$f_{\rm wings} = \frac{\sum\limits_{|v|>\rm HWHM} [S(v) - G(v)]}{\sum\limits_{v} S(v)}, \tag{3.4}$$

where $S(v)$ is the stacked profile, $G(v)$ is the Gaussian model within the FWHM and HWHM = FWHM/2. The sum in the denominator is over the entire profile. The line wing excess can be used to define an effective width of the wings equivalent to the second moment:

$$\sigma^2_{\rm wing} = \frac{\sum\limits_{|v|>{\rm HWHM}} [S(v) - G(v)]\, v^2}{\sum\limits_{|v|>{\rm HWHM}} [S(v) - G(v)]}. \tag{3.5}$$

Note that the definition of the second moment relies on a Gaussian line profile. Since the residual $S(v) - G(v)$ is not Gaussian, this may not have a clear connection to a Gaussian line width. We include it here to compare with the values presented in Stilp et al. (2013a).

We define two additional parameters to describe the profile shapes. While higher-order moments present one avenue for characterizing deviations from a Gaussian shape, we favour developing empirical measures that do not weight the statistic by deviations from the line centroid. These metrics provide a more uniform description of the line profile over its entire extent. First, we define the asymmetry of a profile as the difference between the total flux at velocities smaller than the reference velocity and the total flux at velocities greater than the reference velocity. The difference is then normalized by the total flux over all velocities:

$$a = \frac{\sum\limits_{v>v_{\rm centre}} S(v) - \sum\limits_{v<v_{\rm centre}} S(v)}{\sum S(v)}. \tag{3.6}$$

The sum in the denominator is over the whole profile. Stilp et al. (2013a) define a different asymmetry parameter that, while similar, returns an absolute value. We choose this alternate definition since it retains information on the side of the profile containing more flux, which is determined by the asymmetry in the line wings. In this manner, a is analogous to the skewness, without weighting by the velocity offset and line width.

The final parameter, κ, provides the fractional difference between the central

peak and the Gaussian model within the FWHM, similar to $f_{\rm wings}$:

$$\kappa = \frac{\sum\limits_{|v|<\rm FWHM} [S(v) - G(v)]}{\sum\limits_{|v|<\rm FWHM} G(v)}. \tag{3.7}$$

This is another parameterization for the kurtosis: a negative value indicates the profile is more peaked than the Gaussian, while a positive value is flatter. We note, however, that the kurtosis used in §3.4.3 will primarily be sensitive to the line wing structure, whereas κ describes the shape of the peak since it is only measured within the HWHM relative to a Gaussian of equivalent width.

Figure 3.10 shows the stacked spectra over the entire disk with different reference velocities, and Table 3.4 gives the fitted model parameters. The parameter uncertainties are calculated in one of two ways. For σ and $v_{\rm centre}$, we assume that the uncertainty is set by the channel size; thus the uncertainty is half the channel width (~ 0.1 km s^{-1}). For the rest of the parameters, we perform 100 bootstrap iterations where 1) the data are re-sampled by adding Gaussian noise within the flux uncertainty for each channel, and 2) allowing the estimated peak velocity and Gaussian width to vary by their uncertainties. This propagates the uncertainty in the flux and the assumed Gaussian peak to the other parameters. Due to the high S/N in the stacked profiles, most of the uncertainty comes from the latter source.

Figure 3.10 shows how the use of different reference velocities significantly changes the stacked profile shapes. Stacking based on $v_{\rm rot}$ and $v_{\rm cent}$ gives the widest profiles and both have similar shapes. The shapes of their peaks are close to Gaussian ($\kappa \sim 0$; Table 3.4) and the widths are ~ 10 km s^{-1} in both. The widths of these profiles are dominated by the residual velocity in the $v_{\rm rot}$ model (Figure 3.4) and the bias in $v_{\rm cent}$ due to extended line wings evident in the skewness maps (Figure 3.7). Stacking using $v_{\rm cent}$ or $v_{\rm peak}$ gives profiles with non-zero a parameters (Equation 3.6). This may be highlighting an asymmetry in the low surface-brightness component of the line wings not evident in the position-velocity (PV) slices in §3.5.

The stacked profile from using $v_{\rm peak}$ is clearly non-Gaussian in its central

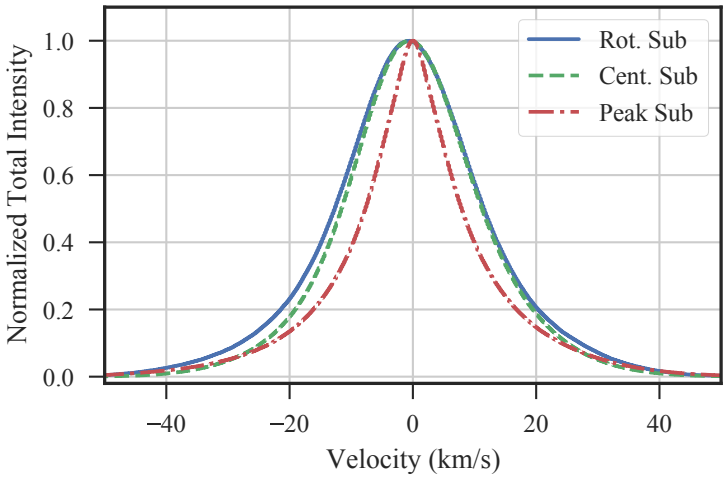

Figure 3.10: Stacked spectra over the entire disk using $v_{\rm rot}$ (blue solid), $v_{\rm cent}$ (green dashed), and $v_{\rm peak}$ (red dot-dashed) as the definitions of the line centre. The peak intensities are set to unity to emphasize the line shapes and the velocity axis is centred at the systemic velocity. Stacking based on $v_{\rm rot}$ and $v_{\rm cent}$ give similar profiles with near Gaussian peaks within the FWHM. However, using $v_{\rm peak}$ as the reference velocity yields a sharper peak shape and more prominent line wings relative to a Gaussian.

Table 3.4: HWHM fit values to the stacked profiles over the entire disk. These profiles are shown in Figure 3.10. The uncertainties are propagated based on the uncertainty in each channel of the profiles, and assuming an uncertainty of half the spectral resolution (~ 0.1 km s^{-1}) for the peak velocity and FWHM line width. Note that $\sigma_{\rm wing}$ may not be treated as an equivalent Gaussian width due to the non-Gaussian shape of the residuals; see Equation 3.5.

	Method of Stacking Spectra		
	Rotation Vel. ($v_{\rm rot}$)	Centroid Vel. ($v_{\rm cent}$)	Peak Vel. ($v_{\rm peak}$)
$\sigma_{\rm HWHM}$ (km s^{-1})	10.3 ± 0.1	9.7 ± 0.1	6.6 ± 0.1
$v_{\rm centre}$ (km s^{-1})	-0.5 ± 0.1	-0.5 ± 0.1	0.0 ± 0.1
$f_{\rm wing}$	0.11 ± 0.01	0.09 ± 0.01	0.26 ± 0.01
$\sigma_{\rm wing}$ (km s^{-1})	$28.9^{+0.5}_{-0.4}$	$26.3^{+1.5}_{-0.4}$	$23.4^{+0.3}_{-0.2}$
a	$-0.009^{+0.011}_{-0.010}$	$0.027^{+0.004}_{-0.003}$	$0.021^{+0.005}_{-0.014}$
κ	$-0.013^{+0.004}_{-0.004}$	$-0.011^{+0.004}_{-0.004}$	$-0.059^{+0.005}_{-0.004}$

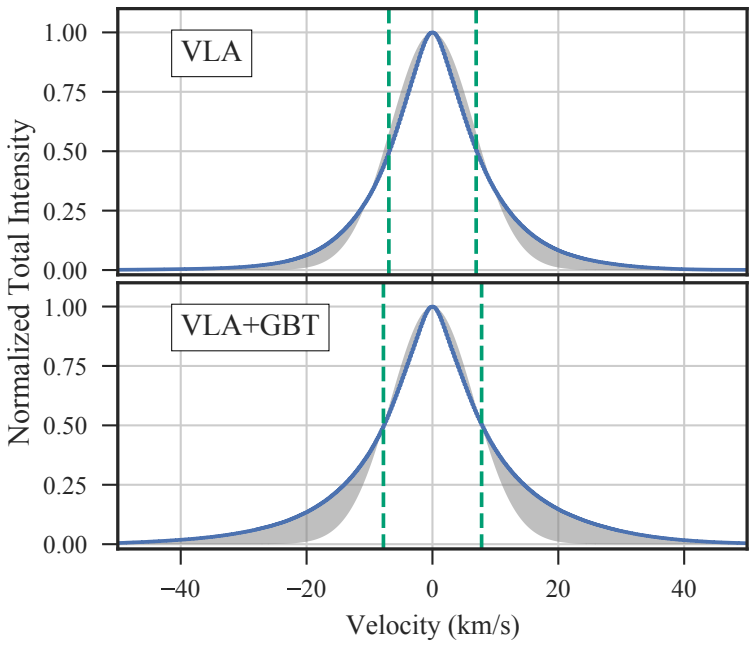

Figure 3.11: Stacked profiles of the VLA (top) and VLA+GBT data (bottom) when centered using $v_{\rm peak}$. The dashed green lines are the FWHM line widths and the shaded gray regions show the difference between the equivalent Gaussian profile and the data. The GBT data mostly adds emission to the line wings and does not affect the shape within the FWHM.

shape ($\kappa = -0.059$). The profile is both narrower and more peaked relative to the other profiles and is similar to the shape found in other studies (e.g., Petric & Rupen, 2007). The narrow and peaked profile contains a larger fraction of flux in the line wings relative to a Gaussian ($f_{\rm wings} = 0.26$). The difference in this stacked profile compared to those from $v_{\rm cent}$ and $v_{\rm rot}$ is the source of broadening in the width, which we discuss further in §3.6.3. If the broadening source has a Gaussian distribution of velocities around the peak, like the the residuals from $v_{\rm rot}$, the central-limit theorem gives that the shape will be close to Gaussian within the FWHM. When using $v_{\rm peak}$, the dominant source of broadening is the uncertainty on $v_{\rm peak}$ (~ 0.1 km s^{-1}), which is small compared to typical HI line widths.

The width and shape of the $v_{\rm peak}$ profile is relatively unchanged without the total power component. In Figure 3.11, we show the $v_{\rm peak}$ stacked profiles with and without the GBT data. Including the GBT data affects the line wings, which are far more prominent in the combined VLA+GBT data set. These differences highlight that most of the line wing emission is driven by a low surface brightness, extended component. This combination of a steep peak within the FWHM and heavy tails makes these shapes difficult to model. They are clearly non-Gaussian, but neither of these profiles has tails strong enough to be well-modeled by a Lorentzian.

The assumed Gaussian shape for sharply peaked profiles violates some of the assumptions for the HWHM model. The fractional difference in the peak shape, defined with κ, indicates that the assumed Gaussian shape overestimates the emission in the peak by 6% for the $v_{\rm peak}$ stacked profile. Though the two-Gaussian model also cannot account for the peak shape (§3.8.4), this highlights the difficulty in defining a simple model for the stacked profile shapes. We discuss the factors that lead to this difficulty in §3.6.3.

3.4.4.1 Radial variations in stacked profiles

We next examine how the stacked profile parameters vary with galactocentric radius, focussing on the $v_{\rm peak}$ stacked profiles. We create stacked profiles in

galactocentric rings with 100 pc widths, as used for the surface density profiles (§3.4.2). Given the opposing line wing directions found in the skewness map (§3.4.3), we also examine stacked profiles over the northern and southern halves separately.

Figure 3.12 shows the three profile parameters describing asymmetries (a) and shape (κ and $f_{\rm wings}$) for the $v_{\rm peak}$ stacked profiles; we explore the line width variations in §3.4.5. The asymmetry parameter a shows a clear difference between the two halves, matching the variations identified in the skewness analysis (§3.4.3). The southern half of the galaxy has an excess of emission at negative velocities while the northern half has an excess at positive velocities. This is qualitatively the same as is shown in the skewness radial profile in Figure 3.8, though we find less variation with radius in a. The decreased radial variation in a results from the σ^3 weighting in the skewness, which makes it sensitive to small line width variations, and from the difference in the order of operations, since stacked profiles are already an averaged quantity.

The shape within the FWHM of the stacked profiles, parameterized with κ, is largely consistent between the two halves. However, κ varies by $\sim 33\%$ with radius in the inner 6 kpc. There are two dips at 2.5 and 3.5 kpc where κ increases, indicating a larger discrepancy from a Gaussian. There are similar features in the kurtosis profile in Figure 3.8, however, those variations suggest that the typical spectrum becomes *closer* to a Gaussian profile. This apparent discrepancy may result from an increase in the fraction of multi-component spectra in these regions. These radii contain portions of the optical spiral arms, particularly the region around NGC 604 in the northern arm and the bright clump in the southern arm (see spectrum c in Figure 3.3). The kurtosis is driven by the shape with respect to the line wings (§3.4.3.1) and multiple components will give a smaller excess kurtosis since the bright component is effectively wider. We see the opposite effect in κ because the spectra are added coherently and multiple components will tend to enhance the line wings.

The fraction of flux in the line wings is on average $f_{\rm wings} \sim 0.22$ in the inner 6 kpc, but there are substantial ($\sim 50\%$) variations. There are peaks at

radii of 2.5 and 3.5 kpc, consistent with variations in κ being driven more by lines-of-sight having multiple spectral components. However, peaks in $f_{\rm wings}$ result from an increase in one half of the galaxy, whereas the variations in κ are consistent between both halves. The variations in $f_{\rm wings}$ cannot then be entirely driven by an increase in multi-component spectra. Another process, possibly outflows from star formation in the spiral arms, may lead to the increase in the line wing fraction.

None of the parameters show a monotonic trend with galactocentric radius in the inner 6 kpc, suggesting that the variations are due to local processes (i.e., outflows, local galactic structure).

A trend is clear, however, in all of the parameters beyond 6 kpc. The line profiles, over both halves, become more asymmetric ($a < 0$), have sharper and narrower peaks (κ and σ decrease) and have increased line wing fractions. A similar trend is seen for the kurtosis profile (Figure 3.8), though the kurtosis *decreases* for the reasons discussed above. At ~ 6.5 kpc, the surface density profile strongly decreases to the edge of the map (Figure 3.6). We attribute these trends to the lack of a bright HI component when averaged on large-scales. The change in asymmetry is due to the beginning of M33's warped disk, which starts to dominate the velocity surface at ~ 7 kpc, consistent with increases in the line width at the edge of the map (Figure 3.4). The variations we find beginning at 6 kpc suggests the warped disk has a significant effect on the spectral shapes before it dominates the large-scale kinematics.

3.4.4.2 Asymmetry between the northern and southern halves

The skewness maps in Figure 3.7 indicate a large-scale asymmetry in profile shapes between the northern and southern halves of the galaxy. By comparing the line wing excess in each half of the galaxy, we can estimate what fraction of $f_{\rm wings}$ is symmetric and asymmetric in velocity.

The $v_{\rm peak}$ stacked profiles in each half and over the entire disk are compared in Figure 3.13. The peaks of all three profiles are essentially identical in shape and width, but the line wings differ in each half. We can re-write the line wing

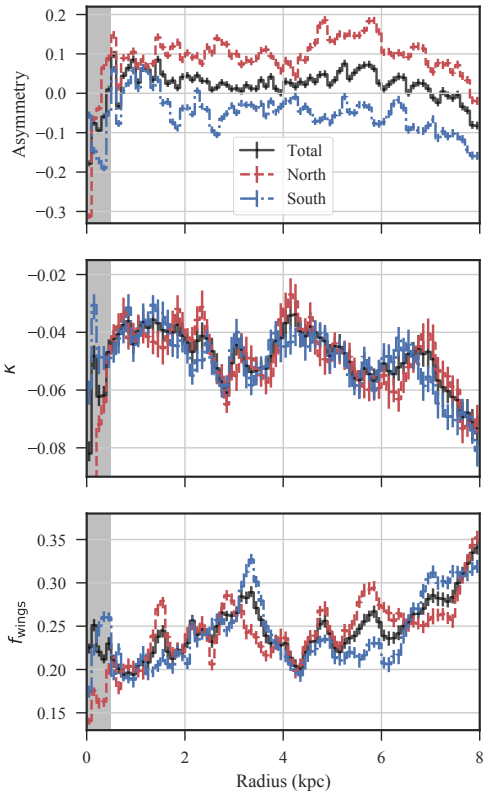

Figure 3.12: Top: Asymmetry, a, of v_{peak} stacked profiles in 100 pc radial bins. The parameter values are shown for stacked profiles over the whole disk (black), and the northern (blue dashed) and southern (red dot-dashed) halves. Middle: The peak shape κ. Bottom: The line wing fraction f_{wings}. The shaded region in both plots highlights the inner 0.5 kpc, where the lack of data points make the estimates more uncertain. The uncertainties in this region are underestimated since they do not account for the additional velocity uncertainty that results from spatial smearing where the rotation curve is steep (Stilp et al., 2013a). The top two panels are qualitatively similar to the skewness and kurtosis profiles from Figure 3.8: the asymmetry varies between the halves of the galaxy, but the peak shape is consistent between both.

fraction (Equation 3.4) as:

$$f_{\rm wing} = \frac{\sum\limits_{|v|>\rm HWHM} [S_{\rm N}(v) + S_{\rm S}(v) - G(v)]}{\sum\limits_{v} S_{\rm N}(v) + S_{\rm S}(v)}, \tag{3.8}$$

where $S_{\rm N}(v)$ and $S_{\rm S}(v)$ are the stacked profiles of the northern and southern halves, respectively. The asymmetric line wing component is the total excess on opposite sides of the peak, located at $v = 0$ km s^{-1} in this case. Based on the total profiles, the southern half will have an excess towards negative velocities, while the northern half will have an excess at positive velocities, both pointing toward the systemic velocity. Thus we define the asymmetric line wing component as:

$$f_{\rm asymm} = \frac{\sum\limits_{v<-\rm HWHM} [S_{\rm S}(v) - S_{\rm N}(v)] + \sum\limits_{v>\rm HWHM} [S_{\rm N}(v) - S_{\rm S}(v)]}{\sum\limits_{v} S_{\rm N}(v) + S_{\rm S}(v)}. \tag{3.9}$$

Note that the asymmetric component is not dependent on the assumed Gaussian peak $G(v)$, though the limits on the sums are set by the HWHM. The symmetric line wing component is the difference between the total and asymmetric components:

$$f_{\rm symm} = f_{\rm wing} - f_{\rm asymm}. \tag{3.10}$$

Because the asymmetric component is not defined as the absolute value of the difference, it is possible to have a negative value if the line wings excesses are opposite of the definition assumed based on Figure 3.13.

Table 3.5 shows the asymmetric and symmetric components of the total profiles from Figure 3.13. Relative to the total line wing fraction, we find that the $\sim$ 70% arises from a symmetric component while 30% stems from an asymmetric one. In both halves, the asymmetric component points toward the systemic velocity of M33 (since $f_{\rm asymm} > 0$) and could trace a rotationally-lagging disk component. We discuss this further in §3.5.

Figure 3.14 shows the symmetric and asymmetric line wing fractions calculated from profiles in 100 pc radial bins. The asymmetric component is largely

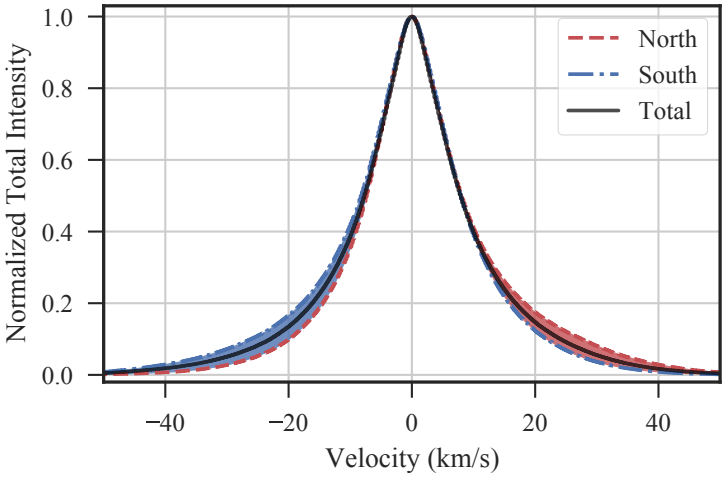

Figure 3.13: The v_{peak} stacked profiles over the whole galaxy (solid black), the Northern half (blue dashed), and the Southern half (red dot-dashed). The line wings asymmetry between the halves are highlighted with the shaded regions, where the colours correspond to the velocity the profiles are skewed towards.

Table 3.5: The line wing fractions from the peak velocity stacked profiles split into the Northern and Southern halves. The total line wing fraction is the same as given in Table 3.4.

	Line Wing Fraction
f_{wing}	0.26 ± 0.01
f_{symm}	0.18 ± 0.01
f_{asymm}	0.08 ± 0.01

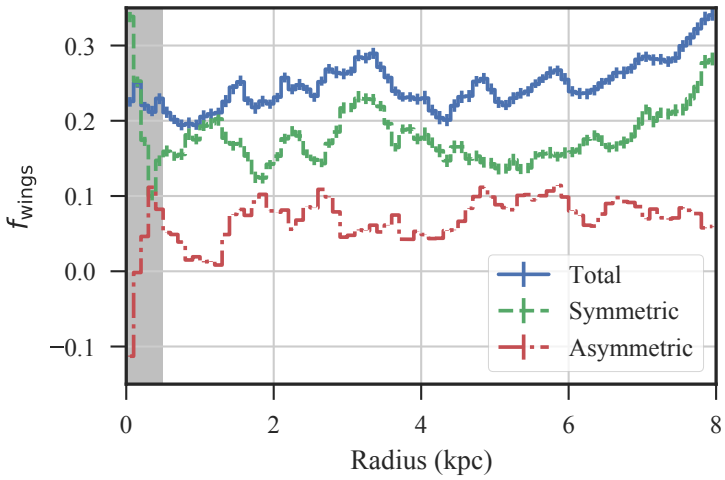

Figure 3.14: The total (blue solid), symmetric (green dashed), and asymmetric (red dot-dashed) line wing fractions determined from Equations 3.8 – 3.10 for $v_{\rm peak}$ stacked profiles in 100 pc radial bins. The inner 0.5 kpc is shown in gray, where the small area and beam smearing cause deviations in the profile properties. The asymmetric component accounts for $\sim$ 30% of the line wing fraction and is roughly constant throughout the disk. The variations in the total line wing fraction are mostly driven by variation in the symmetric component.

constant throughout the entire disk. The symmetric line wing component contains most of the variations previously seen in Figure 3.12 for the total line wing fraction. In particular, the increase in $f_{\rm wings}$ beyond 6 kpc is entirely due to the symmetric component.

3.4.4.3 Stacking profile with peak HI temperature

We also explore the properties of stacked profiles binned by their peak HI temperature. We created 20 stacked profiles in bins with a width of 5% of the peak HI temperature distribution. This provides an equal number of spectra (48900) in each stacked profile. As in the previous section, we consider the profiles stacked with respect to $v_{\rm rot}$, $v_{\rm cent}$ and $v_{\rm peak}$.

Figure 3.15 shows the line widths for each stacking method in 5-percentile bins. Similar to radially-binning the data (§3.4.5), we find an offset between the widths when using different definitions of the line centre, with the peak

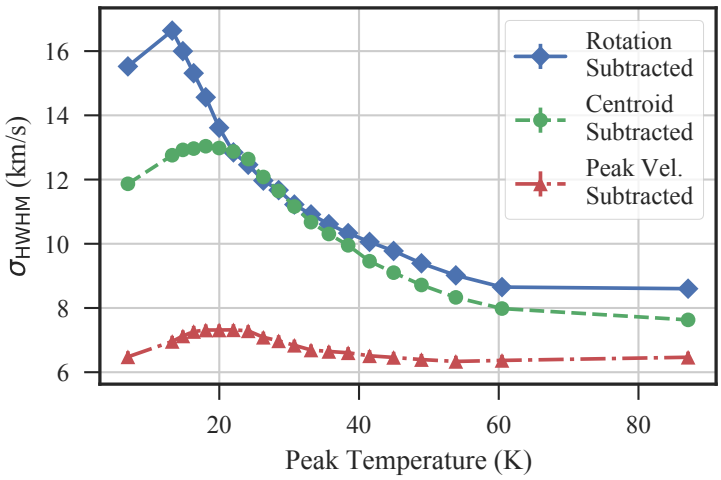

Figure 3.15: The HI line width of profiles stacked based on their peak HI temperature. Each stacked profile covers a range of 5% in the peak temperature distribution — the symbols in the plot indicate the middle of each bin.

velocity stacking providing the smallest widths. Each stacking method has a line width that decreases with increasing peak temperature, though the change in the widths of the $v_{\rm peak}$ stacked profiles is only ~ 1 km s^{-1}.

Despite the small changes we find in the $v_{\rm peak}$ stacked profiles over the range in peak temperature, the shape parameters change substantially. The line wing fraction $f_{\rm wing}$ monotonically decreases with increasing peak temperature, ranging from 0.17–0.32. The shape within the FWHM of the profile, described by κ, also increases with the peak temperature, with values ranging from -0.15 to -0.02. Note that this is the opposite behaviour from the radially-binned profile parameters (Figure 3.12). The combination of the trends in these two parameters and the near-constant line width demonstrates a shift in the profile shape that is near-Gaussian at higher peak temperatures and significantly non-Gaussian to low peak temperatures.

The profile shapes at low peak temperatures (< 20 K) show two distinct components: a narrow sharply-peaked component on top of a broad near-Gaussian component. Since, by area, most of the lines-of-sight with peak temperatures of < 20 K are located at $R_{\rm gal} > 7$ kpc, where the average surface

density begins decreasing (Figure 3.6), we interpret the increased importance of the broad component as arising from an increase in the proportion of emission from M33's warped disk component. This is consistent with our analysis of the skewness maps (§3.4.3), where the reversal in the sign of the skewness near the edges of the map in Figure 3.7 are roughly aligned with the position angle of the warped disk component (Corbelli et al., 2014; Kam et al., 2017). The increase in $f_{\rm wings}$ from the radially-binned stacked profiles in Figure 3.12 is also consistent with most spectra with low peak temperature (< 20 K) arising at $R_{\rm gal} > 7$ kpc.

We also find no clear evidence for HI spectra with flattened peaks in the highest peak temperature bins as would result from optically thick HI emission (Braun et al., 2009; Braun, 2012). If the stacked spectra had flattened tops, κ should decrease with increasing peak temperature, opposite of the consistent increase we find.

3.4.5 Comparing estimates of line width

Figure 3.16 shows the line width as a function of galactocentric radius as derived from the azimuthally-averaged second-moment line width map (Figure 3.4) and the stacked spectra aligned with $v_{\rm rot}$, $v_{\rm cent}$ and $v_{\rm peak}$. The enhanced line widths in the inner 0.5 kpc result from the small areas averaged over and from beam smearing where the rotation curve is steep.

Stacking by $v_{\rm peak}$ gives the narrowest profiles, consistent with the whole-disk stacked spectra in Figure 3.10. The typical width is ~ 7 km s^{-1} in the inner 6 kpc with a shallow gradient; beyond 6 kpc, the line widths decrease to ~ 6 km s^{-1}. Our line widths are consistent with the stacked profile widths found by Druard et al. (2014) using the archival HI VLA observations presented in Gratier et al. (2010).

The $v_{\rm rot}$ and $v_{\rm cent}$ stacked profiles also have similar widths (~ 10 km s^{-1}) to the whole-disk stacked profiles. The widths are consistent with each other throughout most of the disk, suggesting that the line wing bias in $v_{\rm cent}$ in the northern and southern halves of the disk are similar to the dispersion of

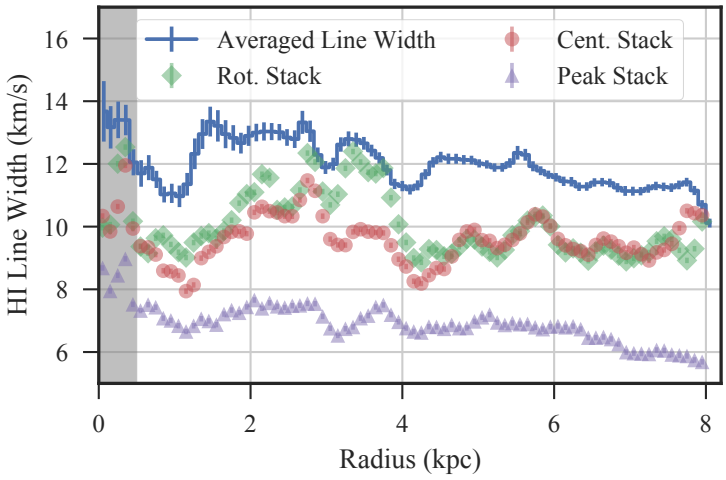

Figure 3.16: HI line width profiles in 100 pc bins for the azimuthally-averaged line width map (solid line; Figure 3.4) and the stacked profiles from $v_{\rm rot}$ (green diamonds), $v_{\rm cent}$ (red circles), and $v_{\rm peak}$ (purple triangles). The shaded gray region highlights the inner 0.5 kpc where the lack of samples and beam smearing cause large uncertainties in the line widths. The different methods give a large range in the HI line width. However, none of the methods shows a strong radial decrease.

velocity residuals from the $v_{\rm rot}$ model. At radii from 2 to 4 kpc, where the spiral arms dominate the galactic structure, the $v_{\rm rot}$ stacked profiles have larger widths than the $v_{\rm cent}$ stacked profiles. This indicates enhanced motion in the spiral arms due to multi-component spectra or coherent flows across the spiral arms, which are not included in the rotation model. Apart from the spiral arm region, there is no trend between $v_{\rm rot}$ or $v_{\rm cent}$ stacked line widths with radius.

We find that the azimuthally-averaged line widths from the second moment have a typical value of $\sim$ 12 km s^{-1} and are larger than all of the stacked profile widths. Since stacking based on $v_{\rm rot}$ yields the typical dispersion to non-circular motion and the $v_{\rm cent}$ is biased by asymmetric line wings, it is worrying that the second moment estimates are substantially larger. Previous studies have shown that the second moment can overestimate the line width relative to stacked profiles (Ianjamasimanana et al., 2012) and Gaussian fits to individual spectra (Mogotsi et al., 2016), particularly in the inner disks of nearby galaxies. The discrepancy we find shows that, when extended line wings are detected throughout the disk, the second moment line widths overestimate the line width. In §3.8.5, we find that the second moment line widths increase by 50% between the VLA and VLA+GBT data sets.

3.5 Extra-planar velocity components

Many nearby galaxies have velocity components offset from galactic rotation, either in the form of discrete HI clouds or extended line wings related to the main disk structure (Sancisi et al., 2008). Both features relate to neutral gas accretion onto the main disks and may play an important role in galaxy evolution. We explore the preferential directions of line wings, as shown in the skewness maps (Figure 3.7), in §3.5.1, and discrete HI sources offset in velocity from the main disk in §3.5.2.

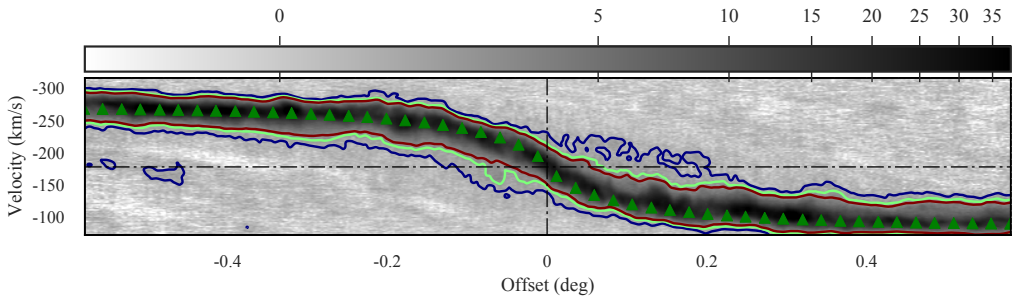

Figure 3.17: A position-velocity slice spatially-averaged over a 200″region centred along the major axis. The x-axis is the offset angle measured from the rotation centre of the disk (Table 3.2). The gray-scale, indicated by the colorbar, is in K, while the contours are at 1, 2, and 3 K, respectively. The green triangles are the velocities from the rotation model (§3.8.3). The horizontal dot-dashed line is at the systemic velocity, and the vertical dot-dashed line indicates the centre of M33. The region near the centre shows significant emission off the galactic rotation curve. The lowest contour highlights two features near the systemic velocity that are independent with emission from the main disk down to the noise level. Across the slice, the contours show faint emission that is preferentially skewed towards the systemic velocity (Table 3.2).

3.5.1 Anomalous velocity component

Figure 3.17 shows a position-velocity (PV) slice along the major axis of the disk from the North to South. Similar to Kam et al. (2017), we find several features either unassociated with the M33's main disk or lagging relative to the rotation velocity. The bright emission from the main disk, shown in gray-scale, is well-matched to the rotation model (green triangles). The contours highlight the extent of the low-surface brightness features. In particular, the 1 K contour, shown in blue, shows the asymmetry in the line wings directed towards the systemic velocity, consistent with what we find from individual spectra (§3.4.3). The extents of the blue contours in Figure 3.17 suggests that the line wing structure is similar on large-scales within the disk, implying that the line wings do not entirely arise from localized outflows from stellar feedback.

The blue contours in Figure 3.17 highlight asymmetrical features that point towards the systemic velocity. This indicates a rotationally-lagging disk com-

ponent that is consistent with gas on high-latitude orbits in a thick disk. The extensive excesses near the centre of the disk may be related to the halo structure. Kam et al. (2017) find extensive "beard"-like structures — velocity elongations in the PV-slice — in this region, which are more evident at lower surface brightnesses.

The excess intensity at lagging velocities extends across the disk, as is also indicated by the skewness of the line wings (Figure 3.7). Similar large-scale skewed profiles are found by Ianjamasimanana et al. (2012) for a sample of nearby galaxies in the THINGS survey (see their Figure 17). Their measure of skewness for individual profiles comes from Gauss-Hermite fitting, where higher-order skewness and kurtosis terms can be included directly in the fit. They find a clear bias in the skewness between the approaching and receding halves for a number of galaxies in their sample.

Similar to the asymmetrical "limbs" of opposing skewness along the edge of the map in Figure 3.7, Ianjamasimanana et al. (2012) find several other systems with similar features towards the outer regions of galactic disks. In §3.4.3, we suggest that these limbs are roughly aligned with the position angle of M33's warped disk component (e.g., Corbelli et al., 2014) and such features may arise from similar warped or lagging components in other nearby galaxies. For example, the direction of the asymmetrical regions found by Ianjamasimanana et al. (2012) in NGC 2403 is consistent with the position angle of its lagging rotational disk (Fraternali et al., 2001, 2002a).

Lagging disk components have been found in a number of nearby galaxies, including NGC 2403 (Fraternali et al., 2002a), NGC 4559 (Barbieri et al., 2005; Vargas et al., 2017), NGC 891 (Barnabè et al., 2006), and NGC 925 (Sancisi et al., 2008; Heald et al., 2011). As mentioned above, NGC 2403 provides an interesting comparison to M33 because of its flocculent structure. Fraternali et al. (2002a) suggest the lagging component has an origin internal to NGC 2403, rather than being primordial, due to the coherent structure and connection to the main disk. NGC 2403 also appears to have been undisturbed by dynamical interactions (e.g., Williams et al., 2013). Fraternali et al. (2002a)

suggests that the lagging component results from a galactic fountain mechanism, whereby hot ionized gas ejected into the halo from the fountain cools into atomic clouds that accumulate and fall back to the disk. Observations of ionized gas tracers confirm the presence of the lagging component (Fraternali et al., 2002b, 2004), consistent with the galactic fountain model. For a sample of nearby galaxies, Sancisi et al. (2008) also find that the extra-planar gas likely results from a galactic fountain, with a small component accreted from the intergalactic medium. Barnabè et al. (2006) find that the lagging component in NGC 891 is well-modeled by a baroclinic pressure model dependent on both the density and temperature of the medium. The model treats the medium as a mixture of a hot homogeneous component surrounding discrete cold HI clouds. From this, they reproduce the observed velocity gradient of the lagging component in NGC 891.

A lagging disk component from the galactic fountain mechanism may be related to cold HI "low-velocity clouds" (LVCs), which have been observed in the Milky Way. Stanimirović et al. (2006) find a population of small ($\sim$ few pc) LVCs with velocities of < 30 km s^{-1} relative to the Milky Way rotation velocity (see also Lockman et al., 2002; Stil et al., 2006). Adopting their typical column densities of clouds ($> 2 \times 10^{19}$ cm^{-2}), we compare this to the average column density of material at $v \sim 30$ km s^{-1} off M33 projected rotation speed ($< 2 \times 10^{18}$ cm^{18}). If the off velocity gas were organized into such clouds, it would have a reasonable filling factor of $f < 0.1$ and be broadly distributed across the disk. Our observations of this gas are consistent with an unresolved population of small clouds at high altitudes above the disk. The material could be raining back down as per a galactic fountain model, or the material could be directly accreting from the hot medium.

Only 1/3 of the line wing emission ($\sim$ 8% of the total atomic gas mass) is found in the asymmetric wing components, which would be identified as the "lagging" gas. The remaining 2/3 of the line wing emission in the symmetric component may require other explanations, especially since the wings are significantly wider than the thermal width of the warm neutral gas. One

possible explanation is neutral gas accretion from the halo. Recently, Zheng et al. (2017a,b) favour a model where $M \sim 10^8\ M_\odot$ of ionized gas in M33's halo is accreting onto the disk at $\sim$ 100 km s^{-1}. They suggest that these features result from the infall of material produced by a galactic fountain or past interactions between M33 and M31. A key observation in this analysis is the split between the symmetric and asymmetric components of the line wings (§3.4.4.2). Our symmetric line wings are consistent with this accretion flow.

3.5.2 Discrete HI clouds

Substantial HI structure has been found within a few degrees of M33 (Grossi et al., 2008; Putman et al., 2009; Keenan et al., 2016), including discrete features in an HI bridge that stretches towards M31 (Lockman et al., 2012). These structures are thought to comprise a population of high-velocity clouds (HVCs) or halo gas surrounding M33, similar to the HVC population in the Milky Way (Sancisi et al., 2008). Studies of these structures require excellent surface-brightness sensitivity and have mostly been limited to single-dish surveys at low resolution. Kam et al. (2017) provides the first resolved study of these features nearest to the main disk, including elongated velocity structures connected to M33's disk and discrete HI clouds offset in velocity. The data used by Kam et al. (2017) has a resolution of 120″, much better than previous deep HI studies.

We search for similar discrete HI structures within the spatial limits of our map. Though most of the HI emission in the aforementioned surveys is outside this area, we find significant extra-planar structure within our data. Figure 3.18 shows the extra-planar emission with integrated intensity maps in the northern and southern halves after subtracting the circular rotation model from the data cube. While most of the previously known HI clouds from single-dish studies fall outside of the area mapped our survey area, we find two structures that are consistent with the catalogue from Keenan et al. (2016).

We find a total atomic mass of $1.3 \pm 0.5 \times 10^7$ M$_\odot$ in the velocity-integrated

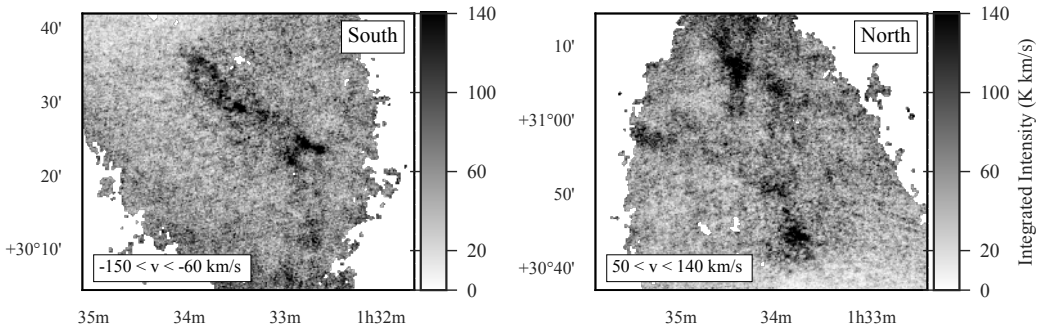

Figure 3.18: Integrated intensity maps of components displaced in velocity from M33's main disk in the southern half (left) and the northern half (right). The maps are integrated over the rotation-subtracted data cube; a velocity of $v = 0$ km s^{-1} is set at the systemice velocity of M33. The emission from the main disk is nearly all confined to velocities of -60 to $+50$ km s^{-1} (see §3.4.4). The left shows the emission structure blue-shifted from M33 (southern half), while the right shows the red-shifted components (northern half).

regions shown in Figure 3.18, which accounts for about 20% of the total HVC and disk-halo mass in M33 (Lockman, 2017), and $\sim 1\%$ of the atomic mass of the main and lagging disks. This mass is a factor of ~ 3 less than the estimated HVC mass in the Milky Way (Putman et al., 2012).

The left panel of Figure 3.18 shows the extent of the blue-shifted components and the right panel shows the red-shifted components. The atomic mass over these regions is $5 \pm 2 \times 10^6$ M$_\odot$ and $8 \pm 3 \times 10^6$ M$_\odot$ for the blue- and red-shifted components, respectively. In both cases, the extra-planar HI emission has significant small-scale structure, with features that were unresolved in previous studies of extra-planar gas in M33 (Westmeier et al., 2005; Grossi et al., 2008; Putman et al., 2009; Keenan et al., 2016). At 80 pc resolution, the structures are filamentary with 'blob'-like concentrations. This is particularly prominent for the blue-shifted components, which appear highly elongated.

The red-shifted integrated intensity shows four structures, one of which is partially cut-off by the edge of the map. The structure nearest to the centre of M33 is a portion of the lagging rotation component highlighted in Figure 3.17, shown by the deviation in the green contour at an offset of $-0.05°$. The last two structures near the top of the disk ($1^{\rm h}34^{\rm m}00^{\rm s}$ $+31°05'$) have been catalogued

by Grossi et al. (2008) as AA1 and Keenan et al. (2016) as AGESM33-1. These structures appear blended at the resolutions of both studies, but are shown here to have multiple faint filamentary structures surrounding the brighter centre. Keenan et al. (2016) report a mass of 1.18×10^6 $M_\odot$ for the blended structure, about 1/8 of the total mass we find over this velocity range.

A second structure identified as a HVC by Keenan et al. (2016), AGESM33-22, is within the spatial extent of our data but is too faint to be detected in our data at 80 pc resolution. The HVC, located at a velocity of -338 ± 15 km s^{-1}, has a column density of 7.2×10^{17} cm^{-2} at a resolution of $3.5'$, well below the sensitivity of our full resolution data.

The structure of the blue-shifted component is dominated by a single filamentary feature that stretches across most of the disk. Thus its mass is that of the blue-shifted region: $5 \pm 2 \times 10^6$ $M_\odot$. The portion nearest to M33's centre is comprised of two connected filamentary structures, forming a loop, that then extends further south. This feature has not been discussed in previous studies, though it is noted in Figure 7 of Sancisi et al. (2008) using the Thilker et al. (2002) archival VLA data. It can also be seen in Figure 3 of Putman et al. (2009) in their Arecibo data at velocities of -159 to -123 km s^{-1}. The Arecibo data indicate that the two components of the filament are indeed connected. The filament then has a projected length of ~ 8 kpc. Similarly, the velocity of the northern tip of the filamentary structure approaches the velocity of the lagging rotational component and the extended structure between 0–0.2° in Figure 3.17. Filamentary HI structures with similar lengths have been found in NGC 2403 and NGC 891 (Sancisi et al., 2008). Fraternali et al. (2001, 2002a) find a mass of 1×10^7 $M_\odot$ for the filament in NGC 2403, about twice the total mass we find in the blue-shifted off-axis emission in M33. Structures like this HI filament in M33 are strong evidence for accretion onto the main disk (Sancisi et al., 2008).

3.5.3 An HI cloud impacting the main disc?

An interesting HI cloud near the northeast edge of the map overlaps in velocity with the main disk, possibly indicating an interaction between an infalling cloud and the disk. We show the spatial structure of this cloud in the top panel of Figure 3.19. The morphology of the cloud is similar to the larger-scale clouds in §3.5.2, with a central blob-like structure with connecting filamentary structure. The prominent filamentary feature in the figure reaches toward the northern edge of the map and overlaps in velocity with the main blob. The blob is about 600 pc in diameter.

The bottom panel of Figure 3.19 shows the average spectrum over the cloud, where the velocity axis is defined with respect to the circular rotation velocity at $v = 0$. There are clearly two spectral components and possibly a third faint component around 60 km s^{-1}. We model the two bright components with Gaussians. The widths of the fitted components are 15.2 ± 0.5 km s^{-1} for the HI cloud and 9.5 ± 0.2 km s^{-1} for the disk.

Putman et al. (2009) show that this cloud overlaps in velocity with a portion of M33's warped disk, labeled as the "northern warp" in their Figure 3, outside of our survey area. What we detect as a separate HI cloud may be part of a larger structure related to the warped disk.

Treating the HI cloud as a discrete structure, we use the cloud's spectral component from Figure 3.19 with the total emission over the cloud to derive a mass of $6.5 \pm 0.5 \times 10^6$ M$_\odot$. This is similar to the total mass found in the blue- and red-shifted HI clouds in §3.5.2. With this cloud included, the total mass of off-rotation components becomes $\sim 2 \times 10^7$ M$_\odot$ and there are likely other fainter off-rotation HI clouds blended with the lagging rotational disk that will also increase this total mass. A full search of these features will require a more thorough spectral decomposition of the data.

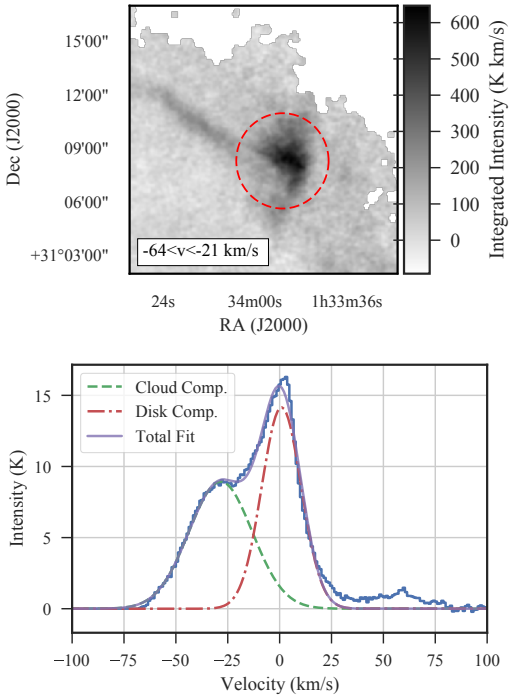

Figure 3.19: Top: An integrated intensity map highlighting the structure of the HI cloud. The red dashed box in Figure 3.3 shows this position relative to the rest of the galaxy. The velocity range is chosen to minimize emission from the main disk. Bottom: The average spectrum within the red-dashed circle in the top panel. The velocity axis is defined with respect to the $v_{\rm rot}$ model from §3.4.1. We fit two Gaussian components to the spectrum, shown as the green-dashed and red dot-dashed components.

3.6 Discussion

Measuring the kinematic properties of the atomic ISM is vital for evaluating and comparing to theoretical models. We compare two common techniques for estimating kinematic properties in extragalactic observations: the second moment of the line shape and spectral-stacking. Our results demonstrate that these common techniques are inherently limited and can yield vastly different results. For example, the method used to estimate the line width can yield values that differ by a factor of two (§3.4.5). In this discussion, we compare our results to those found in nearby galaxies and present a simple Gaussian Mixture Model for describing stacked profile shapes. This model recovers the qualitative profile shape found in previous studies, which we use to compare with previous model interpretations of stacked HI profiles.

3.6.1 Previous work and interpretations

Previous work on extragalactic HI spectra focuses on the HI line width, measured using the second moment (Tamburro et al., 2009), fitting Gaussian models to individual spectra (Petric & Rupen, 2007; Warren et al., 2012; Mogotsi et al., 2016), or stacking spectra over large spatial regions (Young & Lo, 1996; Braun, 1997; Ianjamasimanana et al., 2012; Stilp et al., 2013a). The first method is attractive since second moments are simple to compute, though it assumes the profile is close to a Gaussian. Tamburro et al. (2009) use line widths from the second moment for the THINGS survey and find typical values of 10 ± 2 km s^{-1} across the sample. Since these line widths are larger than the thermal line width expected for the WNM (~ 8 km s^{-1}), the profiles are interpreted as being broadened due to turbulent motions.

Modeling individual spectra by fitting Gaussian components has the advantage of measuring variations down to the resolution of the data, and being able to account for multiple components. Young et al. (2003) and Warren et al. (2012) find that most HI spectra can be modelled by up to two Gaussian components. Both studies find a range of line widths in both components,

suggesting emission components arise from both CNM and WNM phases. The narrow components have widths of < 6 km s^{-1} and are attributed to the CNM, with the line width broadened substantially from the thermal width by turbulence. However, Warren et al. (2012) do not find a clear connection between the location of narrow components and star-forming regions. Using a similar data set as Tamburro et al. (2009), Mogotsi et al. (2016) fit a single Gaussian model to all HI spectra where CO is detected. They find line widths of 10 ± 2 km s^{-1}, consistent with estimates from the second moment.

Alternatively, non-Gaussian line profiles have been modeled using Gauss-Hermite polynomials, which can account for variations in the skewness and kurtosis. While this method often provides a good analytical description, the connections between the model parameters and the underlying physical properties of the medium are unclear (Young et al., 2003). Young & Lo (1996), Young et al. (2003), Ianjamasimanana et al. (2012) and Stilp et al. (2013a) use this model to estimate the velocity at the peak intensity in their data, which are then shifted to a common centre and stacked.

The HI stacking method has the advantage of vastly increasing the S/N in the data, allowing low surface brightness emission to be studied at the expense of removing spatial variations. Nearly all studies find the same basic stacked profile shape: a near-Gaussian peak with enhanced line wings. This shape is not well-described by physically-motivated profile shapes. The enhanced line wings lead to poor single Gaussian fits, yet they are not prominent enough to match a Lorentzian profile. This has led to different modeling approaches with differing interpretations of the results. Young & Lo (1996) model the stacked profiles as a wide and narrow Gaussian component. These two components have widths of 9 km s^{-1} and 3 km s^{-1}, respectively. Since these are similar, though slightly larger, than the expected thermal line widths for WNM and CNM phases, they attribute the model as separating out the two different phases, with some broadening from turbulence. A larger sample of 34 galaxies from the THINGS survey was analyzed by Ianjamasimanana et al. (2012) using the same two-Gaussian model; the sample overlaps with the independent

analyses by Tamburro et al. (2009) and Mogotsi et al. (2016). They find similar results to Young & Lo (1996), with narrow components arising from the CNM with Gaussian widths of 3.4 to 8.6 km s^{-1}, and wide WNM components with widths from 10.1 to 24.3 km s^{-1}. These ranges imply that both the CNM and WNM must be highly turbulent since the line widths are much larger than the expected thermal line widths of these phases. We note that, for comparison with our results, the data from the THINGS survey has a coarser spectral resolution compared to our data, with channel sizes from 1.3 to 5.2 km s^{-1}.

Petric & Rupen (2007) propose an alternate interpretation of the enhanced line wings in stacked profiles. They argue that the flux ratio of the narrow component, if exclusively tracing the CNM, to the wide component should steeply decline with galactocentric radius, matching the observed star formation activity. Instead, they find that the ratio stays roughly constant to large radii. They attribute the central component to a mixture of atomic gas, with enhanced line wings that describe bulk motions, possibly from stellar feedback or infalling halo gas. Stilp et al. (2013a) test this interpretation by modelling the stacked profiles as a single Gaussian peak with line wings measured by their fractional increase relative to the Gaussian. For a sample of 24 dwarf galaxies, they find line widths ranging from $\sim 6-10$ km s^{-1} and $f_{\rm wings} \sim 0.05$ to 0.2.

An alternative model to the previous studies is introduced by Braun et al. (2009) and Braun (2012) where spectra with flattened tops are attributed to a cool HI layer sandwiched between two warmer layers (Braun, 1997). It is difficult to make comparisons to this model without accounting for multi-component spectra when calculating the line width. We defer further discussion of this model to future work.

Most methods have also been used to study radial trends in the HI properties. Tamburro et al. (2009) examine radial variations in the HI line width, finding that the line width tends to decrease rapidly in $R_{\rm gal}$ to R_{25}[6], similar to the radial decline in the star formation rate surface density. The slope beyond

[6]Defined as the semi-major axis of the 25 mag arcsec^{-2} isophote in B-band.

this radius in some galaxies flattens significantly. Stilp et al. (2013b) and Ianjamasimanana et al. (2015) examine radial trends by creating stacked profiles in radial bins. They find the same qualitative result despite using different models: the line width tends to decrease with radius. This trend is found in most nearby galaxies, though with three exceptions: NGC 2403, NGC 2976, and NGC 628 (Ianjamasimanana et al., 2015; Mogotsi et al., 2016). These systems have a shallow radial decrease in the line width throughout the disk. The first two galaxies are flocculent spirals similar to M33. NGC 628 is classified as a grand spiral galaxy, though its mass and star formation rate is within a factor of 2 compared to M33 (Walter et al., 2008).

These studies cover the HI properties of most nearby galaxies, spanning a range of galaxy properties and types. Young & Lo (1996), Young et al. (2003), and Stilp et al. (2013a) restrict their sample to dwarf galaxies, where the low-metallicity properties and lack of large-scale galactic structure may lead to differences in the ISM properties compared to the canonical view of the ISM. Tamburro et al. (2009), Ianjamasimanana et al. (2012), and Mogotsi et al. (2016) include a mix of galaxy types in their sample. Ianjamasimanana et al. (2012) find no difference in the HI line widths of stacked profiles between dwarf and spiral galaxies. Furthermore, Caldu-Primo et al. (2013) analyze the radial line width trends in the HI and CO for a similar set of nearby galaxies and find no common radial variation. These results are consistent with all stacked profiles in these studies having the same qualitative shape (Petric & Rupen, 2007), suggesting that the physical mechanism governing typical line profiles is similar regardless of galaxy type.

Druard et al. (2014) present an analysis of HI stacked profiles in M33, using the archival VLA HI data presented in Gratier et al. (2010). In this work, they emphasize comparisons of the HI to stacked profiles of CO(2-1). In 1 kpc radial bins, they find an average HI stacked profile width of ~ 6.5 km s^{-1} by fitting a single Gaussian to the stacked profiles. The line width decreases slowly with radius — a decrease of just ~ 2 km s^{-1} between the bins at $R_{\rm gal} = 0$–1 kpc and the 7–8 kpc bin — similar to the three galaxies noted above.

3.6.2 Key results in M33

Our results are broadly consistent with previous studies. In particular, we find a line width of 6.7 ± 0.2 km s^{-1} from the $v_{\rm peak}$ stacked profile, consistent with the M33 study from Druard et al. (2014). The line widths of the stacked profiles over radial bins, ranging largely between 5.5 and 7.5 km s^{-1}, are also consistent with Druard et al. (2014). This is encouraging since: (1) the stacked profiles of the archival HI data should have a similarly high S/N as our data, with the only major difference being the lower spectral resolution of the archival data (1.2 km s^{-1}), and (2) a different methodology is used in parameterizing the line shapes. The stacked profiles between the two studies also have a consistent shape, with a kurtosis excess in the central peak and prominent line wings. The kurtosis excess is similar to that found by Braun (1997) and more extreme than other recent studies (e.g., Ianjamasimanana et al., 2012; Stilp et al., 2013a).

The line widths from the $v_{\rm peak}$ stacked profiles are near the lower limit of the range of line widths found by Ianjamasimanana et al. (2012) and Stilp et al. (2013a) for nearby galaxies (see §3.8.4). However, since these profiles have non-Gaussian shapes, we note that the equivalent Gaussian width should be treated as an upper limit on the line width. While the HWHM-scaling method from Stilp et al. (2013a) is appealing because it does not require fitting an analytical model, it still has limitations on how well the profile shape can be described. We discuss the source of these difficulties in §3.6.3.

There are some discrepancies between the range of values we find for the other HWHM model parameters compared to the set of dwarf galaxies in Stilp et al. (2013a) and Stilp et al. (2013b):

1. We find consistently larger $f_{\rm wing}$ values, ranging from 0.25 to 0.35 in radial bins, than the 0.1 to 0.25 range in the Stilp et al. (2013b) sample (Figure 3.12). This range is, however, consistent if we only consider the symmetric line wing fraction (Figure 3.14). Thus the symmetric line wing component, possibly arising from inflowing or outflowing material

(§3.5), is similar with those measured in nearby dwarfs and the difference is due to the presence of a lagging rotational disk in M33.

2. Stilp et al. (2013b) find a large range of asymmetries when stacking spectra in radial bins[7]. We only find large asymmetries in M33 when stacking over radial bins in the northern or southern half, respectively (Figure 3.12). The shape within the HWHM of the stacked profiles in Stilp et al. (2013b) appear close to symmetrical (see their Figure 4), and most of the asymmetry occurs in the line wings.

The estimated line widths from the second moment have an average of ~ 12 km s^{-1}, consistent with the upper limit of the range found by Tamburro et al. (2009). These widths are 50% larger than largest widths of the stacked profiles (Figure 3.16). In §3.8.5, we demonstrate that the second moment is sensitive to the line shape, and can drastically alter the second moment line width when extended line wings are present. Thus, we discourage the use of the second moment for estimating line widths. Similar differences between other line width methods and the second moment are noted in several studies (Ianjamasimanana et al., 2012; Stilp et al., 2013a; Mogotsi et al., 2016).

3.6.3 Mixture models for stacked profiles

As we explain in §3.6.1, the shape of stacked profiles has led to different modelling approaches and different interpretations of their meanings. To illustrate why stacked profiles have shapes that are difficult to model using simple analytic forms, we invoke a Gaussian mixture model as a framework for understanding the line wings and widths. We consider the case where a stacked profile $S(v)$ is an average over an ensemble of Gaussian line profiles with purely thermal line widths:

[7]Note that the asymmetry parameter in their study returns the fraction of asymmetry, bounded between 0 and 1:
$a_{\rm full} = \frac{\sum_v \sqrt{[S(v)-S(-v)]^2}}{\sum S(v)}$

$$S(v) = \sum_i w_i \exp\left[-\frac{(v - v_i)^2}{2\sigma^2_{\mathrm{T},i}}\right] \tag{3.11}$$

where w_i is the weight of each spectrum, v_i is the central velocity and σ_T is the thermal line width for temperature T, i.e., $\sigma_\mathrm{T} = \sqrt{kT/\mu m_\mathrm{H}}$. The weights are normalized such that $\sum_i w_i = 1$. We can describe the shape of the stacked profile $S(v)$ based on the moments of the distribution. We use the derivation from Wang & TaafMixturefe (2015) to describe the first four moments. The centroid (first moment) of the stacked profile is $v_0 = \sum_i w_i v_i$, which depends only on the distribution of the line centre v_i. The variance (second moment) of the stacked spectrum is then

$$\sigma^2 = \sum_i w_i(\sigma^2_{\mathrm{T},i} + v_i^2) - v_0^2 \tag{3.12}$$

and the skewness (third moment) and kurtosis (fourth moment) are respectively

$$\mathrm{skew} = \frac{1}{\sigma^3}\sum_i w_i(v_i - v_0)\left[3\sigma^2_{\mathrm{T},i} + (v_i - v_0)^2\right] \tag{3.13}$$

$$\mathrm{kurt} = \frac{1}{\sigma^4}\sum_i w_i\left[3\sigma^4_{\mathrm{T},i} + 6(v_i - v_0)^2\sigma^2_{\mathrm{T},i} + (v_i - v_0)^4\right] - 3. \tag{3.14}$$

The critical result of these moments is that the stacked profile will depend on the distributions of the line centre and width of the individual spectra that were averaged over. For example, the shape of the stacked profile could be reproduced by adopting a reference probability density function for $w_i = f(T, v_i)$ encoding the amount of material found at a given temperature and velocity offset from the reference velocity.

We investigate how the common features in stacked profiles found in different studies can be attributed to the dependence on the distributions of the line centre and width of individual spectra that comprise the stacked profile.

1. Velocity offsets contribute to the line width. — Equation 3.12 has an equal dependence on the line width of the spectra and the distribution of their central velocities. The differences between the v_rot, v_cent, and v_peak

stacked profile widths is due to this dependence (Figure 3.16). Stacking based on $v_{\rm peak}$ minimizes the variation of the central velocity to within the uncertainty of $v_{\rm peak}$ (0.1 km s^{-1}). It represents the optimal recovery of the mixture of the line widths. The $v_{\rm cent}$ and $v_{\rm rot}$ stacked profiles are instead dominated by the distribution of the central velocities, not the distribution of the line widths. The purpose of the different approaches is discussed further in §3.6.6.

2. Skewness depends on the distribution of line centres. — From Equation 3.13, an asymmetric stacked profile can only occur if the distribution of line centres is also asymmetric since the sign of the skewness is determined by the $(v_i - v_0)$ term. Stacked profiles using $v_{\rm rot}$ or $v_{\rm cent}$ are both close to symmetric about $v = 0$, but $v_{\rm cent}$ can be biased by the asymmetric line wings. To get asymmetric stacked profiles from the former two line centres, an additional component (i.e., the lagging disk) must be present in the line profiles.

3. Stacked profiles of the atomic medium almost always have a line wing excess. — The kurtosis from Equation 3.14 has a strong dependence on the distribution of line centres and the line widths. Since the terms all have even powers, non-Gaussian features in either distribution will always increase the kurtosis, causing the mixture model to have an excess in the tails. For the atomic medium, the distribution of temperatures, which only affect $\sigma_{T,i}$ will be non-Gaussian since it will contain a mixture of CNM and WNM temperatures, with some fraction in the unstable intermediate regime. Thus, without requiring a separate physical process to produce line wings, stacked profiles of the atomic medium will always have an excess in the line wings. The only case where the kurtosis will decrease is if the majority of the spectra have multiple line components. This case is not explicitly handled in this toy model, but would tend to broaden the profile and flatten the peak, resulting in a kurtosis deficit.

Without requiring additional physical processes, this mixture model can

qualitatively reproduce properties of observed stacked profiles based only on a set of Gaussian components. This accounts for the near universal stacked line shape noted in Petric & Rupen (2007) and Stilp et al. (2013a) when stacking based on the peak velocity.

By adopting simple distributions for the line centres and widths, we test whether the mixture model of thermal-Gaussian components can produce simulated stacked profiles with the quantitative properties of observed stacked profiles. We find that reasonable distributions of HI temperatures creates narrower stacked profiles with smaller line wing fractions than the observed stacked profiles. We simulate a stacked profile from the mixture model by sampling 10^4 spectra from a uniform temperature distribution ranging from 10^2–10^4 K. We also sample the central velocities from a Gaussian distribution with width $\sigma = 1$ km s^{-1} to compare with the $v_{\rm peak}$ stacking in our data. The resulting stacked profile is similar in shape to the $v_{\rm peak}$ stacked profiles, with a sharp peak and enhanced line wings. Using the same model used for the stacked profiles, the mixture model profile has a width of $\sigma = 4.8$ km s^{-1}, $\kappa = -0.03$ and $f_{\rm wings} = 0.09$. When using a wider central velocity distribution to compare with the $v_{\rm rot}$ stacked profiles, we find that the mixture model also has a narrower width and smaller $f_{\rm wings}$.

We create another simulated stacked profile, but alter the temperature distribution to range from 10^3–10^4 K. This increases the width of the profile to $\sigma = 5.2$ km s^{-1}, but makes the shape closer to a Gaussian than the observed profiles, with $\kappa = -0.01$ and $f_{\rm wings} = 0.06$. This suggests that a cold atomic component in the stacked profiles is responsible for the high kurtosis excess in observed stacked profiles.

The mixture model demonstrates that the qualitative stacked profile shape can be reproduced from a mixture of Gaussian components and a range of plausible thermal temperatures. This highlights the difficulty in using simple analytical models for describing stacked profiles. We find that observed stacked profiles are wider and have more prominent line wings than the simplified mixture model presented here (Figure 3.3). These differences must arise

from a combination of the CNM/intermediate/WNM fractions, the lagging rotational disk, multi-component spectra, and turbulence. We conclude that stacked profiles reflect this full set of physical effects in M33, rather than being solely the distribution of CNM/WNM temperatures. In particular, there must be significant contributions from a multi-modal distribution of line centres (e.g., from multiple components) to explain the profiles. Because our data can resolve multiple velocity components both spatially and spectrally, we will revisit this analysis in future work through the multi-component fitting to our data, mirroring the approaches used in Galactic studies (e.g., Lindner et al., 2015; Henshaw et al., 2016).

3.6.4 Interpreting line wings

The various models used for HI stacked profiles are driven by different interpretations of the line wings. The two-Gaussian models used by Ianjamasimanana et al. (2012) argues that the line wings trace a turbulent WNM component, while Stilp et al. (2013a) assign the wings as the product of feedback. Our findings suggest that line widths result from a mixture of different physical processes, making it difficult to uniquely identify the source of the line wing excess without a detailed knowledge of the processes that can broaden the line profiles. The mixture model presented in §3.6.3 shows that the excess line wings naturally result from combining a set of Gaussian components. Since both turbulent and thermal motion produce Gaussian profiles, the effect of each component cannot be determined from the line wing fraction without additional information.

Both in individual spectra and stacked profiles (over the approaching or receding halves of the galaxy), we find that line wings are not symmetric and are consistently skewed towards the systemic velocity (Figure 3.8). Using the $v_{\rm peak}$ stacked profiles in the northern and southern halves, we find that about 1/3 of the line wing fraction results from this asymmetric component (Figure 3.14). Based on our comparison of off-rotation and extra-planar structure in §3.5, we identify this component as arising from the lagging rotational disk.

Figure 3.14 shows that the lagging rotational component has roughly an equal contribution to the line wing fraction throughout the disk.

The symmetric component of the line wing fraction, however, shows substantial variation with galactocentric radius, though there is no consistent trend. Instead, the variations appear to result from local processes within the disk. For example, there is an increase in the mid-disk ($R_{\rm gal}$ = 2–4 kpc) where the spiral arms dominate, possibly driven by enhanced recent star formation in the region and bulk non-circular motions driven by the arm mechanism.

The radial trend is inconsistent with the symmetric line wing component being driven solely by stellar feedback. Line wings resulting from stellar feedback should correlate well with tracers of the star formation rate, which drops exponentially with radius in M33 (Heyer et al., 2004; Boquien et al., 2015), as does the H_2 surface density (Druard et al., 2014). Stilp et al. (2013a) find a correlation between the SFR and line wing fraction for stacked profiles over the full galaxy of their sample of dwarf galaxies. However, the correlation is tenuous when the stacking is performed in radial bins (Stilp et al., 2013b). Many of the galaxies in their sample show relatively constant $f_{\rm wings}$, or otherwise peak far from the centre of the disk, as we find in M33. The lack of a radial trend may also be inconsistent with the turbulent WNM interpretation, since turbulent motion is expected to decrease with radius. We discuss this in context with the line width of the stacked profiles in §3.6.5.

The increase in $f_{\rm wings}$ beyond 6 kpc, driven entirely by the symmetric component, emphasizes that multiple processes must contribute to the line wings. The increase in this region is consistent with the behaviour in skewness and kurtosis analyses in §3.4.3, and we attribute the variations in §3.4.4.1 to M33's warped disk component. Corbelli et al. (2014) and Kam et al. (2017) both find evidence for the kinematic effects of the warp at similar radii.

3.6.5 A shallow radial gradient in the line width

As mentioned in §3.6.1, the central regions of many nearby galaxies have a steep radial gradient in the velocity dispersion, with notable exceptions in-

cluding the flocculent spirals NGC 2403 and NGC 2976. From Figure 3.16, we find that all of the line width measurements have a shallow radial trend. The line widths from the $v_{\rm peak}$ stacked profiles decrease by just ~ 1.5 km s^{-1} between a radius of 2 and 8 kpc. This variation is consistent with the radial decrease found in M33 by Druard et al. (2014) using 1 kpc bins.

The measured line widths from the $v_{\rm peak}$ stacked profiles are narrower than the expected thermal line width of ~ 8 km s^{-1} for the WNM (for a temperature of ~ 8500 K; Wolfire et al., 1995). And as found using the Gaussian Mixture model, the excess kurtosis universally found in stacked profiles requires a contribution of narrow Gaussian components with equivalent thermal widths of $T < 1000$ K. The line widths of the stacked profiles should then be some combination of varying thermal line widths, ranging from CNM to WNM temperatures, with turbulent broadening, which plausibly varies as well.

The line width traces the distribution of thermal and kinetic energy in the atomic medium. Cooling and dissipative mechanisms will remove this kinetic energy, so maintaining these line widths will require energy injection. Previous work has used the energy balance in the ISM of other galaxies to determine sources for the energy injection (e.g. Tamburro et al., 2009; Stilp et al., 2013b). All 11 galaxies examined in Tamburro et al. (2009) have a steep line width gradient, which tend to flatten towards the outer disk. They find that feedback via supernova is not sufficient to maintain the observed line widths in the outer regions, but magneto-rotation instabilities (MRI) plausibly could.

In particular, M33 has an increasing (not flat) rotation curve (Figure 3.5), which contrasts with models used in the previous work. Similarly the star formation rate declines exponentially with radius (Heyer et al., 2004; Corbelli et al., 2014), providing insight into the rate of energy injection by core collapse supernova. We estimate the volumetric energy dissipation rate by assuming motions are turbulent with an outer scale set by the disk scale height of the atomic ISM in the galaxy (H, taken to be 100 pc; Tamburro et al., 2009) and

dissipate on crossing times ($\tau_c = H\sigma^{-1}$) for the system:

$$\dot{u}_{\rm diss} = \frac{3}{2}\frac{\Sigma\sigma^2}{H\tau_c} = \frac{3}{2}\frac{\Sigma\sigma^3}{H^2}, \tag{3.15}$$

where Σ is the surface density of the atomic ISM. The rate of MRI energy injection has been estimated as

$$\dot{u}_{\rm MRI} = 0.6\frac{B^2}{8\pi}R\frac{d\Omega}{dR} \tag{3.16}$$

following Mac Low & Klessen (2004a). Here B is the characteristic magnetic field strength, taken to be a constant $B = 8$ μG with radius based on modeling of the non-thermal radio continuum by Tabatabaei et al. (2008). We use our rotation curve model to evaluate orbital frequency Ω at galactocentric radius R. We take the energy injection from core-collapse supernova from Mac Low & Klessen (2004a):

$$\dot{u}_{\rm SN} = \epsilon_{\rm SN}E_{\rm SN}f_{\rm SN}\frac{\dot{\Sigma}_\star}{\langle m\rangle H}. \tag{3.17}$$

Here $\epsilon_{\rm SN} \sim 0.1$ is the efficiency with which supernovae inject energy into the ISM, $E_{\rm SN} = 10^{51}$ erg is the energy of a supernova, $\dot{\Sigma}_\star$ is the star formation rate surface density, $\langle m\rangle = 10^2$ $M_\odot$ is the average mass of stars that forms per core collapse supernova (Tamburro et al., 2009) and H is the scale height of star formation, taken to be 100 pc again. We use the star formation rate maps from Boquien et al. (2015), specifically the combined maps from the Galex FUV and Spitzer/MIPS 24 μm data. We generate a radially averaged star formation rate profile as the median of data in annular bins with width of 50 pc.

In Figure 3.20, we plot the radial profiles of the turbulent energy dissipation adopting a range of line widths from 6 to 12 km s^{-1} and a constant surface density of 8 $M_\odot$ pc^{-2}, which describe our data well. Neither injection mechanism can explain the line widths in isolation. Because of the rising rotation curve and comparatively low shear, MRI driving rates are significantly below the dissipation rate estimates, but the MRI mechanism shows a flat profile consistent with the observed shallow trends in line width and surface density.

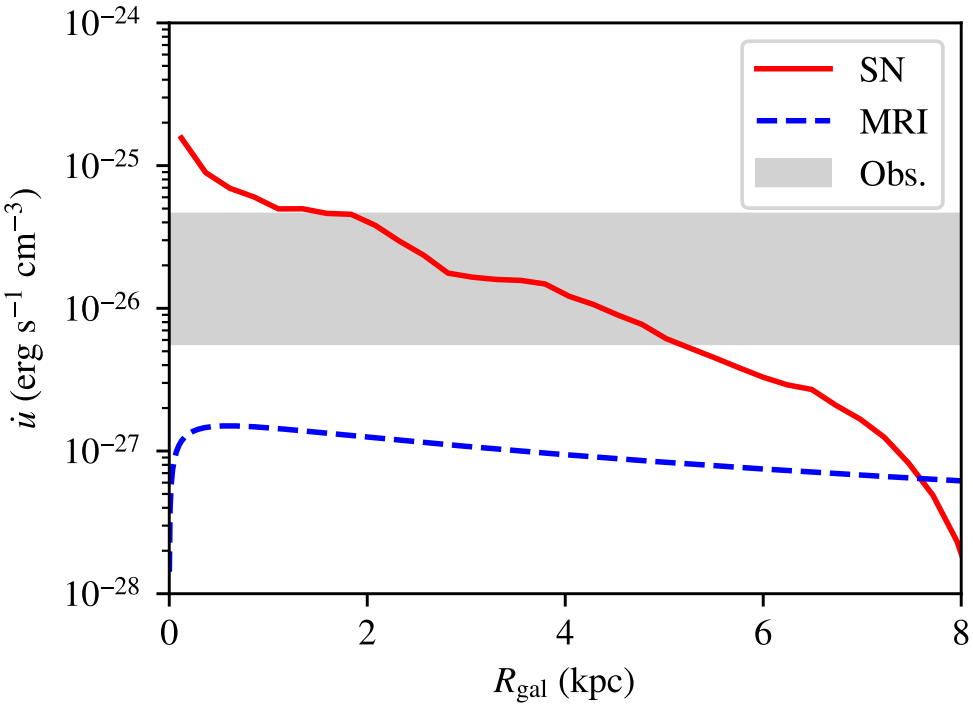

Figure 3.20: Radial profiles of energy injection for core-collapse supernova (SN; solid red line) and magnetorotational instability (MRI; dashed blue line) and estimates of the turbulent dissipation rates in the atomic ISM. Neither mechanism, by itself, matches the range and constant profile of the observed dissipation rates.

The star formation rates falls off exponentially and cannot explain the flat profile of the line widths and surface density but it does broadly agree with the range of required energy injection rates. Appealing to Type Ia supernovae as an additional source of turbulent motion will not change the profile substantially since the old stellar disk in M33 also declines exponentially with a scale length of 1.4 kpc (Regan & Vogel, 1994).

This model adopts a constant atomic gas scale height throughout the observed region, but the scale height likely increases with galactocentric radius. While the radial variation of the scale height of gas in M33 has not been measured, other disk galaxies show that the atomic gas scale height increases linearly by a factor of two over the optical radius of the galaxy, R_{25} (e.g., Yim et al., 2014). For M33, $R_{25} = 9$ kpc (de Vaucouleurs et al., 1991), implying that $\dot{u}_{\rm diss}$ would decrease by a factor of < 4 over the plotted region. The supernova driving would also show a steeper decline, dropping by an additional factor of two. However, the thickness of the star-forming disk is relatively constant (Yim et al., 2014) so the scale height used in $\dot{u}_{\rm SN}$ may remain constant with radius. Even with this assumption, a flaring HI disk does not clearly match the factor of 10^2 decline in supernova driving in this simple model. Using higher resolution HI data, it will become possible to map spatial variations in the scale height through changes in the turbulent properties of the gas (Padoan et al., 2001) and investigate these effects in more detail.

Without a more detailed study, we thus arrive at a similar conclusion as Stilp et al. (2013b), who examine a number of energy sources for driving turbulence in their sample of dwarf galaxies and find that a single source cannot maintain the observed line widths.

3.6.6 Interpretation of different stacking models

The properties of stacked profiles depend on how the spectra are aligned. The appropriate choice of which definition of the line centre to use — the rotation velocity ($v_{\rm rot}$), centroid velocity ($v_{\rm cent}$), or velocity at peak intensity ($v_{\rm peak}$) — depends on the goal of the analysis.

Stacking based on $v_{\rm rot}$ maximizes contributions from non-circular motions in the disk. In M33, the width of the velocity residuals is significant when compared to the typical WNM thermal width. The line width of this stacked profile will measures the average kinetic energy in the ISM in the galactic potential.

The properties of the atomic ISM components at the spatial resolution of the data are best recovered by stacking based on $v_{\rm peak}$ since this minimizes the variance in the line centre distribution for the bright components. Line wings are also the most significant in this method but trace distributions of temperature in the ISM as well as multiple velocity components along the line of sight (§3.6.3). Assuming the statistical uncertainty on $v_{\rm peak}$ is small, the resulting stacked profile will recover the average line shape at the resolution of the data. We explicitly test this in a forthcoming paper (Koch et al., in prep) and our preliminary results show that individual HI components have a median line width of ~ 6 km s^{-1} based on a Gaussian fit, which matches the 6–7 km s^{-1} found for the $v_{\rm peak}$ stacked profiles (Table 3.4). However, in our upcoming work, we find that individual components have widths ranging from 4–10 km s^{-1}. While the $v_{\rm peak}$ profiles appear to recover the typical spectrum's properties, it necessarily removes spatial variations in the widths.

The line width of the $v_{\rm cent}$ stacked profile can be interpreted as tracing the kinematics of the internal motions of the neutral medium, since, for an optically thin tracer, the centroid is the mass-weighted average velocity. The dispersion around this velocity represents the energy of the gas-to-gas relative motions, as opposed to the gas-to-galactic potential motions traced by the rotation-stacked profile. Because $v_{\rm cent}$ is biased by the line shape, however, caution must be used if the line shapes have a preferential skewness as this will artificially broaden the velocity distribution. Creating a map of the skewness (Figure 3.7) or checking the asymmetry of stacked spectra over limited regions of the galaxy (Figure 3.12) can be used to diagnose how substantial the bias will be.

Thus, the line width in each of the stacked profile types result in a dif-

ferent but physically meaningful measure of typical atomic ISM properties. Differences in the interpretations of these line widths present in the literature is partly attributable to real variations, but also arise from the method used for modeling. Most of the literature aims to characterize components of the ISM, however we caution that these interpretations are necessarily uncertain, even with good velocity and spatial resolution. We demonstrate this uncertainty in the source of stacked profile properties with the Gaussian Mixture Model in §3.6.3. Reasonable distributions of ISM properties — for example the temperature — can produce the qualitative properties of observed stacked profiles. However, a realistic mixture model must also include distributions for other physical processes, such as turbulent broadening, outflows from stellar feedback, bulk non-circular motion and lagging rotational motion, amongst others.

For the M33 data, the best recourse is to perform a multi-component Gaussian decomposition on a per line-of-sight basis. Such methods can realistically only work for good-quality observations of nearby galaxies. These detailed analyses are able to separate the effects of different processes by retaining spatial relationships. We recommend that studies seeking to separate physical processes either model individual spectra (e.g., Warren et al., 2012) or average spectra over the size scale of the targetted processes. For more distant targets where the scale of most processes will be unresolved, or for poor S/N data, stacking methods are still useful. Though the stacked profile will combine many physical processes together, it can be used as a powerful measure of similarity between different galaxies. For example, the consistent line widths found by Ianjamasimanana et al. (2012) between dwarf and spiral galaxies suggests that some atomic ISM properties are unchanged by galaxy type.

3.7 Summary

We present new L-band observations of M33 taken by the VLA in C-configuration. The new data set yields a spatial resolution of $\sim 18''$, tracing ~ 80 pc scales,

and significantly improves the sensitivity and spectral resolution compared to previous observations with an rms brightness temperature of 2.8 K in a 0.2 km s^{-1} channel.

1. The new HI VLA data recover 72% of the total HI emission in M33 compared to estimates from single-dish data within the VLA survey region. After combining the VLA data with GBT observations, the total emission matches the emission from the GBT data alone. This gives a total atomic mass of $1.3 \pm 0.2 \times 10^9$ $M_\odot$. The combined data cube is fully-sampled down to 80 pc scales.

2. We fit a circular rotation model to constrain the kinematic parameters on 100 pc scales. The rotation curve is well represented by a Brandt (1960) model with $v_{\rm max} = 110.0 \pm 1.5$ km s^{-1} at $r_{\rm max} = 12.0 \pm 1.3$ kpc. In general, we find good agreement with previously published rotation curves for the inner 8 kpc of the disk.

3. For galactocentric radii $R < 7$ kpc, the azimuthally averaged atomic gas profile has a nearly-constant average surface density of $\Sigma_{\rm HI} = 8\ M_\odot$ pc^{-2}, though the observations highlight local variations of 25% with radius. These variations are seen in both limbs (north and south) of the galaxy.

4. The HI line profiles are consistently non-Gaussian. We parameterize this by calculating higher-order moments — skewness and kurtosis — for the line profiles. We find that HI profiles are asymmetrically skewed towards the systemic velocity, resulting in the northern and southern halves having oppositely-signed skewness. By examining PV-slices, we find evidence for a lagging rotational component in M33, consistent with Kam et al. (2017). We find a typical excess in kurtosis of ~ 0.2, indicating that the typical spectrum has an excess in the line wings relative to a Gaussian.

5. We stack spectra over the entire galaxy, and in 100 pc radial bins. By modeling the profiles as Gaussian peaks with enhanced line wings, we

find that stacking based on the velocity at the peak intensity ($v_{\rm peak}$) of spectra gives the smallest widths of ~ 7 km s^{-1}. This is consistent with arising from a combination of cool and warm atomic gas with some turbulent broadening. Stacking based on the centroid ($v_{\rm cent}$) and rotation ($v_{\rm rot}$) velocities give large linewidths of ~ 10 km s^{-1}, which provides an estimate of internal motions in the atomic medium. However, the centroid is biased by the asymmetric line wings, which will tend to broaden the stacked profile. Line width estimates from the second moment are larger than all of the stacked profile widths and are sensitive to extended line wings.

6. All estimates of the line width show a shallow decrease with radius, dropping $\sim$ 2 km s^{-1} over the inner 8 kpc, though with significant (20%) fluctuations. This is atypical for most nearby galaxies, which have steeper slopes at small galactocentric radii. We find that simple estimates of the volumetric energy dissipation rates from core-collapse supernova and magneto-rotational instability cannot explain the radial trend in the line width. The rising rotation curve and low shear in M33 leads to MRI energy injection rates that are significantly lower than the estimated range from the observed line widths. More careful measurements of the scale height of atomic gas in M33 are needed to make these conclusions robust.

7. The fraction of excess emission in the stacked profile line widths ranges from 9% to 26% depending on the choice of line centre. There is no clear radial trend in the inner 6 kpc, but local variations of 50% are significant. We split the line wing fraction into asymmetric and symmetric parts based on the stacked profiles in the northern and southern halves. We find the asymmetric part, from the lagging disk, accounts for 1/3 of the line wing excess and is nearly constant with radius. The radial variations in the line wings are almost entirely driven by the symmetric part. Beyond 6 kpc, the symmetric part steadily increases, which we

attribute to the start of M33's warped disk.

8. We present a Gaussian Mixture Model to explain stacked profile shapes. Using only a set of Gaussian components with thermal line widths, we can qualitatively reproduce the stacked profile shape found in many previous studies. In particular, the highly peaked centre of the profile can only be reproduced by including line widths expected for the CNM. However, the model line widths are smaller than the observed range, suggesting that a combination of the CNM/intermediate/WNM fractions, turbulence, a lagging rotational disk, and multi-component spectra act to broaden the observed stacked profiles. This large possible combination of physical processes cannot be extracted from the stacked profile without additional information.

9. We identify discrete high-velocity gas structures on the blue- and redshifted sides of the disk. These structures have a total mass of $1.3 \pm 0.5 \times 10^7$ $M_\odot$ in this high-latitude component, about 1% of the total atomic mass in the system. Some of these structures were previously identified with single-dish studies. Here, we can resolve these structures and their extent across the main disk. All high-velocity structures appear as a clump surrounded by fainter filamentary structure. We find two features of particular interest: a long filament in the southern half of the disk with a project length of 8 kpc, and a cloud 600 pc in diameter overlapping in velocity with the main disk, indicating a possible interaction point. These structures highlight the complexities of M33's halo, and possibly a connection with the warped disk component.

10. We find no detections of RRL emission, based on stacking six RRL lines. We set a 3σ upper limit of 3.0 mJy in a $60''$ region towards NGC 604.

In conclusion, the high spectral resolution observations of M33 show the galaxy harbours a kinematically rich atomic medium. With complete spatial sampling down to 80 pc scales, our analysis shows that several common analysis

paths used in extragalactic observations have limitations in describing these spectra. We highlight that overcoming these limitations requires decomposing the spectra into their individual components to gain a full understanding of the atomic ISM kinematics. This provides new opportunities for exploring the data using tools from the Milky Way community.

Scripts to reproduce the reduction, imaging, and analysis are available at https://github.com/e-koch/VLA_Lband.

Acknowledgments

We are grateful for helpful discussions with Doug Johnstone in developing this paper, as well as Adam Ginsburg and Hauyu Baobab Liu for discussions on imaging. We also thank the anonymous referee for their feedback and comments that improved the paper.

EWK is supported by a Postgraduate Scholarship from the Natural Sciences and Engineering Research Council of Canada (NSERC). EWK and EWR are supported by a Discovery Grant from NSERC (RGPIN-2012-355247; RGPIN-2017-03987).

Code Bibliography: CASA (version 4.4 & 4.7; McMullin et al., 2007) — astropy (Astropy Collaboration et al., 2013) — radio-astro-tools (spectral-cube, radio-beam, uvcombine; radio-astro-tools.github.io) — matplotlib (Hunter, 2007) — seaborn (Waskom et al., 2017) — DISKFIT (Spekkens & Sellwood, 2007; Sellwood & Spekkens, 2015) — image_registration (http://image-registration.readthedocs.io)

3.8 Appendix

3.8.1 Imaging approach

At a spectral resolution of 0.2 km s^{-1} and a spatial grid size of 2560^2 pixels needed to cover the entire mosaic, imaging and deconvolution requires significant computational time and power. The size of the resulting cube is $1178 \times 2560 \times 2560$ pixels, giving a size of ~ 29 GB. Rather than imaging the

cube as a whole, we split the data into individual velocity channels, image and deconvolve each channel separately, then recombine the imaged channels into a final data cube. When working on a cluster, this approach allows channels to be simultaneously imaged, providing a significant speed-up in the time required to image the entire cube. Furthermore, since the size of the channel measurement set (MS) is much smaller than the original MS, slow I/O operations are relatively minimized. Using CASA 4.4, we found that imaging a single channel from the complete MS was $\sim 5\times$ slower than when imaging from a split channel MS.

This approach to imaging large data-cubes has a significant bottleneck during the splitting stage: the time required to split all 1178 channels from the complete MS was nearly two weeks on the Jasper cluster[8], which uses a lustre-based file system. In hindsight, we note that this is not an optimized choice of operation; a binary-split algorithm that progressively splits an MS would likely achieve a significant speed-up in the operation, particularly since this can be naturally parallelized. Furthermore, the storage requirements are more than double the size of the original MS. The new CASA task TCLEAN allows for the measurement set to be opened in a read-only mode, largely mitigating the need to split off individual channels. This approach will be used in the future.

3.8.1.1 Image combination

We explore the effects of combining the VLA data with Arecibo (Putman et al., 2009) and Green Bank Telescope (GBT; Lockman et al., 2012) observations to provide short-spacing information. Both are well suited to be combined with our VLA data as they have similar spectral resolution and at least a factor of two spatial overlap in the *uv*-plane. We note that older GBT data were combined with the archival VLA observations presented in Braun (2012), but these data have a spectral resolution of 1.42 km s^{-1} and are not well-suited for combination with the 0.2 km s^{-1} resolution VLA data presented here.

[8]https://www.westgrid.ca/resources_services

The Arecibo data from Putman et al. (2009) has a spatial resolution of $3.'4$ and a spectral resolution 0.4 km s^{-1}. This provides a significant spatial overlap with the VLA data, but also requires up-sampling in the spectral domain, thereby increasing the channel-to-channel correlation. We first spatially-register the data using the cross-correlation method in IMAGE_REGISTRATION[9] and find that no correction is needed. We use the FEATHER_SIMPLE[10] task from the UVCOMBINE package to combine the data. This method of combining the data assumes the single-dish beam is well-approximated by an isotropic Gaussian kernel. However, the Arecibo data contains significant side-lobe structure, which leads to enhanced negative bowling in the feathered image on scales of $\sim 1\ '$. Further, the power-spectrum of the Arecibo data, clipped to match the region covered by the VLA mosaic, shows a drop in power on the scale of the entire image, indicating a large-scale ripple. This ripple is far less significant in the larger region around M33 used by Putman et al. (2009). For these reasons, we did not pursue further image combination with the Arecibo data.

The GBT data from Lockman et al. (2012) has a nominal spatial resolution of $9.'1$ and a spectral resolution 0.16 km s^{-1}. Unlike the Arecibo data, no spectral up-sampling is required to match the VLA data. These data were taken in four $2^\circ \times 2^\circ$ regions centred on M33; aspects of the GBT gridding are presented in §3.2.2. The effective resolution of the data is $9.8'$ (see below), which gives sufficient *uv*-overlap for combining with the VLA data. The GBT beam has minimal side-lobe structure after calibration and the combined images with the VLA data does not show the enhanced bowling we encountered with the Arecibo data. Before feathering, we test if the data are spatially-registered and find that the GBT data require a $3''$ shift in declination. Visually, this shift appears to provide a better combination with the data compared to when no shift is applied. However, we note that $3''$ is within the pointing error for the GBT data.

[9]http://image-registration.readthedocs.io

[10]This matches CASA's feather implementation.

3.8.1.2 Combination tests

Following Chapter 3 of Stanimirovic (1999), we run two tests on the *uv*-amplitudes where the spatial coverage of the VLA and GBT data overlap to 1) ensure a bias is not added from using an incorrect single-dish beam model, and 2) derive a relative calibration factor to obtain a consistent flux-density scale. We define the *uv* overlap region as all points between the 9.′8 beam size of the GBT data and 16.5′– for the shortest baseline of 44 m in the VLA data. This gives ~ 200 overlap points per channel at the grid size of the VLA data ($3'' \times 3''$). To get the GBT amplitudes used below, we deconvolve its Fourier transform by dividing by the Fourier transform of the GBT beam. The GBT amplitudes in the overlap region are then multiplied by the ratio between the beam areas to account for the difference in resolution (Stanimirovic, 1999).

In §3.2.1 of Stanimirovic (1999), an approximate relation between the scale factor and the *uv*-distance k is given:

$$f_{\rm cal} = \left[1 + \frac{\Delta\theta(2\theta_0 + \Delta\theta)}{4{\rm ln}2} k^2\right], \tag{3.18}$$

where θ_0 is the true FWHM single-dish beam size and $\Delta\theta$ is the deviation from the true beam size and assumed to be small. This predicts a linear relation between $f_{\rm cal}$ and k^2 when an incorrect beam size is assumed. The correct beam size should have a slope of zero. We perform this fit for each velocity channel in the data using a robust Theil-Sen estimator to find the slope. A robust fitting method is required since the combined effects of the noise and different emission structure in the channels will naturally produce significant outliers. The Theil-Sen method computes the slope from the median of all slopes between each pair of data points, as is defined in the SCIPY implementation[11]. The slopes for each channel are shown in Figure 3.21 using a GBT beam size of 9.′8. This beam size maximizes the number of channels with slopes consistent with zero. The variations in the slope with different channels – notably the first 100 and last 200 channels – are driven by the

[11] https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.theilslopes.html#scipy.stats.theilslopes

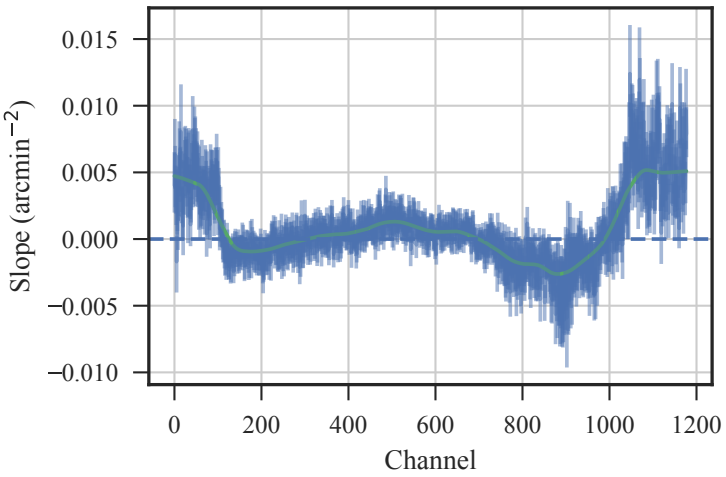

Figure 3.21: Derived slopes of the *uv*-distance versus the ratio between the VLA and GBT amplitudes in the overlap region. A slope consistent with zero suggests the beam model is correct; an incorrect beam model will incorrectly bias the amplitudes, giving a non-zero slope.

lack of emission structure and correspondingly have larger uncertainties. The channels dominated by signal in the *uv*-overlap range are from 200 to 800, and since this region has slopes largely consistent with zero, we adopt a GBT beam size of $9.'8$ for feathering.

With the correct GBT beam size established, we now address the scaling factor between the VLA and GBT amplitudes. We follow a similar procedure to the one described in §3.2 in Stanimirovic (1999). The emission structure in both the VLA and GBT smoothly tapers near the map edges and no additional edge tapering is required. We first fit a line between the GBT and VLA amplitudes, similar to the method used by IMMERGE in MIRIAD. The issue with the linear fitting approach is consistently dealing with outliers. We adopt the Theil-Sen method rather than the L1-minimization used in IMMERGE since the former tends to be insensitive to outliers in both the x and y directions, while the latter is only robust against outliers in the y direction (Wilcox, 2010). We fit a relation to the amplitudes in groups of five velocity channels to increase the number of points. Despite the robust nature of the Theil-Sen

fit, we did not find consistent results across the channel groups. The issue is the severity of the outliers and their fraction relative to the entire data set: if there are too many extreme outliers, the Theil-Sen fit still has some sensitivity to the outliers. Increasing the number of channels simultaneously fit did not show improvement, and fitting the entire set of amplitudes is prohibitive for the Theil-Sen method since it computes the slope for every pair of data points.

Using a distribution-based method, we are able to provide more reliable constraints on the scaling factor. Motivated by §3.2 in Stanimirovic (1999), we first examine the distributions of the GBT amplitudes, the VLA amplitudes, and their ratio across all channels. The GBT and VLA amplitudes reasonably follow a log-normal distribution[12]. Since the ratio of two normal random variables follows a Cauchy (or Lorentzian) distribution, the log of the amplitude ratios can be fit to this form. The scaling factor is then the location of the distribution's peak. Figure 3.22 shows the distribution of the amplitudes across all channels (blue) and the best-fit Cauchy distribution (green). This approach has the significant advantage that outliers are included in the expected model (the distribution tails) and do not require special treatment. The Cauchy distribution is also preferable to the Rayleigh distribution used in Stanimirovic (1999), which gives more weight to the outliers in the right-tail. Adopting a maximum-likelihood approach, we find a scaling factor of 1.02 ± 0.06 using the ratios from all channels. As this is consistent with 1, *we do not apply a scaling factor to the GBT data before feathering.* We also perform the fitting on groups of five channels and find scaling factors consistent with one between channels 200 to 800, where most of the emission is contained.

3.8.2 Choice of velocity surface for the rotational model

Galactic rotation curves, with a few exceptions, are typically derived from velocity surfaces. A key question explored in several publications — notably de Blok et al. (2008) — is: which velocity surface optimally traces galactic

[12]Note that this does not imply a log-normal distribution for the true signal. The log-normal shape results from the mixing between signal, noise, and other telescope effects.

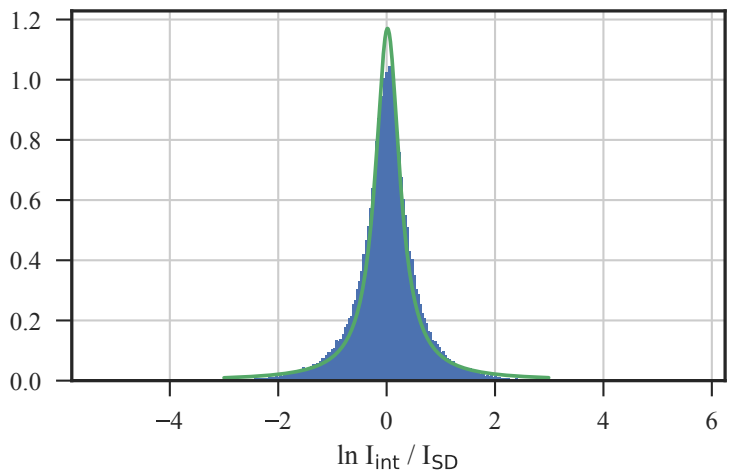

Figure 3.22: The log of the VLA-to-GBT amplitude ratios in the *uv*-overlap region across all channels. The histogram of the ratios is normalized to one and the green line is the best-fit Cauchy distribution. The peak of the distribution is at 1.02 ± 0.06, consistent with a scale factor of one.

rotation? Commonly used methods include the centroid ($v_{\rm cent}$), the velocity at the peak intensity ($v_{\rm peak}$), and Gauss-Hermite fitting. We create velocity surfaces of the former two methods for the VLA-only and combined HI cubes and fit a rotation curve using DISKFIT with the same parameter settings. We did not fit a rotation model for the Gauss-Hermite velocity surface for two reasons. First, using the VLA-only cube, the difference between the Gauss-Hermite centre and $v_{\rm peak}$ was much smaller than the difference to the $v_{\rm cent}$ surface and will yield similar results. The second reason is that the Gauss-Hermite fitting did not perform well with the enhanced line wings for the combined data.

We find that the rotation models using the VLA-only $v_{\rm peak}$ and $v_{\rm cent}$ falls within the uncertainty of the combined VLA & GBT $v_{\rm cent}$ surface, whose rotation curve is presented in §3.4.1. However, the $v_{\rm cent}$ surface from the combined data has a shallower rotation curve. The reason for this discrepancy, and the agreement of the other rotation models, is due to the bias from the line wings. Without the component from the GBT, the VLA-only $v_{\rm cent}$ is not

as biased by line wings and largely matches the $v_{\rm peak}$.

3.8.3 Rotation velocities

We provide a table of the calculated rotational velocities from DISKFIT presented in §3.4.1.

Table 3.6: Circular rotation velocities derived by DISKFIT (§3.4.1). The rotation model is fit to the peak velocity ($v_{\rm peak}$) surface of the VLA+GBT data.

Radius ($''$)	Circ. Velocity (km s^{-1})
3	1.58 ± 6.24
9	11.88 ± 5.26
15	22.71 ± 5.25
21	30.53 ± 5.12
27	29.95 ± 4.47
33	34.20 ± 4.03
39	36.76 ± 3.99
45	44.61 ± 3.79
51	47.80 ± 3.49
57	51.67 ± 3.08
63	52.73 ± 2.96
69	53.24 ± 3.06
75	55.30 ± 3.22
81	57.58 ± 3.30
87	58.70 ± 3.42
93	59.47 ± 3.50
99	60.30 ± 3.59
105	64.32 ± 3.78
111	67.66 ± 3.75
117	69.07 ± 4.03

123	72.05 ± 4.49
129	75.89 ± 4.82
135	74.09 ± 4.91
141	73.44 ± 3.29
147	77.06 ± 2.76
153	77.06 ± 2.80
159	77.82 ± 2.60
165	79.01 ± 2.48
171	80.96 ± 2.59
177	80.34 ± 2.50
183	80.35 ± 2.65
189	82.90 ± 2.66
195	86.23 ± 2.66
201	86.92 ± 2.59
207	86.13 ± 2.63
213	87.43 ± 2.57
219	87.89 ± 2.63
225	87.82 ± 2.35
231	91.08 ± 2.44
237	90.37 ± 2.25
243	88.22 ± 2.09
249	90.67 ± 2.36
255	92.32 ± 2.48
261	93.31 ± 2.26
267	94.15 ± 2.15
273	93.40 ± 2.37
279	94.02 ± 2.44
285	94.91 ± 2.50

291	95.57 ± 2.48
297	95.15 ± 2.22
303	93.19 ± 2.46
309	94.52 ± 2.63
315	95.00 ± 2.77
321	95.72 ± 2.90
327	96.89 ± 2.81
333	98.58 ± 2.73
339	98.09 ± 2.61
345	99.87 ± 2.53
351	99.10 ± 2.41
357	99.01 ± 2.40
363	97.58 ± 2.21
369	98.60 ± 2.26
375	99.61 ± 2.22
381	99.61 ± 2.49
387	102.02 ± 2.40
393	101.84 ± 2.52
399	102.76 ± 2.38
405	102.66 ± 2.62
411	102.57 ± 2.68
417	104.08 ± 2.58
423	103.24 ± 2.39
429	103.17 ± 2.42
435	103.33 ± 2.50
441	103.41 ± 2.64
447	103.91 ± 2.56
453	102.91 ± 2.63

459	105.50 ± 2.47
465	104.05 ± 2.47
471	106.03 ± 2.47
477	104.89 ± 2.57
483	106.68 ± 2.48
489	105.53 ± 2.36
495	105.80 ± 2.33
501	104.03 ± 2.23
507	105.58 ± 2.23
513	105.69 ± 2.26
519	106.32 ± 2.17
525	105.97 ± 2.32
531	105.58 ± 2.20
537	106.18 ± 2.26
543	106.08 ± 2.28
549	106.24 ± 2.36
555	106.39 ± 2.37
561	106.47 ± 2.43
567	107.19 ± 2.44
573	105.88 ± 2.58
579	106.58 ± 2.52
585	106.09 ± 2.56
591	106.30 ± 2.35
597	107.23 ± 2.61
603	104.01 ± 2.82
609	104.50 ± 2.59

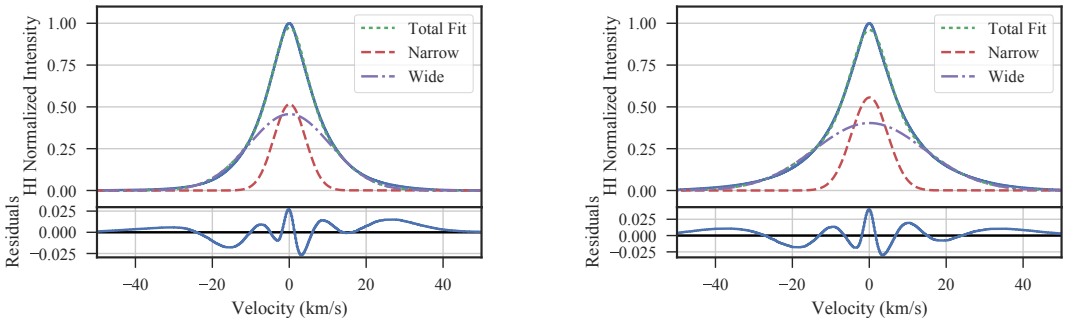

Figure 3.23: Two Gaussian fits to the $v_{\rm peak}$ stacked profiles shown in Figure 3.11. The left panel is a stack of the VLA-only data and the right uses the combined data. The model residuals are shown in the bottom panels.

3.8.4 Modeling super-profiles

We demonstrate an alternative modeling approach for the super-profiles: fitting two stacked Gaussian profiles. Figure 3.23 shows the $v_{\rm peak}$ stacked profiles presented in §3.4.4 with a fitted two-Gaussian model and the model residuals. The model parameters are given below. Note that we do not list uncertainties due to the large covariance between parameters, which we discuss below.

These two stacked spectra drove our decision to use the HWHM scaling from Stilp et al. (2013a), since the central peak is non-Gaussian and the resulting two-Gaussian fit does not provide a better description for the data. Of particular concern is that the model components cannot account for the central peak. The sum of the amplitudes in both fits is ~ 0.95. The narrow and wide amplitudes are 0.51 and 0.55 for the narrow component, and 0.45 and 0.4 for the wide component, for the VLA and combined profiles, respectively. The difference in the components is about the same as the missing peak intensity, making comparisons of their ratios uncertain.

We find widths of 4.1 and 10.5 km s^{-1} for the VLA-only profile, and widths of 4.7 and 14.4 km s^{-1} for the combined profile. The HWHM widths of 5.9 and 6.6 km s^{-1} for these two profiles are $\sim 50\%$ larger than the narrow component. This is consistent with comparing the common galaxies examined with these two methods in Stilp et al. (2013a) and Ianjamasimanana et al. (2012).

Since Gaussian components do not form an orthogonal set, there are large covariances between the parameters. Stilp et al. (2013a) explore this issue in depth and demonstrate that it is this covariance that leads to the overall good fits presented in Ianjamasimanana et al. (2012). Stilp et al. (2013a) show that the individual components from the two-Gaussian fit do not account for the overall profile shape. They demonstrate this by scaling all profiles to a common width, based on the different models, and find that the shapes from the broad and narrow components have significantly more scatter compared to the HWHM estimate.

While the two-Gaussian fits provide an adequate representation of the profiles, there are model-dependencies that impact the connection to the underlying physical parameters, to the extent that the super-profiles can give. Thus we use the HWHM modeling throughout this work.

3.8.5 Issues with the second moment for estimating the line width in Local Group galaxies

In Figure 3.24, we show that there is a stark difference between the averaged second moment line widths with and without the total power component added – the values increase by $\sim 30\%$. This highlights the impact of extended line wings when using the second moment to estimate the line width, making it difficult to directly connect with the underlying physical conditions. The large discrepancy seen here is a product of two factors. The first is the limit of small angular scales in the VLA data for a relatively nearby system. Figure 3.1 shows that the combined VLA and GBT data recover significantly more of the large-scale disk structure in the southern half than the VLA data only. This issue is less important for more distant systems (e.g., Walter et al., 2008). The second factor is the extensive extra-planar component in the warped disk of M33. As we suggest from the skewness maps (§3.4.3), the warped disk component influences the line shape near the edges of our map, increasing the line widths estimated from the second moment. The extent of this factor depends on the galaxy's environment. However, these average line widths

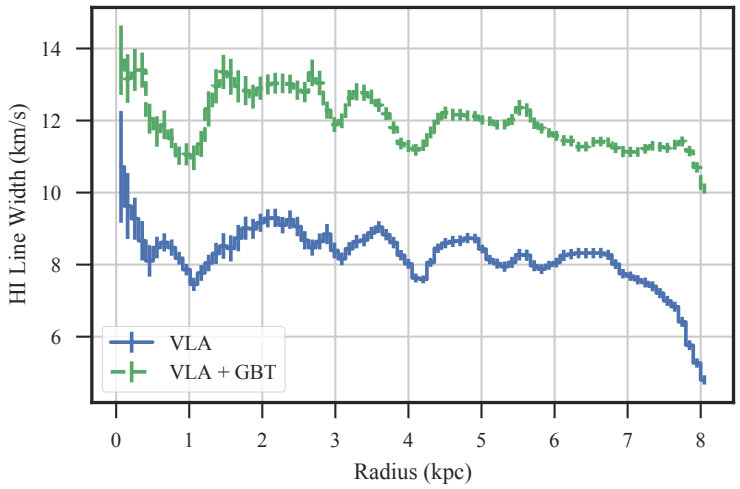

Figure 3.24: The azimuthally-averaged line width estimated from the second moment in 100 pc bins. The combined VLA+GBT data (green dot dashed) are 30% larger than the VLA-only values (blue solid). From the stacking analysis (Figure 3.11, however, we find that adding the GBT component tends to only affect the shape of the line wings. The large increase from the inclusion of the GBT data highlights that estimates from the second moment are very sensitive to the line wing structure.

are consistent with the range found in other HI studies of nearby galaxies (Tamburro et al., 2009; Mogotsi et al., 2016), showing that this effect is not strong enough to make M33 an outlier, or that line widths from the second moment in other galaxies are similarly affected.

Chapter 4

A Dearth of Tophats in M31 & M33: H I Spectra Strongly Prefer Multigaussian over Opaque Model Fits

To be submitted to the *Monthly Notices of the Royal Astronomical Society* as Eric W. Koch et al. 2020, "A dearth of tophats in M31 & M33: HI spectra strongly prefer multigaussian over opaque model fits." This work is in collaboration with J. Chastenet, I. Chiang, A.K. Leroy, E.W. Rosolowsky, K.M. Sandstrom, and D. Utomo.

Abstract

We model 21-cm HI emission spectra from recent VLA observations of M31 and M33 on 100 pc scales and find that the spectra strongly prefer a multi-Gaussian fit over a single opaque model fit (i.e., a top hat shape). This strong preference means that the observations do not support the existence of $\sim$ 100 pc cold HI clouds, which has been suggested in previous work at locations of bright HI in nearby galaxies. Our results are partially driven by the improved spectral resolution (0.42 km s^{-1}) over archival observations ($>$ 1.2 km s^{-1}). Many HI spectra have multiple distinct peaks in our observations, which may be blended at coarser resolution. To verify our model

selection test, we apply the same fitting to a sample of synthetic spectra drawn from the opaque model with observational noise added and confirm that the opaque model is correctly preferred on average. Previous studies compute a mass correction factor of $\sim 30\%$ from inferring the dark HI mass from the opaque model compared to the HI mass from the optically-thin assumption ($\propto$ HI integrated intensity). We show that the mass correction factor is strongly dependent on the definition of goodness-of-fit. We compute mass correction factors from 1%, using the 1% of spectra where the opaque model is favoured over the multi-Gaussian model, to 118% using the criterion from previous studies. Because of this large uncertainty, we suggest that the opaque HI mass is best constrained by HI absorption studies.

4.1 Introduction

Due to the ubiquity of the atomic interstellar medium (ISM) throughout the Milky Way and nearby galaxies, the 21-cm is an excellent tracer of many physical processes, in addition to probing the thermal and kinematic structure of the atomic ISM. These different processes produce confounding and degenerate effects in spectral line observations, and it can be difficult to disentangle the line profiles and recover the physical properties of the atomic ISM. This makes the modeling and interpretation of 21-cm HI spectral lines challenging.

The different states of the atomic ISM can be a particular challenge to account for in observed spectra. The classical interpretation of the atomic ISM is that it is found in two stable states in pressure equilibrium (Field et al., 1969; Wolfire et al., 2003): the warm neutral medium (WNM) and the cold neutral medium (CNM), whose temperature and density differ by factors of ~ 100. These thermodynamic differences result in distinct differences in their observed properties, namely that the WNM at low densities is optically-thin while the CNM at higher densities can become optically thick. In observations, the optically thick CNM is preferentially detected in HI absorption, while HI emission profiles are some combination of CNM and WNM. Additionally, ob-

servations find the presence of atomic gas in an intermediate unstable regime (unstable neutral medium; UNM) pushed out of the stable regimes by a perturbing source, such as turbulence (Heiles & Troland, 2003; Murray et al., 2018). These external sources broaden the observed spectral profiles, yielding line widths in excess of what is expected from thermal motion alone.

One key property measured in most extragalactic HI studies is the atomic ISM mass. Most studies measure the mass by assuming optically-thin HI emission, which allows the HI integrated intensity to be converted to column density. The optically-thin assumption is generally assumed to be valid because the HI in the warm neutral medium (WNM) has a larger volume filling factor relative to the cold neutral medium (CNM; Ferrière, 2001). Since most extragalactic HI studies resolve > 100 pc, the assumption follows from most of the emission coming from the volume-filling component. However, it is difficult to assess the validity of the optically-thin assumption from HI emission alone without constraints on the fraction of CNM/UNM, and thereby the opaque HI mass. This means that optically-thin HI mass measurements are likely underestimated in nearby galaxies.

Milky Way observations provide the strongest constraints on opaque HI from extensive HI absorption studies and provides a rough guide to the phase structure in nearby galaxies (though phase structure is expected to vary with galactic environment Wolfire et al., 1995; Bialy & Sternberg, 2019). Our location in the Galactic disc makes it favourable to measure opaque HI in absorption towards (i) background continuum sources (Heiles & Troland, 2003; Murray et al., 2018; Wang et al., 2020a) and (ii) lines-of-sight where cold gas (CNM/WNM) self-absorbs against a bright WNM background (Gibson et al., 2000; Gibson et al., 2005; Kavars et al., 2005; Soler et al., 2019; Wang et al., 2020b). The former strongly constrains the opacity and spin temperature, while the latter maps the spatial extent of nearby cold atomic gas (Gibson et al., 2005) or its association with molecular clouds (Li & Goldsmith, 2003; Goldsmith & Li, 2005; Krčo & Goldsmith, 2010). In locations without a bright background to self-absorb, cold HI is observed as narrow emission spectra (e.g.,

Haud, 2000) but becomes confused near the Galactic plane where the line-of-sight depth is large. Relative to the optically-thin HI mass, studies comparing HI emission and absorption find that opaque HI requires a mass correction from $\sim 10\%$ (Lee et al., 2015) to 140% (Bihr et al., 2015), depending strongly on the line-of-sight.

Direct constraints on the opaque HI mass in nearby galaxies is more difficult since their much small angular sizes intersect fewer bright background sources to measure HI in absorption against, and the top-down view means there there are fewer lines-of-sight with WNM emission located behind cooler HI to produce strong self-absorption. Because of their proximity, the Local Group galaxies intersect the largest number of background sources and thus have the strongest observational constraints on opaque HI from absorption studies By comparing HI absorption spectra to nearby emission spectra (Dickey et al., 2003), the fraction of atomic gas in the CNM is inferred to range from $\sim 15\%$ (M33; Dickey & Brinks, 1993) to $\sim 40\%$ (M31 & LMC; Braun & Walterbos, 1992; Dickey et al., 1994, 2000). The CNM fraction in the SMC is $\sim 20\%$, near the lower end of this range (Dickey et al., 2000; Jameson et al., 2019). However, the issue in extending these estimates remains the sparse spatial sampling of bright background sources. This makes extrapolating to a total opaque HI mass highly uncertain.

On the other hand, HI emission studies are extremely valuable since they provide the highest linear resolution continuous maps of HI in extragalactic systems (Stanimirovic et al., 1999b; Kim et al., 2003; Braun et al., 2009; Koch et al., 2018c). Because of this high linear resolution, HI emission maps of the Local Group galaxies show significant spectral complexity on < 100 pc scales. In this way, they more closely resemble the complex HI emission spectra in the Milky Way more than the limited spectral shapes that can be recovered from other nearby galaxies (e.g., Walter et al., 2008).

Noting the complexity in Local Group HI spectra, Braun et al. (2009) propose a solution for distinguishing between cold and warm atomic gas from HI emission alone They note that some HI spectra in M31 have a distinct

flattened-top, which can be reproduced by a simple model where a single gas component, broadened by non-thermal (e.g., turbulent) motion, is optically-thick above some column density (see also Rohlfs et al., 1972). If this is the case, it provides a powerful method for measuring the HI opacity since it relies only on HI emission, not absorption. Braun (2012) extend this identification to the M33 and LMC.

The key result from Braun et al. (2009) and Braun (2012) is that the optically-thin approximation underestimates the atomic gas mass by $\sim 35\%$ in M31, M33, and the LMC. This has vast implications for the atomic ISM mass measured using the optically-thin assumption, particularly for nearby galaxies beyond the Local Group where the linear resolution is 100s pc and the HI spectral line shape is poorly constrained. However, comparisons with other tracers suggest that the HI mass correction factors from Braun et al. (2009) are overestimated. Sandstrom et al. (2013) find that this level of opaque HI requires a large change in the dust-to-gas ratio (DGR) between atomic- and molecular-dominated regions, or otherwise require the CO-to-H_2 factor ($\alpha_{\rm CO}$) to increase. While the latter is expected to vary (Bolatto et al., 2013a), a strong effect due from opaque HI suggests its location should correlate with molecular clouds, as the source of bright CO emission. However, many of the opaque HI regions from Braun et al. (2009) and Braun (2012) are not correlated with CO emission. In contrast, the largest HISA structures in the Milky Way are strongly correlated with CO emission and molecular gas (Gibson et al., 2005).

Few other HI studies find flat-top spectra consistent with the opaque HI model, particularly in the Milky Way. Those that do note the difficulty in distinguishing a flat-top from two blended Gaussians (Rohlfs et al., 1972), where the latter may be favoured over a region where some spectra do show two clear peaks while others are too blended to distinguish between (Peek et al., 2011). Further, HI studies toward nearby molecular clouds with $\sim$ pc scale resolution do not find flat-top spectra distributed on molecular cloud scales (Lee et al., 2015). This is an important difference since the Braun et al.

(2009) model requires a high filling factor of cold atomic gas on $\sim$ 100 pc scales (their resolution) to for the model to be valid. Similarly, flat-top HI emission spectra are not prominent on $\sim$ 3 pc scales in the Small Magellanic Cloud (SMC) near bright continuum source where HI absorption is detected (Jameson et al., 2019).

One source for these discrepancies in the HI line shape may be the limited spectral resolution of the HI observations used in Braun et al. (2009) and Braun (2012), which range from 1.4–2.3 km s^{-1}. Assuming a resolved Gaussian component requires sampling with channels $\sim 2\times$ the line width (Koch et al., 2018b), the minimum thermal temperatures for a resolved component with these channel widths is 1300–3600 K, far larger than the the CNM temperatures and requiring a large non-thermal component to broaden the line width. Thus an HI spectrum with a flat-top may result from two or more overlapping Gaussian components where separate peaks become less distinct due to the coarse spectral resolution.

To investigate this issue, we use new HI observations of M31 and M33 with 0.42 km s^{-1} spectral resolution taken with the Karl G. Jansky Very Large Array (VLA). These observations in this work resolve linear resolution scales of 60–300 pc (accounting for galaxy inclination). Specifically, we compare the opaque "flat-top" model introduced by Braun et al. (2009) and Braun (2012) to a multi-Gaussian model. Our initial work with these high-sensitivity observations demonstrate complex HI spectra in M33 (Koch et al., 2018c), suggesting that additional features are observed with 0.42 km s^{-1} resolution that may not be evident in previous observations. With this spectral resolution, these observations are capable of resolving line widths to a lower limit of 120 K cold gas.

We introduce the observations in §4.2 and the model fitting and comparison definition in §4.3. In §4.4, we use the model comparison and present our main result: most HI spectra in M31 and M33 are best-modelled with a multi-Gaussian model. From this result, we discuss in §4.5 the lack of a large-scale opaque HI from emission spectra alone, and how this compares with

previous work in the Milky Way and nearby galaxies. Finally, we summarize our conclusions in §4.6.

Throughout this work, we adopt distances of 744 kpc for M31 (Vilardell et al., 2010) and 840 kpc for M33 (Freedman et al., 2001).

4.2 Observations

The observations in this work are described in Chapter 2. Specifically, we use the 19″ M33 map (VLA C-configuration; also see Koch et al., 2018c) and the new 18″ M31 map (VLA B and C-configuration with a uv-taper), which covers most of the northern half. Both data sets include short-spacing from the Effelsberg-Bonn HI Survey (EBHIS, M31; Winkel et al., 2016) and the GBT (M33; Lockman et al., 2012). Figure 4.1 shows the integrated intensity maps of both galaxies at the resolution used here. The remainder of this sections describes additional data handling required for this work. Table 4.1 summarizes the observational properties and provides a comparison to previous high-resolution 21-cm HI data (see also Thilker et al., 2002; Gratier et al., 2010).

For the M31 map, the bright background radio source B0044+419 falls within the field-of-view and strong HI absorption features from M31 are measured (Braun & Walterbos, 1992). We mask the region around B0044+419 by defining a region 1.5× the beam area to avoid fitting strong absorption features, for which the methods in §4.3 are not suited to model.

To match the M31 spectral resolution, we spectrally downsample the 19″ M33 HI cube from the original 0.21 km s^{-1} from Koch et al. (2018c).

The per-channel rms noise is 2.8 and 2.0 K in the M31 and M33 data, respectively. Since strong emission continues to the edges of the observed area, particularly for M31 (Figure 4.1), the noise varies with the effective primary beam for the mosaic. We weight the noise levels used for the spectral fitting in §4.3 by the primary beam, which increases the noise levels by $\sim 5\times$ in the most extreme cases affecting only a small number of spectra.

Finally, we note that we do not match the angular resolutions of the data set. This is primarily due to M31's large inclination (78° Corbelli et al., 2010), and so the linear resolution is $\sim$ 60–300 pc along the major and minor axes, respectively (similarly noted by Caldú-Primo & Schruba, 2016b). This difference in the linear resolution is less extreme in M33, which has a smaller inclination angle of 55° (Koch et al., 2018c). The linear resolution for M33 is then 80–130 pc. Accounting for variations in the linear resolution, these ranges overlap between the two galaxies, though their angular resolutions differ by $\sim 1''$.

To test this assumption, we use the fitting methods from §4.3 on small regions of M31 and M33 at matched angular resolution. The fitted properties are consistent between the matched and original resolutions.

	Target	Resolution	Linear Resolution	Channel	Noise	$5\sigma_{\rm rms}$ Optically-thin	Observatory
			Major × Minor axis	Width		column density	
		($''$)	(pc)	(km s^{-1})	($\sigma_{\rm rms}$; K)	(cm^{-2})	
Braun et al. (2009)	M31	30	100×500	2.3	1.0	2.0×10^{19}	WSRT
This Work	M31	18	60×300	0.4	2.8	9.8×10^{18}	VLA
Braun (2012)	M33	30	130×205	1.4	2.1	2.6×10^{19}	VLA
This Work	M33	19	80×130	0.4	2.0	7.0×10^{18}	VLA

Table 4.1: Summary of 21-cm HI observations from our work and those from Braun et al. (2009) and Braun (2012, see their Table 1). The M33 VLA observations used in Braun (2012) are presented in Gratier et al. (2010) (see also Thilker et al., 2002). The linear resolution range accounts for the galaxy inclination.

4.3 Modelling HI spectra

We test two models to describe the HI line shape: an optically-thick HI model for one broad component, and a multi-Gaussian model. In this section, we introduce these models, the fitting procedures, and the model selection used to compare between them.

4.3.1 Opaque HI Model

In the absence of a background continuum source, the general line profile for isothermal HI broadened by turbulence is

$$T_{\rm b}(v) = T_{\rm s}\left[1 - \exp\left[-\tau(v)\right]\right], \tag{4.1}$$

where v is the velocity[1], $T_{\rm b}$ is the observed brightness temperature, $T_{\rm s}$ is the HI spin temperature, and $\tau(v)$ is the optical depth profile. Following Braun et al. (2009), the optical depth profile can be expressed in terms of the HI column density ($N_{\rm HI}$; cm^{-2}), $T_{\rm s}$ (K), and line width (σ; km s^{-1}):

$$\tau(v) = \frac{5.49 \times 10^{-19} N_{\rm HI}}{\sqrt{2\pi} T_{\rm s} \sigma} \exp\left[-\frac{(v - v_{\rm p})^2}{2\sigma^2}\right], \tag{4.2}$$

where $v_{\rm p}$ is the velocity at the peak in optical depth. The line width can be expressed in terms of a thermal and non-thermal component, $\sigma = \sqrt{\sigma_{\rm T}^2 + \sigma_{\rm NT}^2}$, where the non-thermal component is presumed to be dominated by turbulence. Similar to Braun et al. (2009), we assume the kinetic and spin temperatures are equal ($T_{\rm k}$ =$T_{\rm s}$) in the absence of additional constraints. In the limit of high $T_{\rm s}$, high σ, or low $N_{\rm HI}$, we recover the optically-thin limit of a Gaussian line shape since $\exp\left[-\tau\right] \approx 1 - \tau$ for small τ.

Equation 4.1 includes an implicit assumption for a high filling factor of cold gas within the telescope beam. If the filling fraction is lower, we expect an observed spectrum will converge to the optically-thin case where the WNM emission dominated the spectrum shape, as is often assumed for nearby galaxies with coarse spatial resolution. If an observed 21-cm HI spectrum is

[1]Or equivalent spectral quantity.

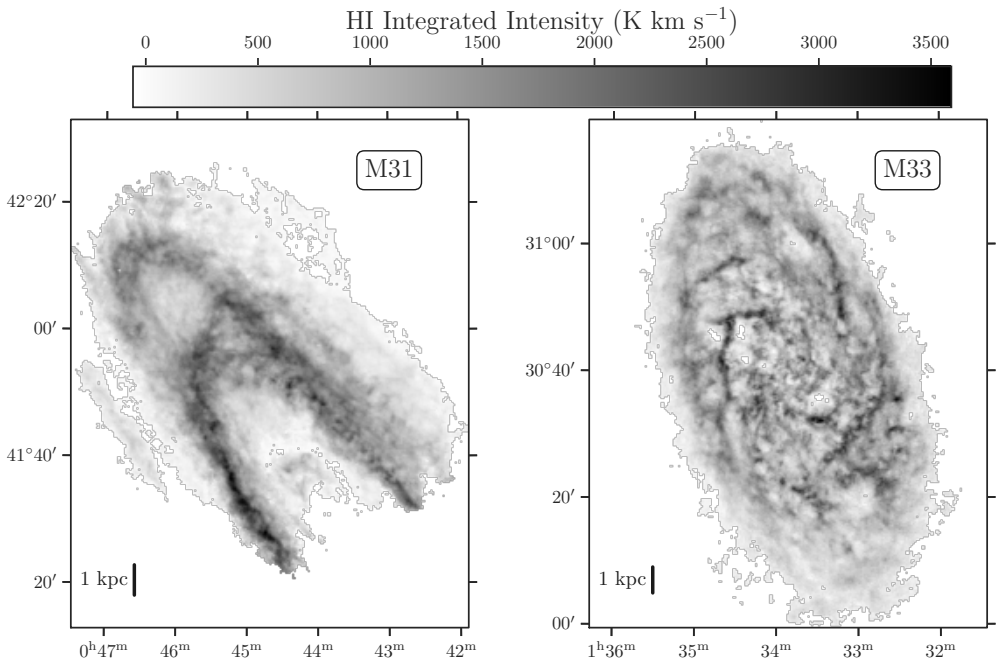

Figure 4.1: HI integrated intensity maps of M31 and M33. Our new VLA HI map of M31 covers most of the Northern half of M31. The M33 VLA HI map is presented and described in detail in Koch et al. (2018c).

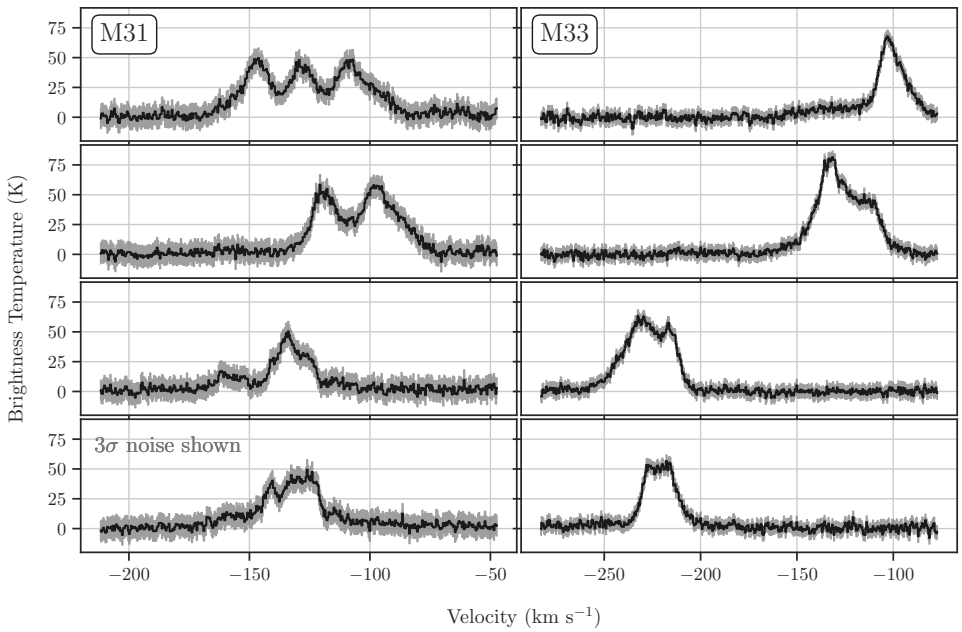

Figure 4.2: Example HI spectra in M31 (left) and M33 (right). The spectral channel width in all spectra is 0.42 km s^{-1}, and the 3σ rms noise levels are shown in gray to emphasize that the spectral complexity is real. In both galaxies, we find distinct spectral features that consistently deviate from a single Gaussian profile. Most of these example spectra have multiple peaks, though notably, the spectra in the bottom row show the general "flat-top" that may result from opqaue HI. By consistently modeling the HI spectra with the multi-Gaussian and opaque models, we quantitatively distinguish between the appropriate model for the array of spectral shapes.

significantly flattened, as Equation 4.1 implies, the opaque HI region must be similar in scale to the resolution. In our observations and those in Braun (2012), this would imply cold atomic "clouds" on ~ 100 pc scales.

Braun et al. (2009) and Braun (2012) fit with τ defined in Eq. 4.2 to HI spectra in M31, M33, and the LMC. An isothermal approximation is assumed to describe the bright HI along a line-of-sight, which follows from the low filling fraction of cold HI clouds from Braun (1997). Braun et al. (2009) use a grid search approach to find the best-fit model, chosen based on minimizing the reduced χ^2. To limit the range of possible models in the grid, the centre of the line was fixed to within ± 5 km s^{-1} of the velocity centroid. With their observations, this allows the centroid velocity in the fit to vary by ~ 2 channels.

To fit the HI spectra here, we make three changes to the approach from Braun et al. (2009) to measure and reduce parameter covariance from Equations 4.1 & 4.2.

1. We fit the model using using a Levenberg-Marquardt algorithm, as implemented in the `scipy` python package (Virtanen et al., 2020), instead of the grid search method from Braun et al. (2009). This approach enables us to estimate the parameter uncertainty and covariance in the fits. Parameter covariances become important as the model approaches the optically-thin limit, where σ, $N_{\rm HI}$, and $T_{\rm s}$ will be strongly covariate due to the lack of constraints on $T_{\rm s}$.

2. To minimize the parameter covariance in the model, we do not split the line width into a thermal and non-thermal component. This change is important when the profile approaches the optically-thin limit and $T_{\rm s}$ is unconstrained, which leads to the thermal contribution only being limited by the upper limit of $T_{\rm s} = 8000$ K for HI. When $T_{\rm s}$ is constrained, the non-thermal line width can be recovered by subtracting the thermal component in quadrature.

3. We further decrease the parameter covariance in the model by simplifying

the optical-depth profile from Equation 4.2 to

$$\tau(v) = \left(\frac{T_\mathrm{p}}{T_\mathrm{s}}\right) \exp\left(-\frac{(v - v_\mathrm{p})^2}{2\sigma^2}\right), \tag{4.3}$$

where T_p is the peak temperature and $\tau_\mathrm{p} = T_\mathrm{p}/T_\mathrm{s}$ is the peak optical depth at v_p. In terms of the column density, $T_\mathrm{p} = 5.49 \times 10^{-19} N_\mathrm{HI}/\sqrt{2\pi}\sigma$. For opaque HI, the observed peak temperature will converge to T_s, while the optically-thin emission will converge to T_p. Using this change of variables removes the added covariance between T_s and σ from the thermal and non-thermal line width components, and introduces a parameter T_p with a similar range to T_s, unlike N_HI which varies over several orders of magnitude. The optically-thin regime is then primarily set by the peak optical depth, $\tau_\mathrm{p} = T_\mathrm{s}/T_\mathrm{p}$.

We keep similar constraints on the velocity of the line centre. Due to the finer spectral resolution of our data, a limit of ± 5 km s^{-1} translates to ± 12 channels.

4.3.2 Multi-gaussian model

A second model for the observed HI spectrum shapes is simply a sum of Gaussian components, equivalent to the optically-thin limit of Equation 4.1:

$$T_\mathrm{b}(v) = \sum_i^N T_{\mathrm{b,i}}(v) = \sum_i^N T_{\mathrm{p,i}} \exp\left[-\frac{(v - v_{0,\mathrm{i}})^2}{2\sigma_\mathrm{i}^2}\right], \tag{4.4}$$

where each component is defined in terms of its peak temperature ($T_{\mathrm{p,i}}$), central velocity ($v_{0,\mathrm{i}}$), and line width (σ_i).

Rohlfs et al. (1972) first point out that the Milky Way HI spectra with flat tops could also be explained by ~ 2 Gaussian components blended together. The difficulty in distinguishing between an optically-thick component and multiple Gaussian components lies in the flexibility of a multi-Gaussian model: two or more Gaussians can be arranged to explain many spectral shapes and do not form an orthogonal basis set. In this section, we describe our approach

for determining an appropriate number of Gaussian components to avoid overfitting, while accounting for spatial continuity to distinguish between one or more overlapping Gaussians.

Several recent studies present algorithmic approaches to (semi-)automated spectral-line fitting fitting (e.g., Haud, 2000; Lindner et al., 2015; Henshaw et al., 2016; Marchal et al., 2019; Keown et al., 2019; Sokolov et al., 2020). The method we use here combines elements from some of these studies to optimize the fitting for our particular data sets, though we highlight that we do not explicitly include spatial information in the fitting (§4.3.2.3). From our early testing trying different methods, we note that most spectral-line fitting algorithms will likely require at least small changes to be optimized for a particular data set or science goal. However, the publicly available code from these projects enables specific optimization with relative ease.

Our multi-Gaussian fitting method has three stages: (1) identifying the number of components and initial parameter guesses (Lindner et al., 2015); (2) an initial non-least squares optimization and internal model selection test; and (3) a nearest neighbour model selection test and final fit. Each stage is described in the preceding sections.

4.3.2.1 Number of components and parameter estimates

The biggest issue in fitting a multi-Gaussian model is choosing the appropriate number of components to fit the spectrum. As mentioned above, part of this difficulty is the correlation between different components. To overcome this issue, the number of components and initial estimates for the components can be recovered by using information from the spectrum shape.

Here, we use the "derivative spectroscopy" used in `Gausspy` following the definition given in Lindner et al. (2015), which we briefly describe below. Estimates of the derivative from the finite-difference between channels are strongly susceptible to noise. To overcome this, a smoothing method can be used to include the information over many channels at each point in the spectrum. Common approaches for this smoothing include Gaussian or median

smoothing (Riener et al., 2020) or total-variation regularization (Lindner et al., 2015). From this smoothing, derivatives provide information to locate and estimate components. A peak in a spectrum corresponds to a (i) negative second derivative, (ii) a zero crossing in the third derivative, and (iii) a positive fourth derivative (Lindner et al., 2015).

Previous works using `Gausspy` identify Gaussian components in two-steps, first by identifying narrow components with a small smoothing length, followed by identifying wide components with a large smoothing length. Previous papers using this method identified optimal smoothing lengths using gradient descent from a training set of spectra (Lindner et al., 2015; Murray et al., 2018; Riener et al., 2020). In our data, we assess the performance of a two-stage identification by-eye on a small set of spectra while varying both the narrow and wide smoothing lengths. Our data contain extremely wide ($\sigma > 30$ km s^{-1}) features, particularly in M31 where the line-of-sight depth through the disc is large, and we find that the two-stage approach used in previous works does not consistently identify spectral features across this large range in line widths.

To account for the wide range in spectral shapes in our data, we modify the identification procedure in `Gausspy` to consider a range of smoothing lengths. From smallest to largest smoothing scale, we identify and estimate that component parameters using the `Gausspy` procedure given above. For progressively larger smoothing scales, wider components are added from the remaining residual. This procedure is the same as previous works using `Gausspy`, just with additional smoothing lengths included.

4.3.2.2 Initial fit and internal model selection

The parameter estimates from the first step are crucial for the multi-component Gaussian fit to converge to a reasonable solution. The hope is that the initial parameter estimates start the minimization algorithm close enough to the global minimum that it will converge quickly and not fall into a local minima that can result from the correlations between Gaussian components. However,

the approximations in step one to get parameter estimates are not always robust and can lead to (i) fits converging to vastly different parameters from the starting point, or (ii) incorrect number of components or their placement. To account for this, we first fit the spectrum using the estimates from stage one, and then perform an internal model selection to ensure the resulting fit is valid. Similar to `Gausspy` (Lindner et al., 2015), we fit spectra using the Levenberg-Marquardt method implemented in the python package `lmfit`[2].

The model selection used here, and in the following sections, relies on the Bayesian information criterion (BIC) fit statistic (Schwarz, 1978). The BIC is defined as a likelihood plus penalization term, where the latter increases in value with the number of free parameters to avoid overfitting:

$$\mathrm{BIC} = \ln(m)\ k - 2\ \ln(\hat{L}), \tag{4.5}$$

where m is the number of data points (velocity channels), k is the number of free parameters, and $\hat{L}$ is the likelihood function. We assume the data uncertainties (σ_n) are independent and normal, and so the log-likelihood function has the standard form of:

$$\ln(\hat{L}) = -\frac{1}{2\sigma_{\mathrm{rms}}^2} \sum_{j=1}^{m} [y_j - T_{\mathrm{b}}(v_j)]^2 + C(\sigma_{\mathrm{rms}}), \tag{4.6}$$

where m is the number of velocity channels, and y_j is the brightness temperature at velocity v_j in the spectrum. The $C(\sigma_{\mathrm{rms}})$ term is a constant term that only depends on the noise, which is a constant for all model comparisons. The preferred model minimizes the BIC. We choose the BIC since it penalizes additional free parameters more strongly than other common statistics (e.g., Akaike information criterion).

There are two steps to the internal model selection:

1. The integral of each component must exceed the 5σ rms noise for a resolved Gaussian line, which removes spurious narrow features due to noise. Following Koch et al. (2018b), we define a resolved line as having

[2] `lmfit` wraps and extends the optimization algorithm from the `scipy` library (Newville et al., 2020).

> 5 channels ($5 \times 0.42 = 2.1$ km s^{-1}) across the full-width-half-max (FWHM). For M31 and M33, this restriction requires an integrated intensity of $\sqrt{\pi/(4 \ln 2)}(2.1 \text{ km s}^{-1})(5\sigma_{\rm rms}) = 31$ and 19 K km s^{-1} respectively. Components below these limits are removed and the spectrum is refit with fewer components.

2. We then iteratively remove components from the fit until the $\Delta\text{BIC}_{N_i,N_f} = \text{BIC}_{N_i} - \text{BIC}_{N_f} < 10$ where N_f and N_i are the number of components before and after removing one component (hence, $N_i > N_f$). Components are removed in lowest to highest integrated intensity. Following Kass & Raftery (1995), we consider $\Delta\text{BIC}_{N_i,N_f} > 10$ to indicate a strong preference for a more complex model with the initial N_i number of components. We highlight that this comparison process continues, if needed, to the $N_f = 0$ model, i.e., no signal in the fitted spectrum.

The updated multi-Gaussian fit uses the minimum model that passes both selection tests.

4.3.2.3 Nearest neighbour model selection and final fit

We introduce one additional step to produce the models used later in this paper. Because we fit each spectrum independently of its neighbours, the fits do not account for the correlation of nearby pixels and models of neighbouring pixels may differ. These small-scale differences are particularly an issue when different numbers of Gaussian components are found, and when strongly overlapping components converge to different solutions due to the large covariance in their parameters.

We account for these differences by comparing each spectrum to the properties of its nearest neighbours. We check and attempt to correct for differences by:

1. We refit the spectrum using the model with the lowest BIC from the nearest 8 pixels. We replace the spectrum's fit if it reduces $\Delta\text{BIC}_{k,l} > 10$,

indicating a strong preference for the neighbouring model at pixel k over the original model at pixel l (Kass & Raftery, 1995).

2. We also refit and compare the spectrum's fit when its neighbours have a different number of Gaussian components. In this case, we refit the spectrum using the models with the lowest and highest number of components. We then re-impose the selection criteria from §4.3.2.2 and replace the solution until $\Delta\mathrm{BIC}_{N_i,N_f} < 10$. Where applicable, both steps are applied.

We apply these checks in a forward and reverse direction, looping through the spatial positions of valid fits along rows of Right Ascensions, followed by column in Declination.

We find that this procedure strongly encourages coherent spatial solutions with a consistent number of components in the model. This procedure is particularly important for spectra with weak signal relative to the noise, where the initial component guesses from §4.3.2.1 are more susceptible to noise. Though we do not make use of this product here (§4.3.3), the models after this neighbour comparison produce far more complete and coherent model of the Milky Way HI foreground in the M31 field. Similar neighbour comparisons are used in other multi-component Gaussian modeling (Haud, 2000; Riener et al., 2020).

Finally, we note that these nearest neighbour comparisons are not ideal for creating spatially-coherent models. Correctly accounting for these differences requires the model to explicitly include its nearest neighbours and be fit simultaneously (Marchal et al., 2019). This simultaneous spatio-spectral modelling is critical when modelling blended components. However, the large data cubes we use here would require significant computational time to model in this manner. We will explore more coherent fitting methods in future work.

The multi-Gaussian models following the neighbour comparisons are the final models we use in our model comparisons (§4.4).

4.3.3 Excluding foreground and off-rotation components

M33 and, primarily, M31 overlap spectral with Milky Way 21-cm HI emission. Further, our data sensitivity is sufficient to detect some off-rotation emission components (e.g., high velocity clouds; Koch et al., 2018c). These sources are contaminants for the model comparison we propose here since only the multi-Gaussian model can account for these additional features.

We remove Milky Way foreground and off-rotation 21-cm HI in two ways. First, we assume the multi-Gaussian model correctly accounts for all detected components, regardless of their source. We then remove components from the multi-Gaussian model which deviate from the centroid velocity by a velocity range Δv. Due to the different galaxy inclinations, we choose different Δv for the galaxies; $\Delta v = 50$ km s^{-1} for M33 (lower inclination) and $\Delta v = 80$ km s^{-1} for M31 (higher inclination). We find that these choices appropriately remove off-rotation features in M33 (Koch et al., 2018c) and remove the majority of Milky Way foreground in M31.

Second, we impose an additional masking to M31 where the red-shifted tip clearly blends with the Milky Way foreground and is not distinguished by the first step. We use the interactively-selected M31 mask described in Section 4.2 to remove components whose velocity centre is outside of the mask. Our results in §4.4 do not show a systematic trend near the red-shifted tip of M31's emission suggesting this separation criterion effectively excludes foreground HI. The spatial footprint of our M33 observations do not include spectra that are strongly blended with Milky Way HI, and so this second step is not applied to the M33 fits.

This separation is imposed on all results in §4.4. From this separation, we re-compute the BIC statistic for the multi-Gaussian fits without the foreground/off-rotation components included in the model. These recomputed BIC values are used in all of our results (§4.4).

We do not, however, recompute the BIC values for the opaque HI fits. In the vast majority of the fits, the BIC statistic did not vary significantly. The

lack of change is due to the original fits not accounting for off-rotation features, since the opaque model is limited to a single component. Because the revised multi-Gaussian models similarly do not include these features, the BIC values are similarly biased.

In the following section, we describe the BIC *difference* (ΔBIC) as the comparison between the models. The removal of foreground/off-rotation components in the multi-Gaussian model effectively equates the bias in the BIC with the opaque HI model, making the ΔBIC a fair comparison for the opaque HI model.

4.4 Which model is preferred for HI spectra?

In this section, we describe the model selection test we use to compare fits between the multi-Gaussian and opaque models, synthetic tests of the model selection test, and the results of the test.

4.4.1 Model comparison using BIC

We use the difference in BIC (ΔBIC) to compare the multi-Gaussian and opaque HI models. Specifically, we define the BIC difference as the multi-Gaussian BIC subtracted from the opaque BIC:

$$\Delta\mathrm{BIC} = \mathrm{BIC}_{\mathrm{Gauss}} - \mathrm{BIC}_{\mathrm{Opaque}}. \tag{4.7}$$

Since a minimum BIC is optimal (Equation 4.5), $\Delta\mathrm{BIC} > 0$ indicates a preference for the opaque model and $\Delta\mathrm{BIC} < 0$ shows a preference for the multi-Gaussian model. Following §4.3.3, the BIC for the multi-Gaussian models excludes foreground and off-rotation components.

Similar to selecting the number of components for the multi-Gaussian model (§4.3.2), we choose the BIC statistic for comparing the models since it penalizes additional free parameters more strongly than other similar statistics (e.g., the Akaike Information Criteria; AIC). Because each Gaussian component adds an additional 3 free parameters, the BIC is more likely to prefer a simpler model

than the AIC, though the statistics will likely prefer a similar model in many cases.

Following Kass & Raftery (1995), we consider $|\Delta\text{BIC}| > 10$ to be strong evidence for a model preference, with smaller values indicating a weak preference that may be spurious. Many of the comparisons exceed $|\Delta\text{BIC}| > 10$ and so our results are not affected by the chosen threshold.

4.4.2 Producing a synthetic fit sample

The model comparison we outline in the previous section is a relative comparison between two specific models. To ensure our model selection test is sensitive to the differences between the two models, we produce a population of 20,000 synthetic spectra using the opaque model (Equation 4.1) with randomly drawn parameters within the parameter range used for the opaque HI observational fits.

For each random draw of parameters, we evaluate the model and fit the synthetic spectrum in the following steps:

1. We produce the true model with Equation 4.1, parameterizing the optical depth with Equation 4.3, evaluated over a spectral axis from −200–200 km s^{-1} with 0.1 km s^{-1} channels. All synthetic spectra are centered at 0 km s^{-1} since the line centre does not affect the model comparison.

2. We then downsample the model by averaging over 0.42 km s^{-1} to match the observations, which produces an idealized observed spectrum (Koch et al., 2018b).

3. We add Gaussian noise to the downsampled model with $\sigma_{\rm rms} = 2.8$ K, matching the noise in the M31 HI map. This produces the equivalent observed spectrum to be fit.

4. Both models are fit to the synthetic observed spectrum using the same parameters and limits used for the observational fits.

The BIC fit statistic and fitted parameters are kept to compare with the distribution of observational properties. Because the models are drawn from the opaque model, From our synthetic fit distribution, we correctly recover a preference for $\Delta\mathrm{BIC} > 0$, with a larger BIC scatter where the opaque model approaches the optically-thin limit. This result validates that our model selection test has sufficient power to distinguish between the multi-Gaussian and opaque models for the observed HI spectra.

4.4.3 A strong preference for multi-gaussian models

Using the comparison framework from §4.4.1, we examine ΔBIC for the population of HI spectra in M31 and M33. Our comparisons in this section demonstrate a strong preference for a multi-Gaussian model.

We first visually examine the fits for a few spectra, chosen specifically to show a range in the inferred peak optical depth from the opaque model. Figure 4.3 shows four example spectra with fits to both models (chosen from a subset of the spectra in Figure 4.2). Each panel in the figure shows the key fit results, including the number of Gaussians (N_{Gauss}), inferred peak optical depth (τ_{p}), and the BIC values are shown for each fit. Following visual inspection, spectra (a–c) clearly demonstrate multiple peaks and strongly prefer the multi-Gaussian model based on the smaller BIC value ($\Delta\mathrm{BIC} < 0$).

The final spectrum (d) has a single bright "pedestal," unlike the other example spectra, and its flat-top is qualitatively the expectation for the opaque HI model (Braun et al., 2009). We find, however, that the multi-Gaussian model has a much smaller BIC value and is therefore preferred.

To test this apparent preference for the multi-Gaussian model, we create ΔBIC maps of both galaxies. Figure 4.4 shows ΔBIC in M31 and M33 where valid fits are found (§4.3). The colour is centered such that $\Delta\mathrm{BIC} = 0$ is shown in gray. Throughout the maps, we find a strong preference for the multi-Gaussian model ($\Delta\mathrm{BIC} < 0$). Most locations where the opaque model is preferred ($\Delta\mathrm{BIC} > 0$) correspond to regions with fainter HI intensity (Figure 4.1). This is in contrast with previous works (Braun et al., 2009; Braun, 2012)

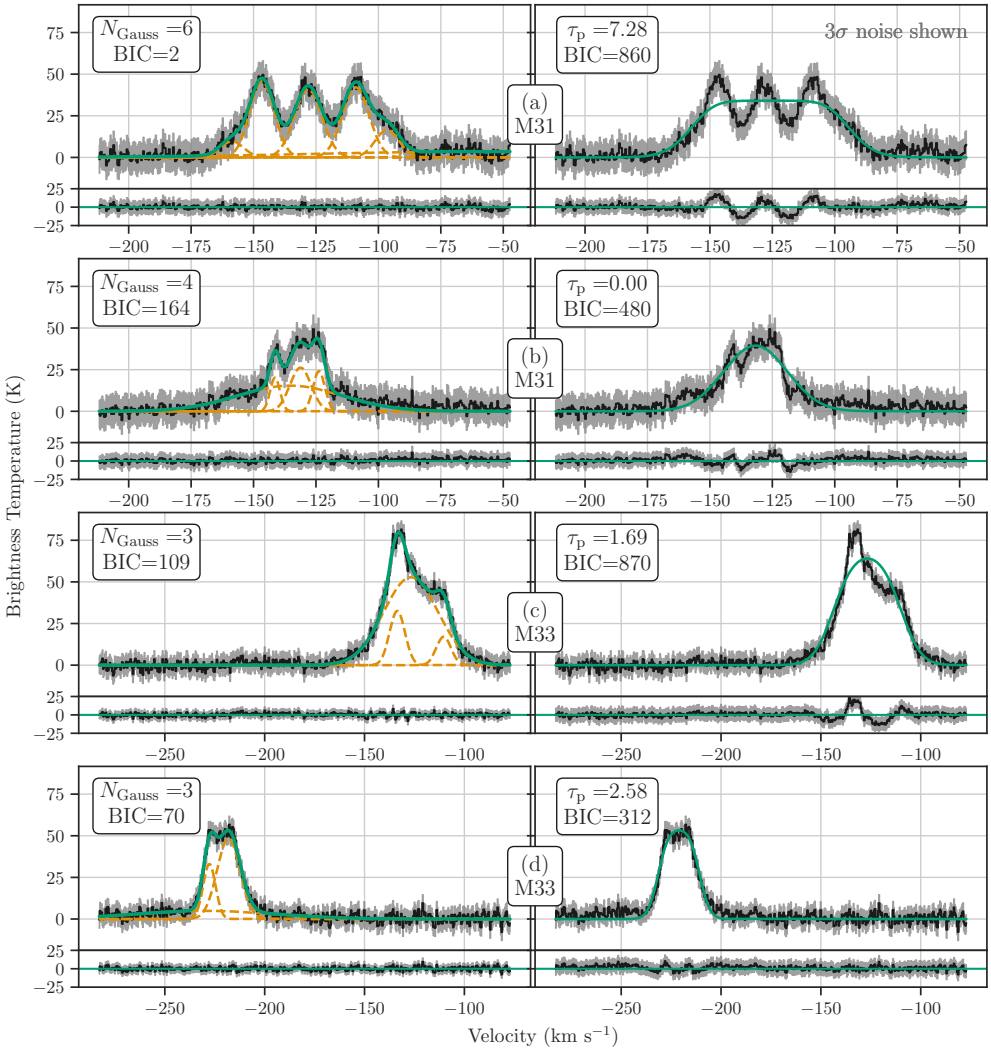

Figure 4.3: Four HI spectra shown in Figure 4.2 with their multi-Gaussian (left) and opaque (right) model fits. The fit residuals are shown in the panel below each spectrum and fit, with the 3σ rms noise shown in gray. The number of Gaussians and the peak opacity ($\tau_{\rm p}$) are shown for each model, respectively, as well as the BIC statistic for each fit. In all the examples, the BIC statistic in minimized for the multi-Gaussian model and is therefore the preferred model. Spectra (a), (c), and (d) are all examples where the opaque model predicts a large opacity ($\tau_{\rm p} > 1$), despite the large residual from the fit. This includes spectrum (d), which visually has a near "flat-top." Despite the larger number of free parameters (9 vs. 4), however, the multi-Gaussian model is still preferred. We assess the validity of this visual inspection by comparing the BIC statistics for the entire population of fitted spectra.

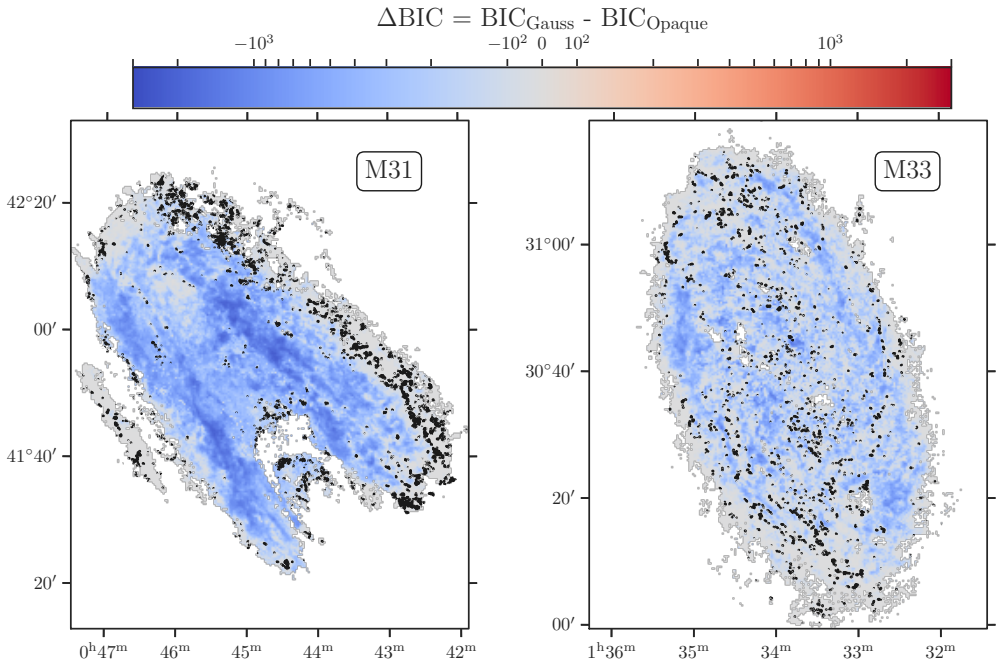

Figure 4.4: Δ BIC values comparing the optically-thick and multi-Gaussian models. Negative values (blue) indicate where $\Delta\text{BIC} < 0$, indicating a preference for the multi-Gaussian model, and positive values (red) show where $\Delta\text{BIC} > 0$ and the optically-thick model is preferred. The black contours show lines of $\Delta\text{BIC} = 0$. Multi-Gaussian models are preferred in nearly every spectra that has bright HI emission (i.e., negative BIC is closely tracing the integrated intensity; Figure 4.1).

where opaque HI regions are highlighted where HI intensity is large (e.g., Braun, 1997).

We see a general trend towards $\Delta\text{BIC} = 0$ near the edge of HI detections in both maps, where the noise increases (e.g., in M31) and the HI emission is fainter. At lower signal-to-noise, the multi-Gaussian model tends to have a single component, and therefore is similar to the opaque model in the optically-thin limit. In these regions, there is a small tendency for $\Delta\text{BIC} < 0$ since a single Gaussian component has three free parameters and the optically-thin limit of the opaque model includes a fourth unconstrained parameter, $T_{\rm s}$, and so the former model is preferred.

To quantify the preference for the multi-Gaussian model, we compare ΔBIC

to the apparent "dark" HI intensity inferred by the opaque HI model. The dark HI intensity, $I_{\rm dark}$, is the integrated difference between the optically-thin and opaque for the model fit by Equation 4.1:

$$I_{\rm dark} = \int \left[T_{\rm p} {\rm exp}\left(-\frac{(v-v_0)^2}{2\sigma_{\rm i}^2} \right) - T_{\rm s} \left(1 - \exp\left[-\tau(v)\right]\right) \right] {\rm d}v, \tag{4.8}$$

where $\tau(v)$ is given in Equation 4.3. We use $I_{\rm dark}$ as it integrates over the three parameters that determine the deviation of the opaque spectrum from a Gaussian shape: $T_{\rm p}$, $T_{\rm s}$, and σ. As we note in §4.3.1, these parameters tend to be highly correlated, so using a metric that integrates over all three measures their combined affect on the line shape. Further, $I_{\rm dark}$ converges to 0 as as $T_{\rm s}$ increases and the model approaches the optically-thin limit.

We stress that using $I_{\rm dark}$ as a physical quantity depends on the opaque model fitting a spectrum well. As we already suggest, the opaque model is often not strongly preferred compared to the multi-Gaussian model. Because of this, we adopt $I_{\rm dark}$ as a useful *metric* to compare with, rather than a physical quantity.

Figure 4.5 shows the ΔBIC statistic distributions for M31, M33, and the synthetic fits plotted against $I_{\rm dark}$. The top row shows the distribution of all fitted spectra. Consistent with the maps in Figure 4.4, most spectra in M31 and M33 show a strong preference for the multi-Gaussian model. When $I_{\rm dark}$ approaches zero, the distribution of ΔBIC values broadens. This broadening shows where the opaque model approaches the optically-thin solution, where some spectra are better modelled with multiple components (e.g., (b) in Figure 4.3) or where the signal-to-noise is low and both models approach a similar fit, as we see at the edges of the maps in Figure 4.4. The fraction of the distribution that prefers one model versus the other is shown in Table 4.2, split into the categories defined in §4.4.1.

Our control sample for these comparisons is the synthetic fits from §4.4.2. Using the same fitting methods, the synthetic distribution shows a stark difference from the observational fits, with the majority of spectra preferring the opaque model (ΔBIC > 0). This is the expected result since the synthetic

spectra are drawn from the opaque model (Equation 4.1). Correctly recovering the preference for the opaque model provides an important test for fits by demonstrating that, with the observational noise and spectral resolution, our fitting methods are sensitive to the difference between an opaque and multi-Gaussian model.

We next examine whether the preference for a multi-Gaussian model persists if we consider only the spectra where the opaque model recovers $\tau_{\rm p} > 1$ (i.e., all locations of inferred optically-thick HI emission). This comparison is shown in the bottom row of Figure 4.5. Because only a sub-set of the ΔBIC are shown, the percentiles defining the contours in the plots are changed, leading to some difference in the observed distributions.

We find the same general trends when considering only the apparently opaque spectra. If the observed HI spectra include a population of well-fit opaque model spectra, we expect to see the distribution shift towards $\Delta{\rm BIC} > 0$, particularly where $I_{\rm dark}$ is largest (i.e, the most opaque emission inferred from the model). We do find this trend for the synthetic fit distribution, however, it is not apparent for large $I_{\rm dark}$ in M31 and M33. We compute the percent of spectra for different model preferences (Table 4.2), which quantitatively shows the behaviour apparent in Figure 4.5.

Finally, we include one further test that explicitly accounts for the single-component limitation of the opaque model (Equation 4.1). Spectrum (d) in Figure 4.3 demonstrates the limits of the opaque model well. The multi-Gaussian model correctly includes a wide component that accounts for faint "tails" in the emission. The opaque model cannot account such features and is therefore more likely to have a larger BIC, even if the model has fewer parameters.

To address this limitation, we recalculate the BIC statistic for both models limited to the channels where $\tau(v) > 0.5$ (i.e., where we expect the opaque model to identify a flattened top corresponding to opaque emission). This limit is chosen based on our synthetic fits where the opaque model measurably deviates from a Gaussian line shape given the noise level and spectral resolu-

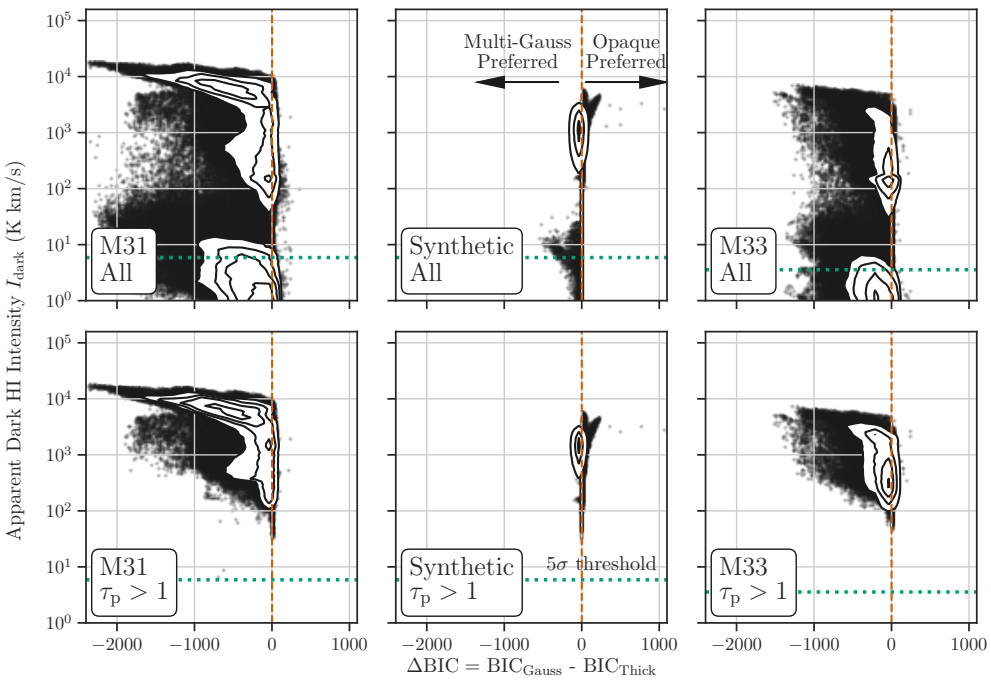

Figure 4.5: Δ BIC of the optically-thick BIC subtracted from the multi-Gaussian model plotted against the apparent dark HI integrated intensity from the optically-thick model (Equation 4.8; i.e., optically-thin equivalent line subtracted by the optical depth-corrected profile). Δ BIC < 0 indicates a preference for the multi-Gaussian model, while Δ BIC > 0 are fits with the optically-thick model preferred. The panels show the plots for M31 (left), synthetic fits generated from Equation 4.1 (centre), and M33 (right). The top row show the results for all spectra, and the second row are fits where the peak optical depth from the optically-thick model fit is $\tau_p > 1$. Each panel shows the 1–4σ_{rms} levels within the contours, and the data points show fit results outside of the 4σ_{rms} contour. The dashed vertical line indicates where ΔBIC $= 0$, and the horizontal dotted line is the 5σ_{rms} sensitivity in integrated intensity for a single 0.42 km s^{-1} spectral channel. The observed fit results strongly prefer a multi-Gaussian model (ΔBIC < 0), while the synthetic sample correctly prefers the optically-thick model used to generate the synthetic spectra (ΔBIC > 0). These differences become more apparent when *only* optically-thick model fits are shown in the second row, demonstrating the opposite trend expected for fits to the optically-thick model. Despite having many more free parameters, the multi-Gaussian model is consistently preferred.

tion in our observations. For the multi-Gaussian model, we consider only the components where $> 75\%$ of their intensity is within range with $\tau(v) > 0.5$.

Figure 4.6 shows the resulting distribution of ΔBIC limited to $\tau(v) > 0.5$. As expected, the synthetic fit distribution shows a preference for the opaque model ($\Delta\text{BIC} > 0$), though a large fraction of the ΔBIC fits fall into the category of $-10 < \Delta\text{BIC} < 10$ where the preference is not clear (Table 4.2). This is due to limiting the number of velocity channels where we calculate the BIC. Thus this trend is expected as the ΔBIC becomes a noisier metric.

Despite ΔBIC being a noisier measure in this test, we still find a strong preference for multi-Gaussian models in both M31 and M33. This result demonstrates that the preference for multi-Gaussian models does not result from the limitation of a single opaque component in Equation 4.1.

This continued result of $\Delta\text{BIC} < -10$ for observed spectra strongly suggests that most HI spectra in M31 and M33 are best modelled with a set of Gaussian components.

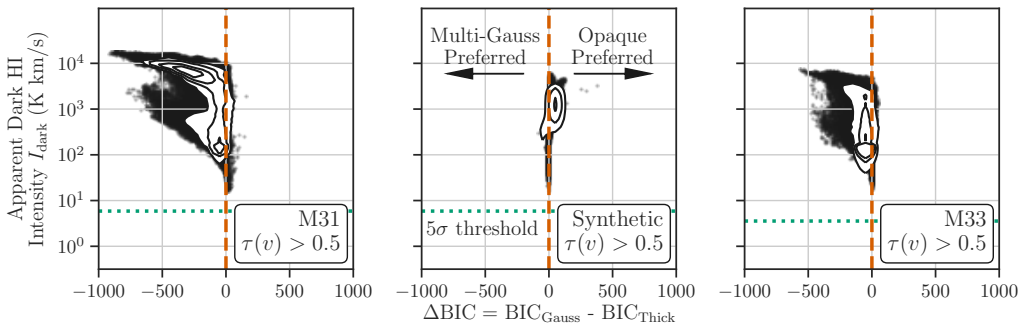

Figure 4.6: Δ BIC plotted against the apparent dark HI integrated intensity from the opaque model, similar to Figure 4.5. Here, we show only the fit statistics calculated where $\tau(v) > 0.5$ from Equation 4.3. This comparison controls for the single-component limitations of the opaque model, which cannot account for multiple peaks in a spectrum. By computing the BIC statistics only where $\tau(v) > 0.5$, this comparison solely compares the models where the "flat-top" is measurable in the spectra. However, these results are similar to the ΔBIC over the whole spectra (Figure 4.5). The synthetic fits, drawn from Equation 4.1 with observational noise added, confirm that our observations are sensitive to opaque and the fitting procedures correctly find $\Delta\text{BIC} > 0$ on average. The observed spectra in M31 (left) and M33 (right), however, have $\Delta\text{BIC} < 0$, indicating a continued strong preference for the multi-Gaussian model.

		Multi-Gaussian preferred	No strong preference	Opaque preferred
		% with $\Delta\mathrm{BIC} < -10$	% with $-10 < \Delta\mathrm{BIC} < 10$	% with $\Delta\mathrm{BIC} > 10$
M31	All (Figure 4.5; top)	82.0	17.5	0.5
	$\tau_\mathrm{p} > 1$ (Figure 4.5; bottom)	88.7	10.5	0.8
	$\tau(v) > 0.5$ (Figure 4.6)	83.7	15.4	0.9
M33	All (Figure 4.5; top)	80.2	19.6	0.2
	$\tau_\mathrm{p} > 1$ (Figure 4.5; bottom)	83.2	16.0	0.7
	$\tau(v) > 0.5$ (Figure 4.6)	74.4	25.0	0.6
Synth.	All (Figure 4.5; top)	1.1	49.9	49.0
	$\tau_\mathrm{p} > 1$ (Figure 4.5; bottom)	1.4	35.0	63.6
	$\tau(v) > 0.5$ (Figure 4.6)	0.6	78.5	20.9

Table 4.2: Percent of spectra for different ranges in ΔBIC in the M31, M33, and synthetic fit distributions. We define three ranges: (i) a strong preference for the multi-Gaussian model ($\Delta\mathrm{BIC} < -10$); (ii) spectral model without a strong preference for either model ($-10 < \Delta\mathrm{BIC} < 10$); and (iii) a strong preference for the opaque model ($\Delta\mathrm{BIC} > 10$). We compute the percentage in the ranges for the three tests including all fits and only those with $\tau_\mathrm{p} > 1$ (Figure 4.5), and the BIC computed only where $\tau(v) > 0.5$ (Figure 4.6). Our synthetic fit distribution correctly shows a preference for the opaque model, from which the spectra are drawn from. However, the vast majority of observational fits in M31 and M33 prefer the multi-Gaussian model for all three comparisons.

4.5 Discussion

Our results suggest a vastly different picture than the dominant opaque HI from Braun et al. (2009). Here, we discuss reasons for this discrepancy, compare the opaque mass correction factors that we determine, and discuss the limitations of our multi-Gaussian modeling.

4.5.1 Discrepancies with previous interpretations

Our results are clearly discrepant from those in Braun et al. (2009) and Braun (2012). These discrepancies result from (i) improved spectral resolution and a higher sensitivity on smaller linear scales and (ii) differences in the modelling methodology.

Table 4.1 summarizes the key differences in our observations versus those from Braun et al. (2009) and Braun (2012). Our observations have a similar sensitivity at moderately higher spatial resolution. However, we expect that the key difference is the finer spectral resolution. As we point out in §4.1, narrow resolved line shapes require ~ 2 channels per line width (σ; Equation 4.4 Koch et al., 2018b), and the range of channel widths from Braun (2012) correspond to thermal temperatures of 1300–3600 K. The 0.42 km s^{-1} resolution from our data is capable of resolving line widths corresponding to 120 K. Thus, the differences in the observations provide some reason for the additional spectral complexity we find relative to Braun et al. (2009). In §4.6.1, we describe an extension of our current work to identify the narrow HI components from the multi-Gaussian modelling.

A second difference in the data sets is the correlation between adjacent spectral channels. While we do not explicitly test the M31 Westerbork data from Braun et al. (2009), their Figures 9, 10, and 14 highlight spectra with clear strong channel-to-channel correlations, suggesting that the effective spectral resolution may be lower than the channel width. This will effectively smooth over narrow spectral features, similar to a moving average filter, and given the spectral shapes we identify (Figure 4.3), is likely to encourage flattened

top profiles in some spectra. We note that highly correlated channels are not apparent in the M33 and LMC data sets in Braun (2012).

The second reason for the discrepancy in our results is from the differences in methodology used for fitting. Braun et al. (2009) do not explicitly model multiple components, and additional residuals after fitting Equation 4.1 are assumed to be optically-thin emission. Further, the model selection criteria they use to distinguish between multiple components is a flat-top is more coarse than our comparison. They define spectra with a reduced chi-square values of $\chi^2_{\rm r} < 25$, based on visual inspection, to be a valid opaque HI fit. However, given there are faint features not modelled in their fits (their Figures 14), it is not clear whether this threshold is sufficient to distinguish between multiple components and faint emission not included in the fit. Braun (2012) reject 4% of their fits in M31 and $< 1\%$ in M33 based on this threshold.

From our fits, we find *no* opaque models would be excluded using the $\chi^2_{\rm r} < 25$ criterion. This difference is likely due to systematics in the different observations as $\chi^2_{\rm r}$ is sensitive to changes in the noise, which is underestimated for highly correlated spectral channels.

This discrepancy further demonstrates the issues with absolute fit statistics when compared between different data sets. It is more robust to use relative fit statistics when comparing models fit to the same data.

4.5.2 Correction factor for optically-thick HI column density

The key outcome from previous studies using the opaque model is the ability to account for "dark" HI over a large spatial area, overcoming limitations of other methods that require HI absorption against background sources. With a constraint on the opacity, the cold opaque atomic gas content can be estimated, yielding a mass correction factor for the atomic ISM relative to the optically-thin mass. In Braun et al. (2009) and Braun (2012), they infer that opaque HI accounts for an additional $\sim 35\%$ of atomic ISM mass in M31 and M33.

We calculate the dark HI intensity from Equation 4.8 and convert to the

atomic mass using the factor 0.0196 $M_\odot\,pc^{-2}$/ (K km s^{-1}) multiplied by the spatial area in each map. Note that this conversion factor includes a 1.36 factor for the He and metals fractions.

Considering only the spectra where the opaque model is preferred (ΔBIC $>$ 0), we find that the correction factor for opaque HI mass is just $\sim 1\%$. However, we expect this small correction factor since most spectra strongly prefer the multi-Gaussian model ($\sim 99\%$), particularly where HI is bright and with the largest optically-thin column density (Figures 4.1 & 4.4) and HI would be more likely to be opaque (e.g., Braun, 1997). Table 4.3 shows the mass correction factors for different selections and the fraction of the fit spectra included.

We stress that our finding of a 1% correction factor does *not* imply that $I_{\rm dark} \sim 0$ in either M31 and M33. Rather, it demonstrates that the dark HI intensity from Equation 4.1 does not provide a good description of HI spectra on $\sim$ 100 pc scales. This follows from our general conclusion that most HI spectra are composed of multiple components, rather than a single cold opaque gas component on these scales.

We further present the mass correction factors for different selection criteria to provide a comparison to the $\sim 35\%$ factor from Braun et al. (2009) and Braun (2012). As shown in Table 4.3, we demonstrate that the choice of model selection or "goodness-of-fit," drastically affects dark HI mass estimates. First, we calculate the mass correction factors from our modelling with no selection criteria, only spectra where the opaque model is preferred (ΔBIC > 0) in our fits, and where the peak opacity from the opaque model fit is $\tau_{\rm p} < 8$. The latter demonstrates how extreme, though allowed, fits to the opaque model are highly sensitive to what is considered a "good" fit. We note that the $\chi^2_{\rm r} < 25$ criterion from Braun et al. (2009) is equivalent to no selection criteria for our fits since none of our fits have $\chi^2_{\rm r} \geq 25$.

With no selection criteria, we recover a correction factor of $\sim 35\%$ for M33, consistent with Braun (2012). In M31, however, we find that many spectra are best fit only by drastically increasing the optical depth above $\tau > 8$, and the

inferred dark HI becomes very large, in some cases exceeding the optically-thin mass. Because of this, we find a mass correction factor of $> 100\%$ relative to the optically-thin integrated intensity estimate. Strictly speaking, such large τ values are acceptable in the model, and Milky Way studies towards extreme star forming regions, like W43, do infer lines-of-sight with a 240% increase in mass compared to the optically-thin mass from the integrated intensity (Bihr et al., 2015). However, such large correction factors seem to be rare in the Milky Way (Wang et al., 2020a).

Figure 4.7 demonstrates that uncertainty on large $\tau_{\rm p}$ fit values is the source of the large mass correction factor in M31. The increase in uncertainty for large $\tau_{\rm p}$ is likely from cases like spectrum (a) in Figure 4.3, where the opaque model fits a wide-envelope that poorly matches the spectrum shape. This uncertainty at large $\tau_{\rm p}$ has a large effect on the inferred dark HI mass. We also note that the fit criterion of $\chi^2_{\rm r} < 25$ does not discriminate these cases for our fits (§4.5.1), and further, that this solution falls within the allowed parameters we and Braun et al. (2009) use. The continuous distribution in $\tau_{\rm p}$, however, suggests there is not a simple cut-off, making estimates of the dark HI mass sensitive to definitions of goodness-of-fit and the parameter space selection.

To show the sensitivity to the parameter space limits, we consider removing fits where the opaque model infers large optical depths ($\tau_{\rm p} > 5$), and approximately where the relative uncertainty increases in Figure 4.7. Table 4.3 shows that the correction factors are drastically lower, with $12^{+46}_{-5}\%$ in M31 and $19^{+78}_{-7}\%$ in M33. Though the correction factors are lower, they are still highly uncertain, as the combined uncertainty on $T_{\rm p}$ and $T_{\rm s}$ produce a non-linear effect on the dark HI mass.

These results demonstrate the large uncertainty on the dark HI mass from the opaque model (Equation 4.1), first from the lack of preferred models compared to the multi-Gaussian model, and second from the large inherent uncertainty on the peak optical depth from the opaque model fits. The latter is highly sensitive to the "goodness-of-fit" threshold and limits on the parameter

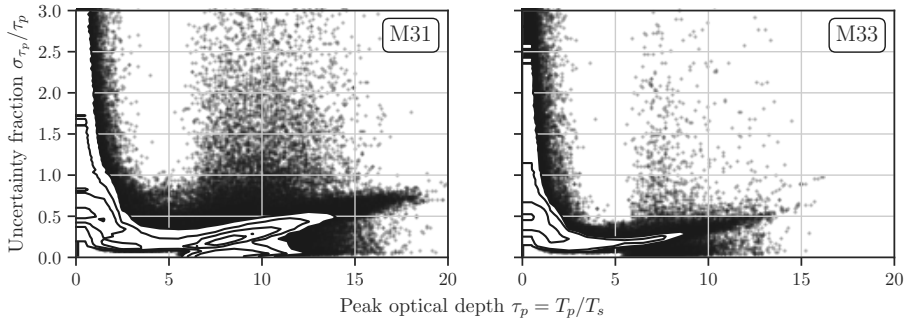

Figure 4.7: Peak optical depth ($\tau_{\rm p} = T_{\rm p}/T_{\rm s}$) versus its uncertainty fraction ($\sigma_{\tau_{\rm p}}/\tau_{\rm p}$) for M31 (right) and M33 (left). Contours and outliers levels are defined in Figure 4.5. In the optically-thin limit, the uncertainty on $\tau_{\rm p}$ is large as $T_{\rm s}$ is unconstrained. However, the uncertainty also becomes larger for $\tau_{\rm p} > 5$, particularly in M31. Large uncertainties at for large $\tau_{\rm p}$ produce large uncertainties in the inferred dark HI mass.

space allowed in the fits. We conclude that dark HI mass estimates from the opaque model should be treated as highly uncertain at best.

4.5.3 Limitations of multi-component Gaussian models

Multi-Gaussian models are inherently uncertain when components overlap because gaussians do not form an orthogonal and independent basis set. This means that the properties of individual components from the multi-Gaussian fit models (Figure 4.3) should be treated carefully for physical interpretations (Rohlfs et al., 1972; Haud, 2000). Our fitting approach uses spatial similarities to improve spectral fitting, which encourages spatial coherence, but remains uncertain where blended components are correlated. We note, however, that this will not affect the overall fit, which is the key for the model comparison we present.

Our multi-Gaussian model also does not identify HISA features, distinguished from the opaque HI model as a clear absorption "dip" like those from Milky Way HI spectra (Gibson et al., 2005). While self-absorption may be detected towards some lines-of-sight, we do not expect our observations to be sensitive to strong HISA features due to the $>$ 60 pc linear resolution.

	Sample Selection	M31	M33
Opt.-Thin Atomic Mass* ($M_\odot$)	–	1.6×10^9	1.0×10^9
Dark Atomic Mass* ($M_\odot$)	All	$2.0^{+2.0}_{-0.7} \times 10^9$	$3.7^{+8.5}_{-1.2} \times 10^8$
% of Spectra		100%	100%
% Mass Difference		118^{+119}_{-40}%	37^{+83}_{-12}%
Dark Atomic Mass* ($M_\odot$)	$\tau_p < 5$	$2.0^{+7.9}_{-0.9} \times 10^8$	$2.0^{+8.0}_{-0.7} \times 10^8$
% of Spectra		85.7%	99.2%
% Mass Difference		12^{+46}_{-5}%	19^{+78}_{-7}%
Dark Atomic Mass* ($M_\odot$)	$\Delta BIC > 0$	$2.1^{+1.5}_{-1.0} \times 10^7$	$8.8^{+4.1}_{-2.9} \times 10^6$
% of Spectra		1.6%	0.8%
% Mass Difference		$1.3^{+0.9}_{-0.6}$%	$0.9^{+0.4}_{-0.3}$%

* Including 1.36 factor for He and metals.

Table 4.3: Dark atomic ISM mass in M31 and M33 estimated with no selection criteria, only spectra with peak optical depths of $\tau_p < 5$ based on the opaque model, and spectra where the opaque model is preferred ($\Delta BIC > 0$; Equation 4.1). We note that all of our fits satisfy the $\chi^2_r < 25$ criterion from Braun et al. (2009), and therefore their selection criterion is equivalent to using all of the opaque HI fits. The table also shows the percent of spectra after the selection criterion is applied, and the percent mass difference compared to the optically-thin mass. The optically-thin atomic ISM mass is calculated from the integrated intensity maps excluding pixels without valid spectral fits. We find large differences in the mass correction factors that depend strongly on the sample selection, and similarly the goodness-of-fit. This strong dependence on sample selection suggests that the dark HI mass from the opaque model (Equation 4.1) is highly uncertain.

Koch et al. (2019b) find the velocity at peak HI and CO temperature are typically consistent within the 2.6 km s^{-1} CO velocity channel size. Strong self-absorption would require additional scatter in this relation, or cold HI line widths narrower than CO, which is not observed in the Milky Way (Wang et al., 2020b). Furthermore, Liu et al. (2019) use CO components to derive HI self-absorption in the LMC and find relatively small corrections on 15 pc scales.

4.6 Summary

Using sensitive 21-cm HI observations at high spectral resolution in M31 and M33, we compare two common spectral models: a multi-Gaussian model and a single-component, opaque-HI model (e.g., Braun et al., 2009). With the improved spectral resolution of our data, we demonstrate a strong preference for the multi-Gaussian model for $> 80\%$ of the spectra. From this model preference, we show that monolithic cold HI clouds are not supported by the data and that the inferred optical depth from an opaque model fit to HI emission alone is subject to significant systematic uncertainty. Our main conclusions from this work are are:

1. Improvements in the spectral resolution of extragalactic 21-cm HI show a wealth of spectra structure. The observations we present extend detailed HI studies that are otherwise largely constrained to within the Milky Way and Magellanic Clouds.

2. We identify a strong preference for a multi-Gaussian model across both M31 and M33 by comparing the difference of the Bayesian Information Criteria (BIC) between the two models. This preference persists when considering only opaque models with a large inferred optical depth ($\tau_{\rm p} > 1$), and when the BIC is computed only over velocities where the opaque model measures $\tau(v) > 0.5$.

3. We compare these results to a synthetic set of 20,000 spectra drawn

from the opaque model (Equation 4.1), with randomly drawn parameters within the observed fit parameter space, which we sample at 0.42 km s^{-1} spectral channels and add Gaussian noise equivalent to the observations. By applying the same fitting procedures as the observed spectra, we demonstrate that our model selection criteria correctly prefers the opaque model over a multi-Gaussian model. This test ensures that our observational comparisons are sensitive enough to distinguish between the two models.

4. Considering only the observed spectra where the opaque model is preferred, we determine that opaque HI, determined solely according to this model, contributes just $\sim 1\%$ of HI mass that is missed compared to the mass assuming the optically-thin assumption, far smaller than the 30% correction inferred from studies which consider only the opaque model on 100 pc scales (Braun et al., 2009; Braun, 2012).

Our results indicate that estimates of the opaque HI mass from the opaque emission model are not reliable on 100 pc scales. The large difference in the inferred dark atomic HI mass highlights the crucial role of model assumptions on interpretations of the atomic ISM. From these results, we suggest that HI opacity is best constrained by comparing HI absorption to background sources with nearby emission (e.g., Dickey et al., 2003), or through HI self-absorption (e.g., Gibson et al., 2005; Wang et al., 2020b).

4.6.1 Continuing work

This chapter describes recent and still unpublished work. For publication, we anticipate expanding the current work in two ways.

First, §4.4 will present more information on the Gaussian components from the multi-Gaussian model. This is important as a key portion of the paper is ultimately how much CNM is in the observed galaxies. We do not expect to directly estimate this fraction, but constraints are plausible based on the fraction of narrow HI components.

The primary difficulties in calculating these constraints are (i) blended Gaussian components, and (ii) the completeness of the multi-Gaussian fitting method. From §4.5.3, blended components are uncertain without explicit spatial coherence in the fitting (Marchal et al., 2019). To address this, we will split the population of components into blended and distinct sets, where the latter produces a distinct shape in the second derivative of the multi-Gaussian model.

The completeness threshold is likely driven by the minimum area constraint for components (§4.3.2.3), which favours bright narrow and faint wide components at the two extremes. This introduces a completeness threshold that varies with the peak temperature and line width, and may be further altered when components are blended. We are measuring the completeness fraction using false source injection of, at first, distinct narrow components on the observed set of HI spectra.

The second expansion of this work will incorporate an estimate of the covering fraction of cool narrow HI components (relative to the observed galaxy area) to compare with Braun (1997) in §4.5. This latter work provides an assumption of a low CNM covering fraction, which Braun et al. (2009) uses as motivation for a single opaque HI component along most lines-of-sight. This comparison requires usable constraints on the population of cool HI components from the first point above.

Chapter 5

Relationship between the Line Width of the Atomic and Molecular ISM in M33

Koch, E.W et al. (2019). "Relationship between the line width of the atomic and molecular ISM in M33," *MNRAS*, 485, 2324–2342.

Abstract

We investigate how the spectral properties of atomic (HI) and molecular (H_2) gas, traced by CO(2-1), are related in M33 on 80 pc scales. We find the HI and CO(2-1) velocity at peak intensity to be highly correlated, consistent with previous studies. By stacking spectra aligned to the velocity of HI peak intensity, we find that the CO line width ($\sigma_{\rm HWHM} = 4.6 \pm 0.9$ km s^{-1}; $\sigma_{\rm HWHM}$ is the effective Gaussian width) is consistently smaller than the HI line width ($\sigma_{\rm HWHM} = 6.6 \pm 0.1$ km s^{-1}), with a ratio of $\sim$0.7, in agreement with Druard et al. (2014). The ratio of the line widths remains less than unity when the data are smoothed to a coarser spatial resolution. In other nearby galaxies, this line width ratio is close to unity which has been used as evidence for a thick, diffuse molecular disk that is distinct from the thin molecular disk dominated by molecular clouds. The smaller line width ratio found here suggests that M33 has a marginal thick molecular disk. From modelling individual lines-of-sight,

we recover a strong correlation between HI and CO line widths when only the HI located closest to the CO component is considered. The median line width ratio of the line-of-sight line widths is 0.56 ± 0.01. There is substantial scatter in the HI–CO(2-1) line width relation, larger than the uncertainties, that results from regional variations on < 500 pc scales, and there is no significant trend in the line widths, or their ratios, with galactocentric radius. These regional line width variations may be a useful probe of changes in the local cloud environment or the evolutionary state of molecular clouds.

5.1 Introduction

Across large samples of nearby galaxies, several studies show a tight correlation between the surface density of molecular (H_2) gas and star formation rate (SFR) surface density (Kennicutt, 1998; Leroy et al., 2008; Bigiel et al., 2008; Kennicutt et al., 2011), and a lack of correlation with the atomic (HI) gas surface density (Bigiel et al., 2008; Schruba et al., 2011). This result shows that star formation is primarily coupled to the molecular gas, rather than the total (HI + H_2) gas component.

A critical, potentially rate-limiting, step in the star formation process is then the formation of molecular gas. Several mechanisms have been proposed that lead to conditions where molecular gas can readily form (Dobbs et al., 2014). These mechanisms for forming the molecular interstellar medium (ISM) are predicted to act over scales ranging from individual molecular clouds to galactic scales. Recent star formation models have sought to predict the atomic-to-molecular gas fraction from the local environment properties (Blitz & Rosolowsky, 2006; Krumholz et al., 2009; Ostriker et al., 2010; Krumholz, 2013; Sternberg et al., 2014; Bialy et al., 2017) and recover observed properties to within a factor of a few (Bolatto et al., 2011; Jameson et al., 2016; Schruba et al., 2018).

To observe signatures of the molecular ISM, we require observations that resolve giant molecular cloud (GMC) scales (< 100 pc) in both the atomic

and molecular gas. Only within the Local Group can current 21-cm telescopes resolve GMC scales, making studies of M33, M31, and the Magellanic Cloud critical for understanding how the molecular ISM forms. In this paper, we use HI and CO(2-1) observations of M33 with a resolution of 80 pc to study the spectral properties of the atomic and molecular ISM.

Previous high-resolution studies of HI and CO, used as a tracer of H_2, in the Local Group have identified spectral-line properties that are correlated between these tracers. Wong et al. (2009) and Fukui et al. (2009) compared the HI to CO properties in the Large Magellanic Cloud (LMC) on 40 pc scales. They found that HI and CO spectral properties are correlated, with a close relationship between the velocities at peak intensity and a suggestive correlation between the HI and CO line widths. However, they also found that the HI temperature and column density are poor predictors for the detection of CO, suggesting that a significant amount of HI emission arises from atomic gas not associated with the molecular gas.

On larger scales (> 100 pc) where individual clouds are unresolved, several studies have found evidence of a large-scale molecular component, possibly unassociated with CO emission from GMCs on small scales. Garcia-Burillo et al. (1992) found CO emission ~ 1 kpc from the plane of the disk in the edge-on galaxy NGC 891, providing direct evidence for a "molecular halo." More recently, Pety et al. (2013) find evidence for a diffuse molecular disk based on interferometric data (~ 50 pc resolution) recovering only $\sim 50\%$ of the flux from single-dish data. They suggest that the remaining emission is filtered out by the interferometer and must be from larger scales. Using a similar comparison between interferometric and single-dish data, Caldú-Primo et al. (2015) and Caldú-Primo & Schruba (2016a) identify a wide velocity component in the CO that is only recovered in single-dish data on scales > 500 pc.

There is also growing evidence for a significant diffuse molecular component in the Milky Way. Dame & Thaddeus (1994) find excess CO emission in the line wings that may be similar to the wide velocity components in nearby galaxies (Caldú-Primo et al., 2015; Caldú-Primo & Schruba, 2016a). Roman-

Duval et al. (2016) find 25% of the Milky Way molecular gas mass is in diffuse ^{12}CO emission that is extended perpendicular to the Galactic plane beyond the ^{12}CO emission where denser gas is detected.

Spectral analyses have found connections between the HI with the bright dense and faint diffuse CO components. The different CO components are highlighted through different analyses, with individual lines-of-sight primarily tracing the bright CO emission, while analyses that study an ensemble of spectra through stacking recover the faint CO emission. Comparing these analyses shows that the properties of the bright and faint CO emission differ. Fukui et al. (2009) find CO line widths in the LMC on 40 pc scales that are $\sim 30\%$ of the HI line widths along the same lines-of-sight. Similar ratios between the CO and HI are found by Wilson et al. (2011) for 12 nearby galaxies on scales from $\sim$ 200–1200 pc, though the HI line widths are estimated at a different resolution from the CO. On similar scales ($\sim$ 200–700 pc), with matched resolution between the HI and CO, Mogotsi et al. (2016) found that the CO line widths ($\sigma_{\rm CO} = 7.3 \pm 1.7$ km s^{-1}) are consistently narrower than the HI ($\sigma_{\rm HI} = 11.7 \pm 2.3$ km s^{-1}) for a number of nearby galaxies. The average ratio of ~ 0.6 between the line widths is much larger than the ratio from Fukui et al. (2009) on smaller scales (~ 0.3).

Stacking analyses consistently have broader CO line widths than those from individual spectra. Combes & Becquaert (1997) found comparable HI and CO line widths in two nearby face-on galaxies ($i < 12^{\circ}$). They suggested that the HI and CO emission trace a common, well-mixed kinematic component that differs only in the phase of the gas. Using the same data as Mogotsi et al. (2016), Caldu-Primo et al. (2013) also found similar line widths between the HI and CO ($\sigma_{\rm CO} = 12.0 \pm 3.9$ km s^{-1} and $\sigma_{\rm HI} = 11.9 \pm 3.1$ km s^{-1}) for a number of nearby galaxies. Caldu-Primo et al. (2013) concluded that the wide CO component arises from a faint, large-scale molecular component that is too faint to be detected in individual lines-of-sight. However, a stacked spectrum is broadened due to scatter in the line centre (Koch et al., 2018c), particularly when HI velocities are used to align the CO spectra (Schruba

et al., 2011; Caldu-Primo et al., 2013). Characterizing methodological sources of line broadening is critical for understanding the spectral properties of the diffuse molecular component.

In M33, there are differing results regarding a diffuse molecular component. Wilson & Scoville (1990) inferred the presence of diffuse molecular gas from interferometric data recovering $\sim$ 40% of the flux from single-dish observations. Wilson & Walker (1994) supported this conclusion by demonstrating that the high ^{12}CO to ^{13}CO line ratio does not result from different filling factors between the two lines. Later, Combes et al. (2012) found a non-zero spatial power-spectrum index on kpc scales and suggested that it arises from a large-scale CO component.

Rosolowsky et al. (2003) and Rosolowsky et al. (2007) also found additional CO emission that did not arise from GMCs, similar to Wilson & Scoville (1990). However, Rosolowsky et al. (2003) localized 90% of the diffuse emission to within 100 pc of a GMC and suggested that this diffuse emission is from a population of small, unresolved molecular clouds that are too faint for their interferometric observations to detect.

These previous results in M33 and other nearby galaxies suggest that detailed studies of molecular clouds and their local environments may need to account for the presence of diffuse CO emission or bright HI emission along the line-of-sight that is unrelated to the molecular cloud. In this paper, we characterize the relationship between the spectral properties of HI and CO in M33 on 80 pc scales by stacking spectra and modelling individual lines-of-sight. We then critically compare these two different analyses, constraining how methodological line broadening and unrelated HI or CO emission affects the properties of stacked spectra. M33 is an ideal system for this comparison as we can connect studies of HI and CO performed on larger scales ($>$ 100 pc) to those on small scales ($<$ 50 pc).

M33's flocculent morphology also lies between the nearby galaxies in previous studies, with a sample of more massive spiral galaxies in the lower-resolution studies and the irregular morphology of the LMC observed at higher-

resolution. The HI in M33 is an ideal tracer of the flocculent spiral structure. The bright HI is aligned in filaments, similar to the "high-brightness network" identified in other galaxies (Braun, 1997).

We compare the atomic and molecular ISM using HI observations obtained with the Karl G. Jansky Very Large Array (VLA) by Koch et al. (2018c) and the CO(2-1) data from the IRAM 30-m by Druard et al. (2014), as described in §5.2. The HI data have a beam size of 20″, corresponding to physical scales of ∼80 pc at the distance of M33 (840 kpc; Freedman et al., 2001). Our study builds on work by Fukui et al. (2009) and Druard et al. (2014) by utilizing improved 21-cm HI observations and new techniques for identifying spectral relationships. We focus on comparing M33's HI and CO distributions along the same lines-of-sight, where we explore the difference in velocity where the HI and CO intensity peaks (§5.3.1), how the line widths of stacked line profiles compare to those measured at lower resolutions (§5.3.2), and the distribution of HI and CO line widths from fitting individual spectra (§5.3.3). We then compare the properties from these two analyses and discern where sources of discrepancy arise (§5.3.4). Our results show that M33 does not have a significant diffuse molecular disk. We discuss this result and compare to previous findings in §5.4.

5.2 Observations

5.2.1 HI VLA & GBT

We utilize the HI observations presented in Koch et al. (2018c) and provide a short summary of the observations here. Figure 5.1 shows the HI integrated intensity map. The observations were taken with the VLA using a 13-point mosaic to cover the inner 12 kpc of M33. The data were imaged with CASA 4.4 using natural weighing and deconvolved until the peak residual reached 3.8 mJy beam^{-1} (7.1 K) per channel, which is about 2.5 times the noise level in the data. The resulting data cube has a beam size of 20″ × 18″, a spectral resolution of 0.2 km s^{-1}, and a 1-σ noise level of 2.8 K per channel. This

spectral resolution is a factor of $\sim$ 13 finer than the CO(2-1) data (§5.2.2), leading to significantly less uncertainty in the velocity at peak intensity in the HI compared to the CO(2-1).

We combine the VLA data with GBT observations by Lockman et al. (2012) to include short-spacing information[1]. We feather the data sets together using the UVCOMBINE package[2], which implements the same feathering procedure as CASA. Thus the HI data used in this work provide a full account of the HI emission down to $\sim$80 pc scales.

5.2.2 CO(2-1) IRAM 30-m

We use the CO(2-1) data from the IRAM-30m telescope presented by Druard et al. (2014). Figure 5.1 shows the region covered by these observations, along with the zeroth moment contours. A full description of the data and reduction process can be found in their §2; a brief summary is provided here. Portions of the map were previously presented by Gardan et al. (2007), Gratier et al. (2010), and Gratier et al. (2012). The data have an angular resolution of 12″, corresponding to a physical resolution of $\sim$48 pc, and a spectral resolution of 2.6 km s^{-1}. Because IRAM 30-m is a single dish telescope, the data are sensitive to all spatial scales above the beam size and does not require the feathering step used with the HI (§5.2.1).

The CO(2-1) cube is a combination of many observations that leads to spatial variations in the noise. The rms noise level differs by a factor of a few in the inner $\sim$ 7 kpc of M33's disk (see Figure 6 in Druard et al., 2014). We adopt the same beam efficiency of $0.56/0.92 = 0.61$ from Druard et al. (2014) for converting to the main beam temperature. The average noise per channel is 33.3 mK in units of the main beam temperature. Since we focus only on the line shape properties, we do not require a conversion factor to the H_2 column density in this paper.

The spectral channels are moderately correlated due to the spectral re-

[1]Described in Chapter 3

[2]https://github.com/radio-astro-tools/uvcombine

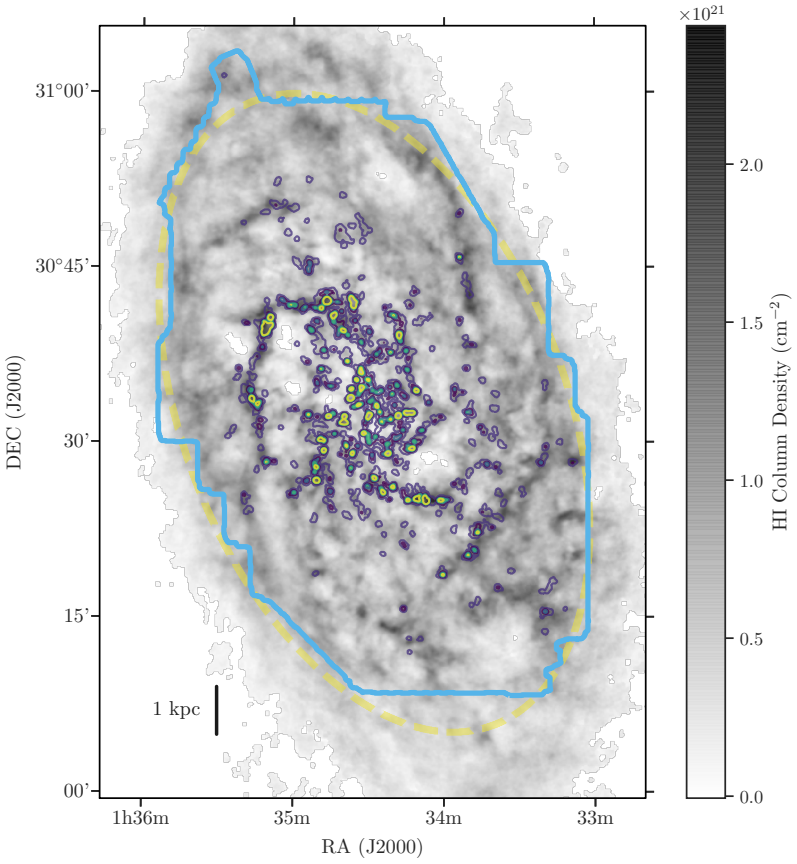

Figure 5.1: HI and CO(2-1) (contours) column density maps at a resolution of $\sim$ 80 pc (20″). The HI column density assumes optically thin emission and is corrected for inclination (Koch et al., 2018c). The CO(2-1) contour levels (from blue to yellow) indicate surface densities of 900, 1400, 1900, and 2400 K km s^{-1}. The light-blue line indicates the extent of the CO(2-1) data, and the dashed yellow line shows the $R_{\rm gal}$ = 7 kpc galactocentric radius. Qualitatively, the CO(2-1) emission tends to be located with bright HI.

sponse function of the instrument. Along with broadening due to finite channel widths, the spectral response function correlates nearby channels and broadens the spectra. This broadening can be accounted for by modelling the known spectral response function and accounting for the channel width (Koch et al., 2018b). Adopting the correlation coefficient of $r = 0.26$ determined by Sun et al. (2018) for these data, and using the empirical relation from Leroy et al. (2016), we approximate the spectral response function as a three-element Hanning-like kernel with a shape of $[k,\ 1-2k,\ k]$, where $k = 0.11$ is the channel coupling. The use of the spectral response function in spectral fitting is described further in §5.3.3.1.

Throughout this paper, we use a spatially-matched version of these CO(2-1) data convolved to have the same beam size as the HI data . The data are spatially-convolved and reprojected to match the HI data, which lowers the average noise per channel to 16.0 mK. The spectral dimension is not changed. We create a signal mask for the data by searching for connected regions in the data with a minimum intensity of 2-σ that contain a peak of at least 4-σ. Each region in the mask must be continuous across three channels and have a spatial size larger than the full-width-half-maximum of the beam.

5.3 HI-CO spectral association

We examine the relation between HI and CO(2-1) spectra using three comparisons: (i) the distribution of peak velocity offsets; (ii) the width and line wing excess, and shape parameters of stacked profiles; and (iii) the line widths of both tracers from a limited Gaussian decomposition of the HI associated with CO(2-1) emission. Unless otherwise specified, the line width refers to the Gaussian standard deviation (σ) and not the full-width-half-maximum (FWHM $= 2\sqrt{2\ \ln\ 2}\sigma$).

5.3.1 Peak velocity relation

We first determine the spectral relation between the HI and CO(2-1) by comparing the velocity of the peak temperatures along the same lines-of-sight. We refer to this velocity as the "peak velocity." Figure 5.2 compares the absolute peak velocity difference between HI and CO(2-1) versus the peak CO temperature. Most lines-of-sight have peak velocities consistent between the HI and CO(2-1). The standard deviation of the velocity difference, after removing severe outliers with differences of > 10 km s^{-1}, is 2.7 km s^{-1}. This is similar to the 2.6 km s^{-1} channel width of the CO(2-1) data, suggesting that the peak velocities are typically consistent within the resolution of the CO(2-1) data. Since the peak velocities are defined at the centre of the velocity channel at peak intensity, recovering a scatter in the peak velocity difference of $\sim \pm 1$ channel is reasonable. The much narrower HI channel width (0.21 km s^{-1}) accounts for significantly less scatter than the CO(2-1) channel width.

Previous HI-CO studies find a similar correlation between the peak velocities of these tracers and have used the HI to infer the peak velocity of CO(2-1) with the goal of detecting faint CO (Schruba et al., 2011; Caldu-Primo et al., 2013). The brightest CO(2-1) peak intensities tend to have smaller velocity differences between the HI and CO, also consistent with the relation found on 40 pc scales in the LMC by Wong et al. (2009).

The distribution in Figure 5.2 has several outliers with velocity differences of > 10 km s^{-1}, far larger than what would be expected from a Gaussian distribution with a width of 2.7 km s^{-1}. These outliers account for 3% of the lines of sight and result from locations where the HI spectrum has multiple components and the CO(2-1) peak is not associated with the brightest HI peak (Gratier et al., 2010). In these cases, the CO(2-1) peaks are well-correlated with the peak of the fainter HI component (§5.3.3). This result is important when stacking spectra (§5.3.2) aligned with respect to the peak HI temperature. When the CO(2-1) peak is not associated with the brightest HI peak, the CO(2-1) stacked profile will be broadened and could potentially be

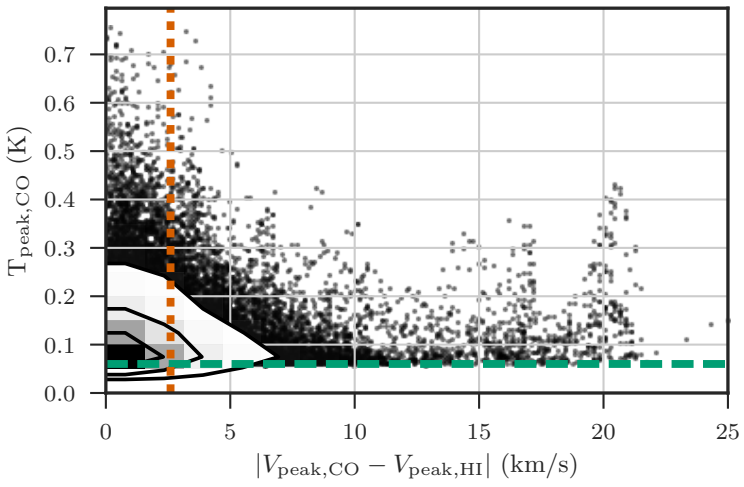

Figure 5.2: Distribution of the peak CO(2-1) brightness (with S/N > 3) versus the absolute difference in the HI and CO(2-1) peak velocities. The shaded region and contours indicate the regions containing the 1, 2, and 3-σ limits of the distribution of points. Individual points show outliers beyond 3σ. The dashed horizontal line is the 3σ rms noise cut-off in the CO(2-1) data imposed to avoid spurious outliers in the velocity difference. The dotted vertical line is the CO channel width of 2.6 km s^{-1}. The CO velocities are preferentially located at or near the HI velocities. However, there remains a number of high S/N outliers with a large velocity difference. These outliers occur when the CO emission is associated with a different HI component than the brightest HI peak.

asymmetric if the CO peaks are preferentially blue- or red-shifted from the HI component. We explore these effects in §5.3.4.

We conclude that the HI peak velocity can nearly always be used to infer the CO(2-1) peak velocity.

5.3.2 Stacking analysis

By stacking a large number of spectra aligned to a common velocity, we can examine a high signal-to-noise (S/N) average spectrum of each tracer. Since the signal will add coherently, while the noise will add incoherently, the stacked profiles are ideal for identifying faint emission that is otherwise not detectable in individual lines-of-sight (Schruba et al., 2011). These high S/N spectra can

be used to compare the line profile properties of the HI and CO(2-1).

We examine stacked profiles of HI and CO(2-1) aligned with respect to the HI peak velocity since the HI is detected towards nearly every line-of-sight, the HI peak velocity describes the peak CO(2-1) velocity well (§5.3.1), and the velocity resolution of the HI data is much higher than the CO(2-1) data. We align the spectra by shifting them; we Fourier transform the data, apply a phase shift, and transform back. This procedure preserves the signal shape and noise properties when shifting by a fraction of the channel size[3]. The channel size is a particular issue for the CO(2-1) data, since the channels have a width of 2.6 km s^{-1} and the signal in some spectra only spans $\sim$5 channels.

Figure 5.3 shows the stacked profiles, where spectra within a radius of 7 kpc are included. The HI stacked profile has a kurtosis excess relative to a Gaussian, with enhanced tails and a steep peak. These properties of HI stacked profiles are extensively discussed in Koch et al. (2018c). The CO(2-1) stacked profile has a qualitatively similar shape but is narrower than the HI profile and has a smaller line wing excess.

As in Koch et al. (2018c), we model the profiles based on the half-width-half-maximum (HWHM = FWHM/2) approach from Stilp et al. (2013a). The model and parameter definitions are fully described in Koch et al. (2018c); we provide a brief overview here. The HWHM model assumes that the central peak of the profiles can be described by a Gaussian profile whose FWHM matches the profile's FWHM, which is well-constrained in the limit of high S/N. This model sets the Gaussian standard deviation ($\sigma_{\rm HWHM} = {\rm HWHM}/\sqrt{2\ln 2}$) and central velocity ($v_{\rm peak}$) of this Gaussian, which we refer to as $G(v)$.

The following parameters describe how the observed profile, $S(v)$, compares to the Gaussian model.

The line wing excess expresses the fractional excess relative to the Gaussian

[3]Implemented in spectral-cube (https://spectral-cube.readthedocs.io)

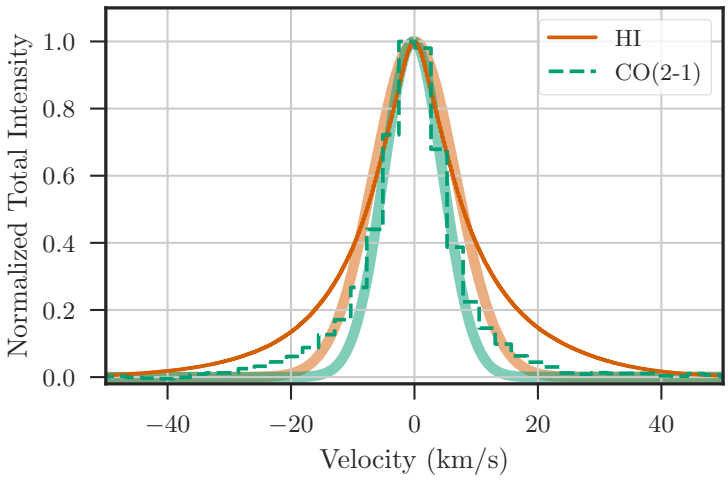

Figure 5.3: The HI (orange solid) and CO(2-1) (green dashed) stacked profiles shifted with respect to the HI peak velocity. The thick faint lines show the Gaussian model for each tracer. The HI stacked profile is wider and has a larger line wing excess than the CO(2-1) stacked profile.

outside of the FWHM:

$$f_{\rm wings} = \frac{\sum\limits_{|v|>\rm HWHM} [S(v) - G(v)]}{\sum\limits_{v} S(v)} \ . \tag{5.1}$$

This excess in the line wings can also be used to find the "width" of the wings using a form equivalent to the second moment of a Gaussian:

$$\sigma^2_{\rm wings} = \frac{\sum\limits_{|v|>\rm HWHM} [S(v) - G(v)]\, v^2}{\sum\limits_{|v|>\rm HWHM} [S(v) - G(v)]} \ . \tag{5.2}$$

Since the line wing excess will not be close to Gaussian in shape, $\sigma_{\rm wings}$ does not have a clear connection to a Gaussian width.

The asymmetry of a stacked profile is defined as the difference in total flux at velocities greater than and less than $v_{\rm peak}$, normalized by the total flux:

$$a = \frac{\sum\limits_{v>v_{\rm peak}} S(v) - \sum\limits_{v<v_{\rm peak}} S(v)}{\sum S(v)} \ . \tag{5.3}$$

This makes a analogous to the skewness of the profile.

The shape of the peak is described by κ, defined as the fractional difference between the central peak and the Gaussian model within the FWHM:

$$\kappa = \frac{\sum_{|v|<\mathrm{HWHM}} [S(v) - G(v)]}{\sum_{|v|<\mathrm{HWHM}} G(v)} \ . \tag{5.4}$$

This describes the kurtosis of the profile peak, where $\kappa > 0$ is a profile more peaked than a Gaussian and $\kappa < 0$ is flatter than a Gaussian. We note that the kurtosis typically describes the line wing structure, however since these regions are excluded in our definition, κ describes the shape of the peak.

Since we adopt a semi-parametric model for the stacked profiles, deriving parameter uncertainties also requires a non-parametric approach. We use a bootstrap approach presented in Koch et al. (2018c) to account for the two significant sources of uncertainty:

1. *Uncertainty from the data* – These uncertainties come directly from the noise in the data. In each channel, the uncertainty is $\sigma_{\mathrm{RMS}}/\sqrt{N_{\mathrm{spec}}}$, where N_{spec} is the number of spectra included in the stacked spectrum. We account for this uncertainty by resampling the values in the stacked spectrum by drawing from a normal distribution centered on the original value with a standard deviation equal to the noise. We then calculate the parameters from the resampled stacked spectrum at each iteration in the bootstrap.

2. *Uncertainty due to finite channel width* – The finite channel width introduces uncertainty in the location of the peak velocity when not explicitly modelled for with an analytic model (Koch et al., 2018b). Since the HWHM model is a semi-parametric model that does not account for finite channel width, we need to adopt an uncertainty for the peak velocity and the inferred line width. We use the HWHM model on very high S/N stacked spectra and assume that the true peak velocity is known to be within the channel of peak intensity. To create an equivalent Gaussian

standard deviation[4] for the uncertainty, we scale the rectangular area in the peak channel to the fraction of the area under a Gaussian within ± 1–σ, which gives $\sigma_{v_{peak}} = 0.34\Delta v$[5]. The HWHM model estimates the width σ_{HWHM} based on v_{peak} and thus we adopt the same uncertainty for both parameters. To estimate the uncertainties on the other parameters in the HWHM model, we sample new values of v_{peak} and σ_{HWHM} from normal distributions with standard deviations of $0.34\Delta v$ in each bootstrap iteration. These two parameters set the Gaussian shape used to derived the parameters in Equations 5.1–5.4.

Based on the bootstrap sampling used in these two steps, we estimate the uncertainties on the remaining HWHM model parameters. Since the stacked profiles have high S/N, the parameter uncertainties are dominated by the uncertainty in v_{peak} and σ_{HWHM}, and since the CO(2-1) channel size is much larger than the HI (2.6 versus 0.2 km s^{-1}), the CO(2-1) uncertainties are much larger.

Table 5.1 provides the parameter values and uncertainties from the HWHM model for the stacked profiles in Figure 5.3. The CO(2-1) line width is 4.6 ± 0.9 km s^{-1}, which is 70% of the HI width of 6.6 ± 0.1 km s^{-1}.

Using the same CO(2-1) data at the original $12''$ resolution, Druard et al. (2014) create CO(2-1) stacked profiles and fit a single Gaussian component to the profile. They find a line width of $\sigma = 5.3 \pm 0.2$ km s^{-1}, which is 0.7 km s^{-1} larger than our measurement using the HWHM method. This discrepancy results from the different modelling approaches used; fitting our CO(2-1) stacked profile with a single Gaussian component gives a line width of 5.4 ± 0.9 km s^{-1}, consistent with Druard et al. (2014), because the fit is influenced by the line wings.

The profiles are consistent with the same v_{peak}, as is expected based on the strong agreement between the peak velocities (§5.3.1). The scatter in the peak

[4] In Koch et al. (2018c), we used Δv as the uncertainty. Since the HI channels are much narrower than σ_{HWHM}, this change has little effect on the HI uncertainties.

[5] $A\,(2\,\sigma_{v_{peak}}) = A\,(0.68\Delta v)$, where A is the value in the stacked spectrum and 0.68 is the fraction of the area when integrating a Gaussian from -1σ to $+1\sigma$.

velocity difference will primarily *broaden* the spectrum, rather than create an offset in $v_{\rm peak}$. We test the importance of this source of broadening in §5.3.4, but note that the non-Gaussian shape of the stacked profiles makes correcting for this broadening non-trivial. Because of this, we do not apply a correction factor to $\sigma_{\rm HWHM}$ to account for the spectral response function and channel width since the stacked profiles have a non-Gaussian shape.

The HI profile is more non-Gaussian in shape than the CO(2-1). The HI profile has a larger line wing excess and a non-Gaussian peak ($\kappa < 0$), consistent with the stacked profiles in Koch et al. (2018c).

The large uncertainties on the CO(2-1) shape parameters make most not significant at the 1-σ level, or are consistent within 1-σ of the HI shape parameters. Within the uncertainty, the CO(2-1) stacked profiles are symmetric about the peak ($a = 0$) and have a Gaussian-shaped profile within the HWHM ($\kappa = 0$). The only significant CO(2-1) shape parameter is the line wing excess $f_{\rm wings}$, which is non-zero at the 2-σ level and consistent with the HI $f_{\rm wings}$ within 1-σ. However, there are additional systematics that may contribute to the CO(2-1) $f_{\rm wings}$, including broadening from the distribution of the peak HI and CO velocities (§5.3.1) and the IRAM 30-m error beam pickup (Druard et al., 2014). We discuss the former contribution in more detail in §5.3.4. Druard et al. (2014) estimate that the error-beam pickup contributes 2.5×10^6 $M_\odot$, or 1.1×10^4 K km s^{-1}, using their conversion factor $X_{\rm CO} = 4 \times 10^{20}$ cm^{-2}/(K km s^{-1}) and a brightness temperature ratio of 0.8 between the $J = 2$-1 and $J = 1$-0 CO transitions. The error beam flux may then contribute up to 45% of the 2.4×10^4 K km s^{-1} line wing excess. We further assess whether the error beam flux contributes to the line wings in §5.3.2.1.

Assuming that the error beam does contribute 45% of the CO line wing excess, M33 appears to exhibit weaker line wings compared to those measured in M31 by Caldú-Primo & Schruba (2016a). They characterized the CO stacked profiles with a Gaussian model and found that single-dish CO observations in M31 are best fit by two Gaussian components. Their wide Gaussian com-

Table 5.1: HWHM model parameters for the HI and CO(2-1) stacked profiles for $R_{\rm gal} < 7$ kpc at different spatial resolutions. The uncertainties are propagated assuming an uncertainty of half the channel width and the uncertainty of each point in the spectrum is the standard deviation of values within that channel scaled by the square-root of the number of beams.

	HI	CO(2-1)
80 pc (20″) resolution		
$\sigma_{\rm HWHM}$ (km s^{-1})	6.6 ± 0.1	4.6 ± 0.9
$v_{\rm peak}$ (km s^{-1})	0.0 ± 0.1	-0.2 ± 0.9
$f_{\rm wings}$	$0.25^{+0.01}_{-0.01}$	$0.21^{+0.12}_{-0.10}$
$\sigma_{\rm wings}$ (km s^{-1})	$24.0^{+0.3}_{-0.4}$	18^{+5}_{-3}
a	$0.021^{+0.014}_{-0.005}$	$-0.05^{+0.09}_{-0.14}$
κ	$-0.059^{+0.004}_{-0.003}$	$0.02^{+0.06}_{-0.09}$
160 pc (38″) resolution		
$\sigma_{\rm HWHM}$ (km s^{-1})	8.0 ± 0.1	5.9 ± 0.9
$v_{\rm peak}$ (km s^{-1})	-0.1 ± 0.1	-0.4 ± 0.9
$f_{\rm wings}$	$0.19^{+0.01}_{-0.01}$	$0.14^{+0.10}_{-0.09}$
$\sigma_{\rm wings}$ (km s^{-1})	$29.3^{+0.3}_{-0.4}$	20^{+8}_{-4}
a	$0.012^{+0.007}_{-0.012}$	$-0.07^{+0.20}_{-0.10}$
κ	$-0.022^{+0.004}_{-0.003}$	$0.00^{+0.05}_{-0.06}$
380 pc (95″) resolution		
$\sigma_{\rm HWHM}$ (km s^{-1})	8.9 ± 0.1	7.2 ± 0.9
$v_{\rm peak}$ (km s^{-1})	0.0 ± 0.1	-0.2 ± 0.9
$f_{\rm wings}$	$0.19^{+0.01}_{-0.01}$	$0.15^{+0.09}_{-0.07}$
$\sigma_{\rm wings}$ (km s^{-1})	$32.1^{+0.3}_{-0.4}$	22^{+4}_{-3}
a	$0.033^{+0.010}_{-0.015}$	$-0.07^{+0.20}_{-0.10}$
κ	$-0.035^{+0.004}_{-0.035}$	$-0.01^{+0.05}_{-0.05}$

ponent would be related to the line wing excess, i.e., a large $f_{\rm wings}$, in our formalism[6]. For the sake of comparison with Caldú-Primo & Schruba (2016a), we fit a two-Gaussian component to the CO(2-1) stacked profile and find line widths of 3.8 ± 0.9 km s^{-1} and 10.9 ± 0.9 km s^{-1} for the narrow and wide Gaussian components, respectively. The narrow line width is similar to the 3.2 ± 0.2 km s^{-1} found by Caldú-Primo & Schruba (2016a), however their wide component is much narrower, with a width of 6.1 ± 0.6 km s^{-1}.

5.3.2.1 Radial stacked profiles

We explore trends with galactocentric radius by creating stacked profiles within radial bins of 500 pc widths out to a maximal radius of 7 kpc, matching the coverage of the CO(2-1) map. We use a position angle of 201.1° and inclination angle of 55.1° for M33's orientation, based on the HI kinematics from Koch et al. (2018c). The radial stacking uses the same procedure for the HI profiles as in Koch et al. (2018c), but with 500 pc radial bins instead of 100 pc due to the smaller filling fraction of CO(2-1) detections relative to the HI. The stacked profiles are modeled with the same HWHM model described above.

Figure 5.4 shows the line widths ($\sigma_{\rm HWHM}$) of the stacked profiles over the galactocentric radial bins (values provided in Table 5.4). We quantify the relation between galactocentric radius and the line widths by fitting a straight line. We exclude the innermost bins (< 1 kpc) where beam smearing has a small contribution (§5.6.1). To account for the line width uncertainties, we resample the line widths in each bin from a Gaussian distribution with a standard deviation set by the uncertainty (Table 5.4) in 1000 iterations. We then estimate the slope and its uncertainty using the 15$^{\rm th}$, 50$^{\rm th}$, and 85$^{\rm th}$ percentiles from the distribution of 1000 fits. We find that the HI line widths decrease with galactocentric radius (-0.14 ± 0.01 km s^{-1} kpc^{-1}), consistent with the stacking analysis in 100 pc bins by Koch et al. (2018c). The decrease in the CO line widths with radius is insignificant at the 1-σ level ($-0.16 \pm$

[6]A relation between $f_{\rm wings}$ and the wide Gaussian component would be a function of the amplitudes and widths of both Gaussian components, and the $\sigma_{\rm HWHM}$ used here.

0.16 km s^{-1} kpc^{-1}). We note, however, that cloud decomposition studies of the CO(2-1) find a shallow line width decrease with galactocentric radius (Gratier et al., 2012; Braine et al., 2018).

Many nearby galaxies have a similar shallow radial decline in the HI and CO line widths, outside of the galaxy centres (Caldu-Primo et al., 2013; Mogotsi et al., 2016). Enhanced line widths are observed in galactic centres that result from a significant increase in the molecular gas surface density or the presence of a bar (e.g., Sun et al., 2018), or due to beam smearing where the gradient of the rotation velocity is significant on the scale of the beam. M33 is a lower mass spiral galaxy and lacks a strong bar, making it likely that the moderate line width increase within $R_{\rm gal} < 0.5$ kpc is due to beam smearing (§5.6.1).

A clear difference between our results and those by other studies is that the ratio between the CO and HI line width is ∼0.7, differing from the ratio of ∼ 1.0 typically measured in other systems on 0.2-0.7 kpc (Combes & Becquaert, 1997; Caldu-Primo et al., 2013). Since the observation of comparable CO and HI line widths is used as an indicator of a thick molecular gas disk, we discuss this topic in detail in §5.4.

Most of the shape parameters from the HWHM model do not show significant trends with galactocentric radius for the CO(2-1) spectra and are insignificant at the 1-σ level. The line wing excess ($f_{\rm wings}$) is significant for radial bins within $R_{\rm gal} < 5$ kpc. At larger radii, the line wings become less prominent, though the CO detection fraction also sharply decreases at these radii and systematic effects—for example, from baseline fitting—strongly affect the stacked profile shapes beyond the HWHM. This leads to the negative line wing excess at $6 < R_{\rm gal} < 6.5$ kpc and the large excess in the $6.5 < R_{\rm gal} < 7$ kpc bin. The stacked profiles in Figure 12 of Druard et al. (2014) also clearly suffer from these effects.

In the previous section, we estimated that error beam pick-up can account for up to 45% of $f_{\rm wings}$ in the stacked profiles for $R_{\rm gal} < 7$ kpc. Due to galactic rotation, the error-beam pick up will be asymmetric between the halves of the galaxy. To test for this asymmetry, we also stack spectra in galactocentric

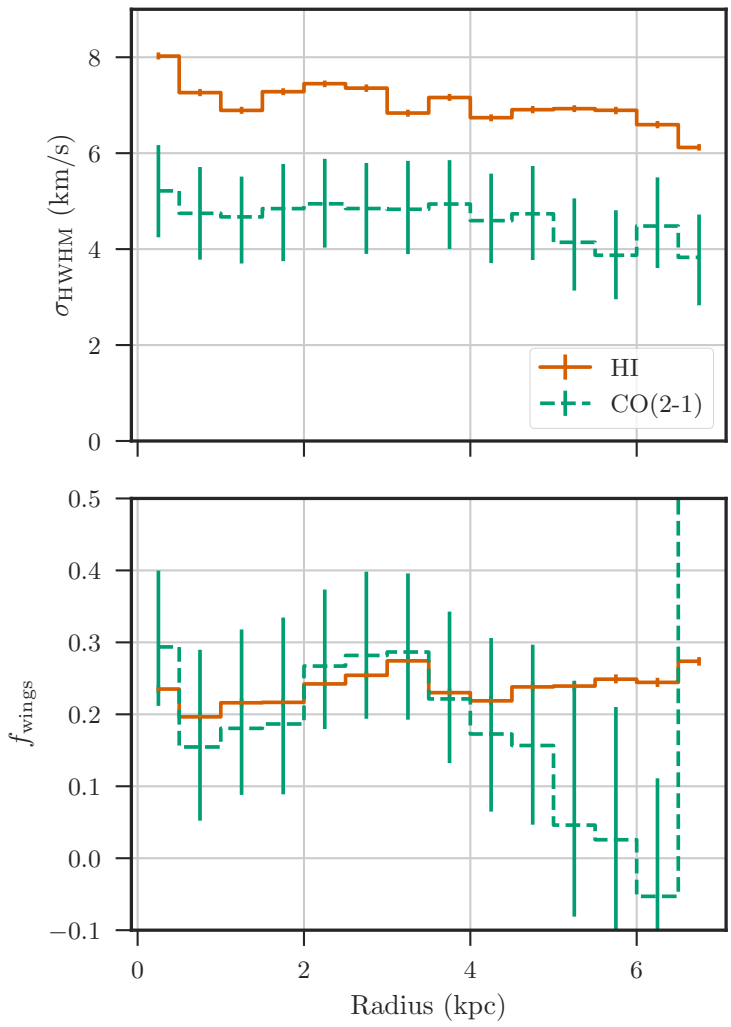

Figure 5.4: Stacked profile line widths ($\sigma_{\rm HWHM}$; top) and the fractional line wing excess ($f_{\rm wings}$; bottom) measured in 500 pc radial bins. The widths are based on the HWHM approach from Stilp et al. (2013a), and the errors are from a bootstrap approach described in the text. The widths from both tracers show a shallow radial decline and have a consistent line width ratio of ∼0.7.

radial bins separated into the northern and southern halves. Though with significant uncertainties, the asymmetry of the northern half stacked profiles is consistently more negative than the southern half stacked profiles, which have asymmetries that are either positive or near zero. This discrepancy between the halves shows that error beam pick-up accounts for some of the line wing excess.

The other model parameters are consistent between the northern and southern halves of M33.

5.3.2.2 Stacked profiles at coarser resolution

We further investigate how stacked profile properties change with spatial resolution by repeating our analysis on data smoothed to a resolution of 160 pc ($38''$) and 380 pc ($95''$). This allows for a more direct comparison to studies of stacked profiles on larger physical scales (Caldu-Primo et al., 2013). At each resolution, we recompute the HI peak velocities and create stacked HI and CO profiles at that resolution. Table 5.1 gives the HWHM model parameters for these lower-resolution stacked profiles.

The line widths of both HI and CO(2-1) increase at coarser spatial resolution. Based on stacking over the entire galaxy within $R_{\rm gal} = 7$ kpc, we find $\sigma_{\rm HI} = 8.0 \pm 0.1$ km s^{-1} and $\sigma_{\rm CO} = 5.9 \pm 0.9$ km s^{-1} at a resolution of 160 pc, and $\sigma_{\rm HI} = 8.9 \pm 0.1$ km s^{-1} and $\sigma_{\rm CO} = 7.2 \pm 0.9$ km s^{-1} at a resolution of 380 pc. The CO(2-1) line widths have a larger relative increase than the HI ones, which results in increased ratios of $\sigma_{\rm CO}/\sigma_{\rm HI} = 0.70 \pm 0.18$, 0.74 ± 0.15 and 0.81 ± 0.12 at scales of 80, 160 and 380 pc data, respectively. However, the large uncertainties on the line width ratios makes this increase insignificant at the 1-σ level. Using CO observations with higher spectral resolution will decrease these uncertainties and can determine whether this trend is significant.

There are two sources of line broadening that affect $\sigma_{\rm HWHM}$ as the resolution becomes coarser: (i) the dispersion between the HI and CO(2-1) peak velocities, and (ii) beam smearing. Line broadening from the former source is due to aligning the CO(2-1) spectra by the HI peak velocity. At a resolution

of 80 pc, we estimate the standard deviation of the peak velocity difference to be 2.7 km s^{-1}, as described in §5.3.1. Using the same procedure at the coarser resolutions, we find 1-σ standard deviations of 3.1 and 3.3 km s^{-1} at a resolution of 160 and 380 pc, respectively. The increase in the peak velocity difference moderately increases with resolution, but cannot account for the increase in the line widths at coarser resolution.

To address increased line broadening from beam smearing at coarser resolution, we repeat the line stacking in 500 pc radial bins at each resolution. Figure 5.5 shows the stacked line widths ($\sigma_{\rm HWHM}$) at the three spatial resolutions. As the resolution becomes coarser, there is a steeper radial gradient in the line widths of both HI and CO, particularly for $R_{\rm gal} < 1$ kpc. This radial trend qualitatively matches our estimate of beam smearing from §5.6.1. We determine how much of the line width increase with resolution is due to beam smearing with the area-weighted line broadening estimates calculated in §5.6.1. For resolutions of 80 and 160 pc, the broadening from beam smearing is similar, with estimates of $2.0^{+2.1}_{-1.8}$ km s^{-1} and $1.5^{+1.7}_{-0.8}$ km s^{-1}, respectively. The similar levels of beam smearing at 80 and 160 pc imply that the increase in the line width with resolution is not due to beam smearing.

At a resolution of 380 pc, beam smearing contributes significantly to the stacked line widths. The area-weighted line broadening from beam smearing is $2.8^{+1.0}_{-1.0}$ km s^{-1}. Treating the stacked profiles as Gaussian within the HWHM, we assume the line broadening can be subtracted in quadrature from the line width. Applying this correction gives line widths of 8.4 ± 0.9 km s^{-1} for HI and 6.6 ± 3.6 km s^{-1} for CO(2-1). The CO(2-1) line width does not constrain whether the increased line widths are from beam smearing due to the uncertainty from the channel width. However, the 0.9 km s^{-1} increase in the HI line width between the 160 and 380 pc data is much larger than the uncertainty and can entirely be explained by beam smearing.

On scales of 380 pc and larger, beam smearing becomes the dominant contribution to the line width. With increasing scale, the stacked HI and CO profiles will approach a common width since the stacking is performed with

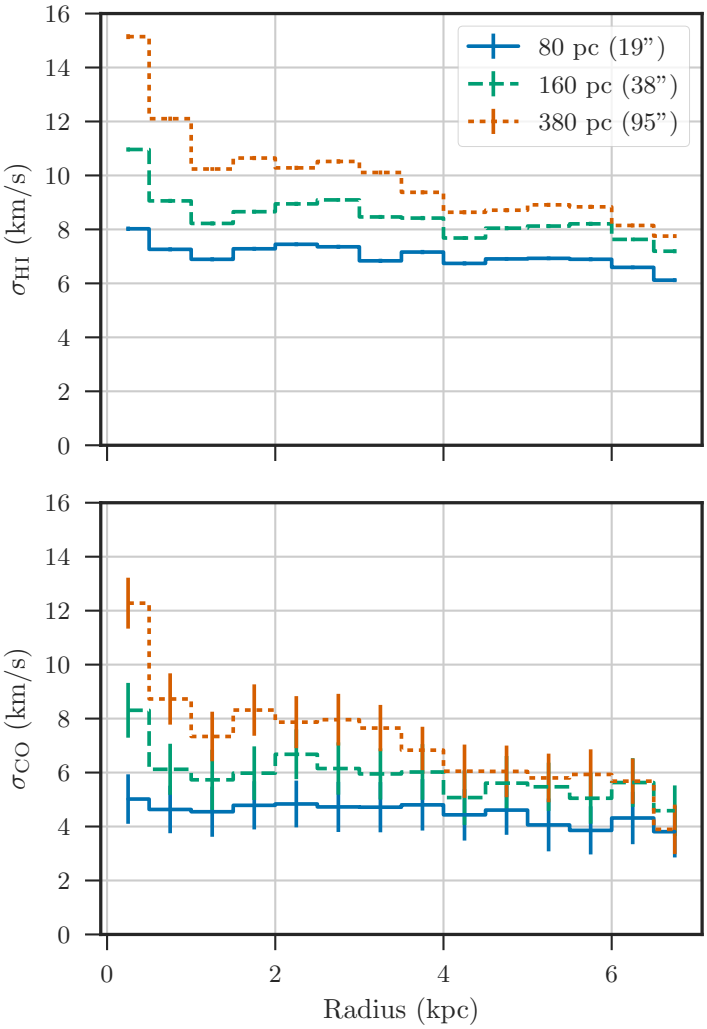

Figure 5.5: Stacked profile line widths ($\sigma_{\rm HWHM}$) measured in 500 pc radial bins at three different spatial resolutions for HI (top) and CO(2-1) (bottom). Line widths measured at coarser resolution have a steeper radial gradient due to beam smearing (§5.6.1).

respect to the HI and will be set by the rotation curve. On smaller scales, systematics in the stacking procedure and beam smearing cannot account for the measured increase in the line widths.

We find that the CO line widths remain smaller than the HI on scales up to 380 pc, which is within the range where the stacking study by Caldu-Primo et al. (2013) find equivalent CO and HI line widths. We discuss this difference in the ratio of the line widths in M33 to other nearby galaxies in §5.4.

5.3.3 HI-CO line of sight comparison

Stacked profiles provide a high S/N spectrum whose properties trace the average of the ensemble of stacked spectra. However, stacking removes information about the spatial variation of individual spectra and distributions of their properties. By fitting individual lines-of-sight, the distributions of line shape parameters that lead to the shape of the stacked spectrum can be recovered.

Describing the HI line profiles with an analytic model is difficult. As we demonstrate in Koch et al. (2018c), the typical HI line profile in M33 is non-Gaussian due to multiple Gaussian components and asymmetric line wings. The weak relations between the CO and HI integrated intensities found by Wong et al. (2009) in the LMC suggest that much of the HI emission may be unrelated to the CO. With this in mind, and since the CO spectra can typically be modelled with a single Gaussian at this resolution and sensitivity, we use the location of CO emission as a guide to decompose the HI spectra and model the component most likely related to the CO. This is an approximate method; a proper treatment of modeling individual HI spectra requires a robust Gaussian decomposition (Lindner et al., 2015; Henshaw et al., 2016), which is beyond the scope of this paper.

5.3.3.1 Fitting individual spectra

We relate the HI and CO by using a limited decomposition of the HI based on the spatial and spectral location of CO(2-1) emission. Towards lines-of-sight with CO detections, we determine the parameters of the single Gaussian

component that is most closely related to the CO emission.

There are 19,796 spectra where CO(2-1) emission is detected above 3-σ in three consecutive channels and is within $R_{\rm gal} < 7$ kpc. We model these spectra with the following steps:

1. We fit a single Gaussian to the CO(2-1) spectrum, accounting for broadening due to the channel width and the spectral response function using forward-modelling, which is described in §5.6.2. Forward-modelling accounts for line broadening due to the channel width and the spectral response function of the CO(2-1) data (§5.2.2).

2. The peak velocity of the CO(2-1) fit defines the centre of a search window to find the nearest HI peak. The window is set to a width of three times the FWHM of the CO(2-1) fitted width.

3. Since narrow extragalactic HI spectra have widths > 2 km s^{-1} (Warren et al., 2012), we first smooth the HI spectrum with a 2 km s^{-1} box-car kernel. Within the search window, we identify the closest HI peak to the CO(2-1) peak velocity[7].

4. Using the peak temperature of the identified HI peak, we search for the HWHM points around the peak to define the HI fit region.

5. We fit a Gaussian to the un-smoothed HI spectrum within the HWHM points of the identified peak.

This approach assumes that HI spectra are comprised of a small number of Gaussian components with well-defined peaks. Since these restrictions are severely limiting, we define a number of checks to remove spectra that do not satisfy the criteria. A line-of-sight is included in the sample if it meets the following criteria:

[7]We did not find evidence for flattened HI spectra in Koch et al. (2018c) on 80 pc scales due to self-absorption so we do not search for absorption features aligned with the CO.

1. The HI peak is within the CO(2-1) search window, defined above. Based on visual inspection, HI spectra that contain velocity-blended Gaussian components near the CO(2-1) peak will not satisfy this criterion, and our naive treatment will fail to identify a single HI component.

2. The CO line width is larger than one channel width (2.6 km s^{-1}).

3. Ten faint ($T_{\rm HI,peak} < 15$ K) HI spectra have a fitted peak associated with noise. The resulting HI fitted profiles have narrow widths ($\sigma < 3$ km s^{-1}) and are removed from the sample.

4. The fitted HI peak velocity falls within the HWHM region or its Gaussian width is smaller than 12 km s^{-1}. This step removes HI spectra with velocity-blended Gaussian components that significantly widen the Gaussian width and are not treated correctly with this method. We set the width threshold based on visual inspection.

5. The CO(2-1) line width is less than 8 km s^{-1}. A small fraction of CO(2-1) spectra have multiple velocity components, and their fits to a single Gaussian all yield widths larger than 8 km s^{-1}.

These restrictions yield a clean sample of 15,153 spectra that we analyze here, 76% of the eligible spectra. Table 5.2 gives the properties of the line width distributions. The uncertainties from each fit are from the covariance matrix of the least-squares fit. We further validate the use of single Gaussian fits in §5.6.3 and show examples of the fitting procedure.

5.3.3.2 Relations between fitted parameters

We now examine the distributions of fitted HI and CO line parameters to identify which parameters are related.

The fitted peak velocities are strongly correlated (Kendall-Tau correlation coefficient of 0.97), consistent with the peak velocity over the LOS shown in Figure 5.2. The agreement is improved, however, since the outliers (> 10 km s^{-1}) in Figure 5.2 are removed by only fitting the closest HI component rather

Table 5.2: Mean line widths from the line-of-sight spectral fitting at different resolutions. The uncertainties correspond to the 15^{th} and 85^{th} percentiles, respectively.

Resolution (pc)	σ (km s^{-1})		Fitted $\sigma_{\text{CO}}/\sigma_{\text{HI}}$
	HI	CO(2-1)	
80 (20″)	$7.4^{+1.7}_{-1.3}$	$4.3^{+1.5}_{-1.0}$	0.56 ± 0.01
160 (38″)	$8.4^{+1.8}_{-1.2}$	$5.0^{+1.4}_{-1.1}$	0.57 ± 0.01
380 (95″)	$11.0^{+2.7}_{-2.0}$	$7.3^{+2.4}_{-1.6}$	0.63 ± 0.01

than the brightest one. Comparing the fitted peak velocities, the largest velocity difference between the HI and CO is 9.5 km s^{-1}. The standard deviation between the fitted peak velocities is 2.0 km s^{-1}, narrower than the 2.7 km s^{-1} standard deviation from the line-of-sight peak velocity distribution from §5.3.1.

We also find that the peak HI temperatures where CO(2-1) is detected increase from $R_{\text{gal}} < 2$ kpc to $R_{\text{gal}} > 2$ kpc. The brightest peak HI temperatures (> 80 K) are primarily found in the spiral arms, or spiral arm fragments, at $R_{\text{gal}} > 2$ kpc. The lack of spiral structure in the inner 2 kpc may lead to the lack of peak temperatures over $\sim$80 K. These results imply that the HI and CO peak temperatures are not strongly correlated, consistent with the small correlation coefficient of 0.1 we find using the Kendall-Tau test. The weak correlation in peak temperatures is consistent with Wong et al. (2009), who find that HI peak temperature is poorly correlated to CO detections in the LMC.

The peak CO temperature has a negative correlation with CO line width, as would be expected for a Gaussian profile with a fixed integrated intensity. However, other studies using these CO(2-1) data do not recover this negative correlation. Gratier et al. (2012) find a positive, though weak, correlation

between the peak CO temperature and the line width from a cloud decomposition analysis. A similar correlation is found by Sun et al. (2018), who estimate the line widths with the equivalent Gaussian width determined from the peak temperature and integrated intensity of a spectrum (Heyer et al., 2001; Leroy et al., 2016). The discrepancy between our results and these other works is due to requiring three consecutive channels 3-σ above the rms noise. This biases our LOS sample, leading to incomplete distributions in the peak temperature and integrated intensity.

Figure 5.6 shows that there is a clear relation between $\sigma_{\rm CO}$ and $\sigma_{\rm HI}$. Though there is significant dispersion in the relation, there is an increasing trend between the line widths of CO(2-1) and HI. We find median line widths of 4.3 and 7.4 km s^{-1} for CO and HI, respectively, on 80 pc scales. The CO line width distribution is near-Gaussian with a skew to large line widths, with 15$^{\rm th}$ and 85$^{\rm th}$ percentiles of 3.3 and 5.8 km s^{-1}. The HI distribution is more skewed to larger line widths compared to the CO distribution, and has 15$^{\rm th}$ and 85$^{\rm th}$ percentile values of 6.2 and 9.2 km s^{-1}. The variations in $\sigma_{\rm HI}$ and $\sigma_{\rm CO}$ are larger than the typical uncertainties of 0.2 and 0.6 km s^{-1}, respectively.

We highlight the importance of restricting where the HI is fit to in §5.6.3.1, where we show that fitting the whole HI spectrum leads to significant scatter in the HI line widths that severely affects the relationships between the line widths we find here.

Caldú-Primo & Schruba (2016a) perform a similar restricted analysis of single Gaussian fitting to CO spectra of M31 at a deprojected resolution of $80\,{\rm pc} \times 380\,{\rm pc}$. From their combined interferometric and single-dish data, they find typical CO line widths of 4.3 ± 1.3 km s^{-1}, consistent with the range we find.

We characterize the relationship between line widths by fitting for the line width ratio using a Bayesian error-in-variables approach (see Section 8 of Hogg et al., 2010). The goal of this approach is to fully reproduce the data in the model by incorporating (i) the line width uncertainties into the model and (ii) a parameter for additional scatter perpendicular to the line in excess of

the uncertainties[8]. We find that $\sigma_{\rm CO} = (0.56 \pm 0.01)\ \sigma_{\rm HI}$, which is shown in Figure 5.6 as the the green dashed line. The scatter parameter in the model is fit to be 0.52 ± 0.02 km s^{-1}, demonstrating that the scatter in the line width distributions exceeds the uncertainties. This additional scatter represents real variations in the line width distributions.

We next examine whether changes in the line width with galactocentric radius can lead to additional scatter in the line width distributions. Similar to the stacking analysis, we fit the line width relation within 500 pc galactocentric bins out to a radius of 7 kpc and find no variations in the average widths of the component with galactocentric radius, consistent with the shallow radial decrease from the stacked profile analysis (§5.3.2). By examining these possible sources of scatter in the line width distributions, we find that none of the sources can fully account for the scatter and that there must be additional variations not accounted for by the relationships of the fitted parameters. We discuss the source of the scatter further in §5.3.3.4.

5.3.3.3 The line width ratio at coarser spatial resolution

Similar to the stacked profile analysis (§5.3.2.2), we repeat the line-of-sight analysis when the data are smoothed to 160 and 380 pc. The same fitting procedure is applied, with similar rejection criteria for poor fits. However, we found that the line widths of valid fits to the data when smoothed to 380 pc can exceed the imposed cut-off values of 12 and 8 km s^{-1} for HI and CO(2-1) due to additional beam smearing on these scales (§5.6.1). Based on visual inspection, we increase these cut-off values to 17 km s^{-1} and 12 km s^{-1}.

Table 5.2 shows the line width distributions at these scales. As we found with the stacked profile widths, the line widths increase at coarser resolution. The line width remains strongly correlated on these scales.

We fit for the line width ratios of the low-resolution samples and find values of 0.57 ± 0.01 and 0.63 ± 0.01 for the 160 and 380 pc resolutions, respectively. The fitted ratios indicate that the line width ratio is relatively constant

[8]This parameter tends to 0 when no additional scatter is required to model the data.

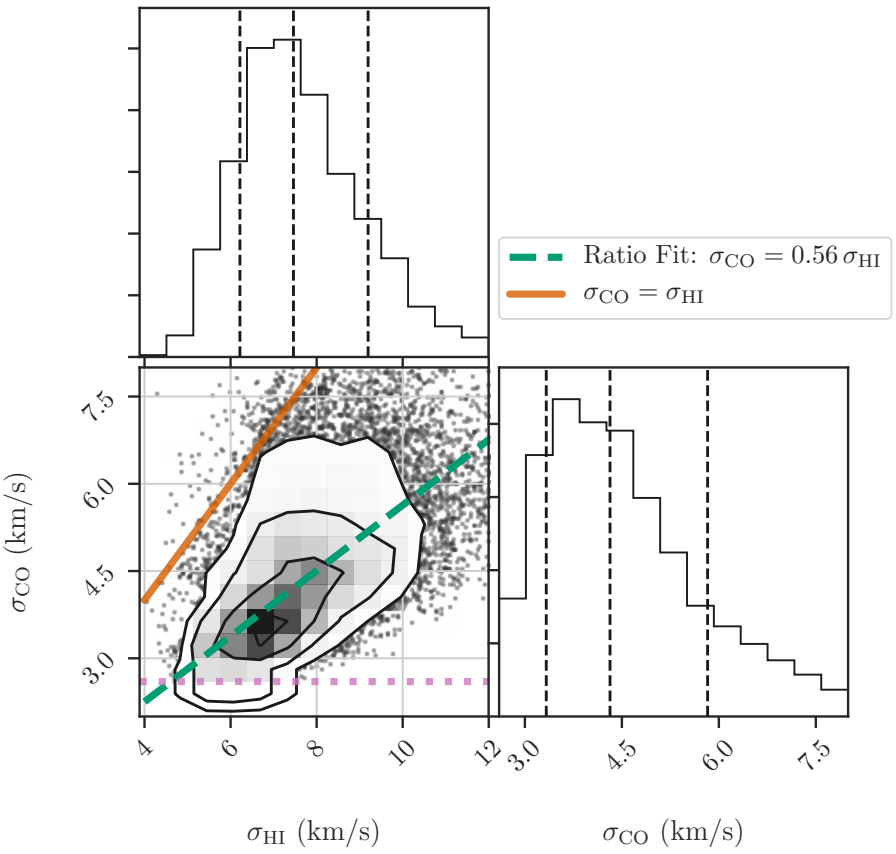

Figure 5.6: Gaussian line widths of individual HI and CO(2-1) profiles. The one-dimensional histograms show the distributions of σ_{HI} (top left) and σ_{CO} (bottom right) with vertical dashed lines indicating the 15th, 50th and 85th percentiles, respectively. The joint distribution is shown in the bottom left panel. Contours show the area containing data within the 1- to 4-σ limits of the distribution and black points show outliers beyond 4-σ, as described in Figure 5.2. The orange solid line is the line of equality. The horizontal dotted line indicates the CO(2-1) channel width of 2.6 km s^{-1}; no samples are included below this width. The green dashed line shows the fitted ratio of 0.56 ± 0.01. We note that our definition of a 'clean' component sample restricts HI line widths to be less than 12 km s^{-1}, and CO line widths must be less than 8 km s^{-1}. Typical uncertainties are 0.6 and 0.2 km s^{-1} for the CO and HI, respectively. A clear relation exists between the HI and CO line widths with intrinsic scatter.

with increasing spatial resolution when analyzed on a line-of-sight basis. We compare these line width ratios to those from the stacking analysis in §5.3.4.

The line width ratios we find are moderately smaller than the line-of-sight analysis by Mogotsi et al. (2016) for a sample of nearby galaxies. Fitting single Gaussians to HI and CO(2-1) spectra, they find a mean ratio of 0.7 ± 0.2 on spatial scales ranging from 200–700 pc. In contrast, the line width ratio we find is significantly steeper than extragalactic studies at higher physical resolution. In the LMC, Fukui et al. (2009) fit Gaussian profiles to both tracers where the CO peaks in a GMC and find a much shallower slope of 0.23 at a resolution of 40 pc.

5.3.3.4 Regional variations in the HI-CO line widths

To further investigate the observed correlation between HI and CO(2-1) line widths and the source of the scatter in this relationship, we highlight the positions of the line widths from three regions in Figure 5.7. These regions each have peak CO temperatures above the 75^{th} percentile, and so the observed scatter is not driven by the correlation between peak CO temperature and line width. By examining many regions on ~ 100 pc scales, including the three examples shown, we find that the line widths remain correlated on these scales, but the slope and offset of the line widths varies substantially. These regional variations are the source of the additional scatter required when fitting the HI-CO line width relation (§5.3.3.2).

By averaging over these local variations—over the full sample, in radial bins, and at different spatial resolutions—we consistently recover similar line width ratios. The lack of radial variation indicates that 500 pc radial bins provide a large enough sample to reproduce the scatter measured over the whole disk. If these regional variations arise from individual GMCs, the HI-CO line widths may indicate changes in the local environment or the evolutionary state of the cloud. If the latter is true, the lack of a radial trend is consistent with the radial distribution of cloud evolution types from Corbelli et al. (2017), which are well-mixed in the inner 6 kpc (see also Gratier et al., 2012).

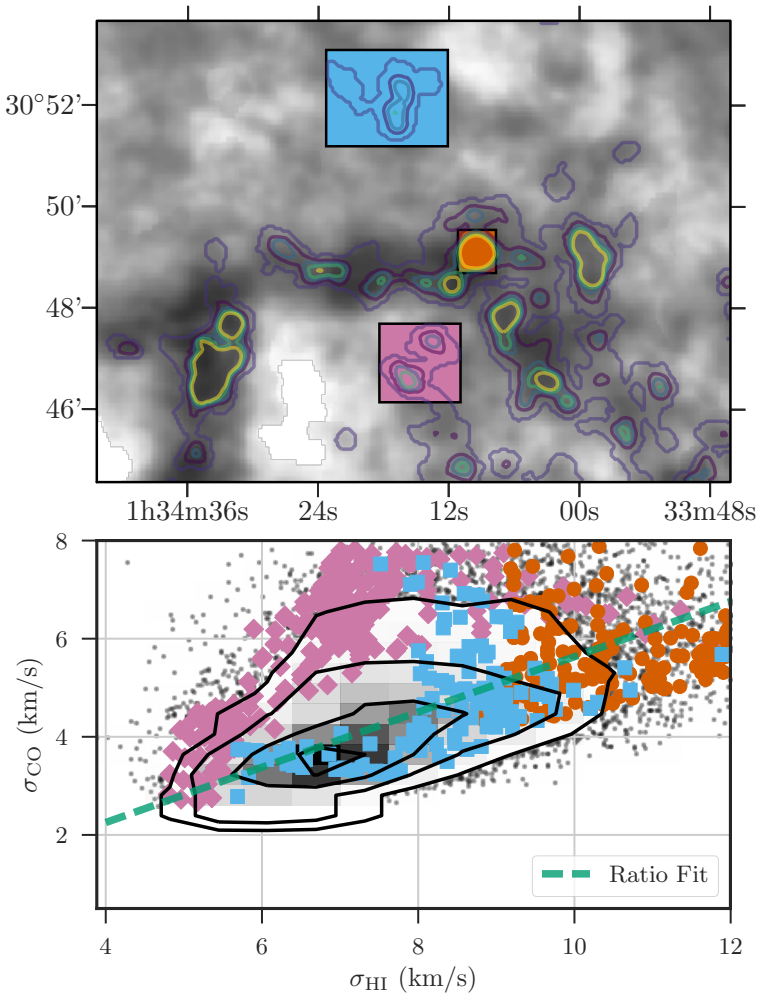

Figure 5.7: Top: The region around the Northern arm is shown, where the grayscale is the HI column density and contours are the CO(2-1) column densities with same levels shown in Figure 5.1. The coloured boxes indicate the line widths highlighted below. Bottom: The HI-CO(2-1) line width relation from Figure 5.6 with line widths highlighted according to their region. The typical uncertainty of the fitted line widths is 0.2 km s^{-1} for HI and 0.6 km s^{-1} for CO. The HI and CO line widths remain correlated within individual regions, but their position in the line width plane varies with displacements larger than their uncertainties. This suggests the line width relation is sensitive to environmental properties or the evolutionary state of GMCs.

5.3.4 Spectral properties from stacking versus individual lines-of-sight

Previous studies of HI and CO have found differing results between line stacking and fitting individual spectra. While there are some discrepancies in the spectral properties we find, our results from these two methods are more similar than the results from other studies. In this section, we explore the sources of discrepancy between the stacking and line-of-sight fit results and argue that most of the sources are systematic, due to the data or the analysis method.

The stacking analysis and line-of-sight fitting each have relative advantages and disadvantages. Stacked profiles provide an overall census of HI and CO(2-1) without conditioning on the spatial location of the emission. However, variations in the centre and width of individual spectra—along with asymmetries and multiple components—will lead to larger line wings than a Gaussian profile of equivalent width. This result is shown using a mixture model in Koch et al. (2018c).

The line-of-sight (LOS) analysis retains spatial information, providing distributions of spectral properties that can be connected to different regions. However, the simplistic decomposition of the HI of this analysis requires the detection of CO along the line-of-sight, and so only provides an estimate of HI properties where CO is detected. If HI where CO is detected differs from the global population of HI, the properties we find may not describe the typical HI line properties.

We determine the source of discrepancies in our stacking and fitting results by creating stacked profiles from the fitted LOS sample and their Gaussian models. There are five stacking tests we perform that are designed to control for variations in $\sigma_{\rm HWHM}$ or $f_{\rm wings}$. Table 5.3 gives the values for these parameters for each of the tests. We described the purpose of each stacking test and their derived line profile properties below:

1. *Stacked fitted model components aligned with the fitted CO mean velocity* – The fitted Gaussian components, not the actual spectra, are stacked.

This removes all emission far from the line centre and will minimize $f_{\rm wings}$ in the stacked profiles. Indeed, $f_{\rm wings}$ for HI and CO(2-1) are both significantly smaller than for the stacked profiles in §5.3.2 (Table 5.1). Stacking based on the fitted CO mean velocity will minimize the CO $\sigma_{\rm HWHM}$, while increasing the HI $\sigma_{\rm HWHM}$ due to the scatter between the HI and CO mean velocities. The CO $\sigma_{\rm HWHM}$ is narrower than all of the other stacked profiles, including those from §5.3.2, and is consistent with the mean CO LOS fitted width of $4.3^{+1.5}_{-1.0}$ km s^{-1} (Table 5.2).

2. *Stacked fitted model components aligned with the fitted HI mean velocity* – This stacking test is identical to (i), except the fitted HI mean velocities are used to align the spectra. Aligning the spectra with the HI mean velocity will decrease the HI $\sigma_{\rm HWHM}$ and increase the CO(2-1) $\sigma_{\rm HWHM}$, consistent with the measured properties.

3. *Stacked spectra in the LOS sample aligned with the fitted HI mean velocity* – The spectra in the LOS sample, rather than the model components used in (i) and (ii), are stacked aligned with the fitted HI mean velocities. The CO spectra in the sample are required to be well-modelled by a single Gaussian component, but there is a modest increase in $f_{\rm wings}$ to 0.07., larger than in tests (i) and (ii) The HI spectra, however, have significant line structure that is not modelled for, leading to a vast increase in $f_{\rm wings}$ to 0.19. The line widths of the HI and CO stacked profiles are the same within uncertainty.

4. *Stacked spectra in the LOS sample aligned with the* HI *peak velocity* – The spectra used in (iii) are now aligned with the HI peak velocities from §5.3.1. These stacked profiles are equivalent to the precedure used in §5.3.2 using only a sub-set of the spectra. This subset contains some of the outlier points in Figure 5.2, however $\sigma_{\rm HWHM}$ and $f_{\rm wings}$ do not significantly change from (iii). The outliers in the peak HI and CO velocity difference distribution do not contribute significantly to $\sigma_{\rm HWHM}$ or $f_{\rm wings}$.

5. *Stacked spectra in the entire LOS sample, including rejected fits, aligned with the* HI *peak velocity* – Finally, we create stacked profiles for the entire LOS sample considered in §5.3.3, including the LOS with rejected fits. This test is equivalent to (iv) with a larger sample. Relative to (iv), the HI $\sigma_{\rm HWHM}$ and $f_{\rm wings}$ both moderately increase, as expected when including LOS potentially with multiple bright spectral components. The CO stacked spectrum $\sigma_{\rm HWHM}$ marginally increases compared to (iv), however, $f_{\rm wings}$ increases by 33% to $f_{\rm wings} = 0.12$. This increase is driven in part by the CO spectra with multiple components.

From these tests, we can identify the source of the LOS and stacking discrepancies.

5.3.4.1 Smaller CO line-of-sight fitted line widths than from stacking

The larger CO stacked line widths are due to the scatter between the HI and CO peak velocity (Figure 5.2). This is demonstrated by comparing tests (i) and (ii), where the former is consistent with the median CO line width from the LOS fitting. The larger CO line width from stacking will lead to an overestimate of the HI-CO line width ratio.

5.3.4.2 Larger HI line-of-sight fitted line widths compared with stacking

The stacked HI line width towards LOS with CO detections (Table 5.3) is consistently larger than the HI stacked line width from all LOS (Table 5.1). There are two possible causes for this discrepancy. First, the HI where CO is detected has larger line widths than the average from all HI spectra. This source requires a physical difference in the atomic gas properties where molecular clouds are located, possibly related to the HI cloud envelope (Fukui et al., 2009). Alternatively, the HI components fitted here may be broadened by overlapping velocity components since our analysis does not account for this. However, from visually examining the fits to the LOS sample, most HI spectra

Table 5.3: Stacked line width ($\sigma_{\rm HWHM}$) and line wing excess ($f_{\rm wings}$) from the spectra used in the line-of-sight analysis (§5.3.3.1). The line widths do not strongly vary when changing the line centre definition or when the Gaussian model components are stacked rather than the actual spectrum. However, $f_{\rm wings}$ is sensitive to whether the full spectra or the models are used. The CO $f_{\rm wings}$ is also more sensitive to the how the stacking is performed than the HI.

	HI	CO(2-1)
(i) Fitted model components aligned to CO Model v_0		
$\sigma_{\rm HWHM}$ (km s^{-1})	7.6 ± 0.1	4.2 ± 0.9
$f_{\rm wings}$	$0.03^{+0.01}_{-0.01}$	$0.05^{+0.17}_{-0.13}$
(ii) Fitted model components aligned to HI Model v_0		
$\sigma_{\rm HWHM}$ (km s^{-1})	7.4 ± 0.1	4.6 ± 0.9
$f_{\rm wings}$	$0.03^{+0.01}_{-0.01}$	$0.03^{+0.16}_{-0.12}$
(iii) LOS spectra aligned to HI Model v_0		
$\sigma_{\rm HWHM}$ (km s^{-1})	7.4 ± 0.1	4.8 ± 0.9
$f_{\rm wings}$	$0.19^{+0.01}_{-0.01}$	$0.07^{+0.10}_{-0.12}$
(iv) LOS Spectra aligned to HI $v_{\rm peak}$		
$\sigma_{\rm HWHM}$ (km s^{-1})	7.3 ± 0.1	4.7 ± 0.9
$f_{\rm wings}$	$0.20^{+0.01}_{-0.01}$	$0.08^{+0.15}_{-0.12}$
(v) All LOS spectra aligned to HI $v_{\rm peak}$		
$\sigma_{\rm HWHM}$ (km s^{-1})	7.6 ± 0.1	4.8 ± 0.9
$f_{\rm wings}$	$0.22^{+0.01}_{-0.01}$	$0.12^{+0.14}_{-0.11}$

would need to have highly overlapping components for the average HI LOS line width to be broadened, and this does not seem likely for most spectra (§5.6.3.2). In order to definitely determine which of these sources leads to the larger HI line widths, we require decomposing the HI spectra without conditioning on the location of the CO emission. However, this analysis is beyond the scope of this paper. We favour larger HI line widths in molecular cloud envelopes as the source of this discrepancy.

5.3.4.3 Sources of the line wing excess

These five tests provide restrictions on the source of the line wing excess in both tracers. In the HI, $f_{\rm wings}$ is only changed when the model components ((i) and (ii)) are stacked rather than the full HI spectra ((iii)–(v)). This result is consistent with the line structure and wings evident in individual HI spectra, as explored in Koch et al. (2018c). The scatter in the fitted velocities (HI or CO) can account for $f_{\rm wings} \sim 0.03$ in the stacked HI profiles.

For the CO stacked profiles, there are variations in $f_{\rm wings}$ from multiple sources. There are small contributions to $f_{\rm wings}$ from the scatter in the HI fitted mean velocities (Test (ii); $f_{\rm wings} = 0.03$) and the scatter between the HI and CO peak velocities or fitted mean velocities (Tests (iii) & (iv); $f_{\rm wings} = 0.01$). Multi-component CO spectra account for $f_{\rm wings} \lesssim 0.04$–$0.05$ from comparing Tests (iii) and (iv) to Test (v). This estimate is an upper limit since we do not control for contributions from real line wings versus multiple spectral components. Finally, comparing Tests (iii) and (iv) to Test (ii), excess line wings can directly account for $f_{\rm wings} = 0.04$–0.05.

The different sources of line wing excess in the CO stacked profiles implies that there is marginal evidence for CO line wings. As described above, stacking systematics and multi-component spectra can account for $f_{\rm wings} \lesssim 0.08$–$0.09$, roughly half of the line wing excess of $f_{\rm wings} = 0.21$ from the stacked profile towards all LOS (Table 5.1). The discrepancy between the total $f_{\rm wings}$ for the CO stacked profiles and the systematics is then 0.13, though the large $f_{\rm wings}$ uncertainties can account for the remaining line wing excess. Including

the error beam contribution of up to 45% of the line wing excess (§5.3.2), we find that $\sim$ 80% of the CO line wing excess can be accounted for without requiring the presence of real CO line wings. However, due to the estimated uncertainties, we cannot rule out their presence.

5.4 A marginal thick molecular disc in M33

Studies of CO in the Milky Way and nearby galaxies find evidence of two molecular components: a thin disk dominated by GMCs, and a thicker diffuse molecular disk. Our results, however, suggest that M33 has a marginal thick molecular component, unlike those found in other more massive galaxies, based on (i) finding smaller CO line widths relative to the HI and (ii) the marginal detection of excess CO line wings. In this section, we compare our results to previous studies and address previous works arguing for a diffuse component on large-scales in M33.

Evidence for a diffuse molecular component has been demonstrated with extended emission in edge-on galaxies (e.g., NGC 891; Garcia-Burillo et al., 1992), separating ^{12}CO emission associated with denser gas in the Milky Way (Roman-Duval et al., 2016), comparing the flux recovered in interferometric data to the total emission in single-dish observations (Pety et al., 2013; Caldú-Primo et al., 2015; Caldú-Primo & Schruba, 2016a), and large CO line widths in nearby galaxies (Combes & Becquaert, 1997; Caldu-Primo et al., 2013; Caldú-Primo et al., 2015).

In M33, a diffuse molecular component has been suggested based on the CO flux recovered in GMCs (Wilson & Scoville, 1990), comparing the ^{13}CO to ^{12}CO spectral properties (Wilson & Walker, 1994), and a non-zero CO power-spectrum index on kpc scales (Combes et al., 2012). Rosolowsky et al. (2007) find that 90% of the diffuse CO emission is located $<$ 100 pc to a GMC, and suggest the emission is due to a population of unresolved, low-mass molecular clouds (see also Rosolowsky et al., 2003).

Our stacking analysis, and the stacking analysis in Druard et al. (2014),

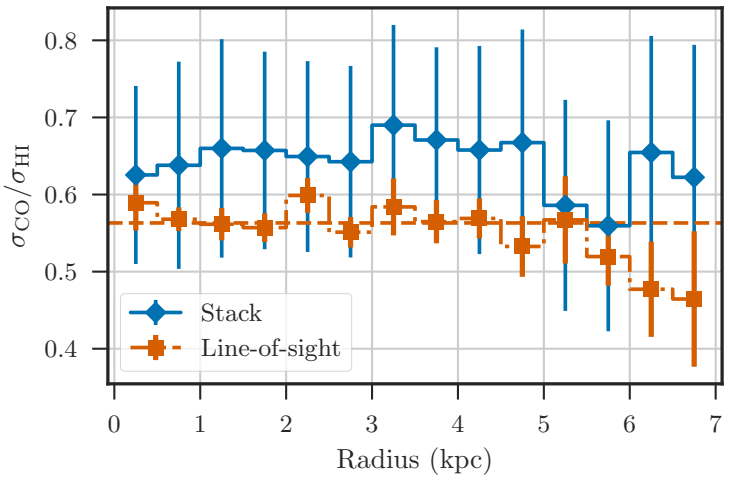

Figure 5.8: Line width ratio from stacked profiles (blue solid diamonds) and the average from the line-of-sight fits (orange dot-dashed squares) versus galactic radius at 80 pc (20″) resolution. The 1-σ uncertainties on the stacked widths are dominated by the CO(2-1) channel width. The errors on the line-of-sight fits are the standard deviation in the radial bin divided by the square root of the number of independent components. The horizontal dashed line is the fitted ratio 0.56 shown in Figure 5.6. Both methods have line width ratios smaller than unity, unlike other (more massive) nearby galaxies (Caldu-Primo et al., 2013), suggesting M33 lacks a significant thick molecular disk.

shows that the CO line widths are consistently smaller than the HI, unlike the line widths from most nearby galaxies on $\sim$ 500 pc scales (Combes & Becquaert, 1997; Caldu-Primo et al., 2013). Figure 5.8 summarizes our results by showing the CO-to-HI line width ratios of the stacked line widths and the radially-binned averages from the line-of-sight analysis at 80 pc resolution. The ratios from the stacked profiles are consistently 10% larger than the average of the line-of-sight fits due to using the HI peak velocities to align the CO spectra (§5.3.4). The line width ratio increases but remains less than unity when the data are smoothed to resolutions of 160 and 380 pc (§5.3.2.2), scales comparable to some of the data in Caldu-Primo et al. (2013).

Stacking analyses of CO by Caldú-Primo et al. (2015) and Caldú-Primo & Schruba (2016a) find a wide Gaussian component to CO stacked profiles that arises only in single-dish observations. The high-resolution (80 $\times$ 350 pc)

CO observations of M31 from Caldú-Primo & Schruba (2016a) constrain this wide component to scales of $\sim$ 500 pc and larger. Coupled with the large CO line widths, these results suggest the wide Gaussian component is due to a thick molecular disk. In our analysis, the wide Gaussian component would contribute to the line wing excess ($f_{\rm wings}$)[9]. In M33, we find a qualitatively similar line wing excess to Caldú-Primo & Schruba, however, up to $\sim$ 80% of the excess is due to stacking systematics and error beam pick-up from the IRAM 30-m telescope (§5.3.4.3; Druard et al., 2014). The remaining fraction of the CO line wing excess is small, and would correspond to a much smaller contribution from a wide Gaussian component compared to those found by Caldú-Primo et al. (2015) and Caldú-Primo & Schruba (2016a).

These results strongly suggest M33 has a marginal thick molecular disk and is instead more consistent with the findings from Rosolowsky et al. (2007) where diffuse CO emission is clustered near GMCs and may be due to unresolved low-mass clouds (flux recovery with spatial scale with these CO data is explored in Sun et al., 2018). There remains ambiguity about the diffuse molecular component in M33 from other analyses, and whether M33 is the only nearby galaxy with a marginal thick molecular disk. We address these issues in the following sections. First, we demonstrate that the large-scale CO(2-1) power-spectrum identified in Combes et al. (2012) can be explained by the exponential disk scale of the CO emission rather than a thick molecular disk. We then note the similarity of the line width ratios found by Caldu-Primo et al. (2013) for NGC 2403 and our results.

5.4.1 Comparison to a thick molecular disk implied by power-spectra

Previous work by Combes et al. (2012) found that the power-spectra of HI and CO integrated intensity maps of M33 have shallow indices extending to kpc scales, with distinct breaks near $\sim$ 120 pc where the indices becomes steeper.

[9]We stress that the flux in a wide Gaussian component will not equal $f_{\rm wings}$. However, an increase in the amplitude or width of the wide Gaussian component will be positively correlated with $f_{\rm wings}$.

The non-zero slope on kpc scales suggests there is significant large-scale diffuse emission in M33 from both HI and CO, in contrast with our findings for CO from the line widths.

To explain the non-zero index on large scales, we compare the large-scale distribution of emission in M33 for CO and HI. Bright CO emission is broadly confined to individual regions on scales comparable to the beam size (Figure 5.1) and has a radial trend in the average surface density that is well-modelled by an exponential disk with a scale length of ~ 2 kpc (Gratier et al., 2010; Druard et al., 2014). This radial trend implies that the detection fraction per unit area of CO also depends on radius, providing additional power in the power-spectrum on $\sim$ kpc scales. On the other hand, HI emission is widespread throughout the disk and the surface density is approximately constant within the inner 7 kpc (Koch et al., 2018c). The difference in the large-scale radial trends of HI and CO will affect the large-scale ($\sim$ kpc) parts of the power-spectrum.

We demonstrate how the exponential CO disk affects the power-spectrum by calculating the two-point correlation function of the GMC positions from Corbelli et al. (2017). By treating the CO emission as a set of point sources at the GMC centres, any large-scale correlations must result from spatial clustering, rather than extended CO emission. Figure 5.9 shows that the two-point correlation function of the cloud positions has a non-zero correlation up to scales of $\sim$ 2 kpc, similar to the exponential CO disk scale. This non-zero correlation on these scales will correspond to a non-zero CO power spectrum slope, thus demonstrating that the large-scale power-spectrum does not imply the presence of a thick molecular disk.

This result may also explain the large change in the CO power-spectrum index across the ~ 120 pc break point found by Combes et al. (2012). Distinct breaks in the power-spectra, and other turbulent metrics, are useful probes of the disk scale height (Elmegreen et al., 2001; Padoan et al., 2001). The power-spectrum index is predicted to change by +1 on scales larger than the disk scale height since large-scale turbulent motions are confined to two dimensions

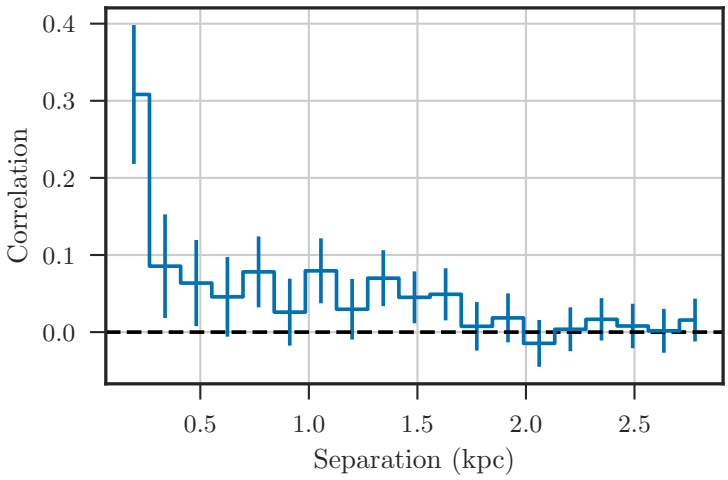

Figure 5.9: Two-point correlation function of the GMC positions from Corbelli et al. (2017) measured in 150 pc bins. The uncertainties are estimated from 2000 bootstrap iterations. Most of the correlation occurs on < 150 pc scales, but the structure of CO emission due to the disk gives non-zero correlations up to ~ 2 kpc scales.

in the disk (Lazarian & Pogosyan, 2000). Combes et al. (2012) find the CO power-spectrum index changes by +2.2 across the break, significantly larger than the expected change of +1. For the HI, the index change of +0.8 across the break is much closer to the expected change.

The similarity Combes et al. (2012) find between the CO and HI break points in the power-spectra also differs from the disk scale heights implied by the line widths we find. For ratios $\sigma_{\rm CO}/\sigma_{\rm HI} < 1$, the disk scale height of CO should be smaller than the HI; we can approximate the ratio of the disk scale heights from the line width ratio. For $R_{\rm gal} < 7$ kpc, the stellar surface density is larger than the total gas surface density in M33 (Corbelli et al., 2014). Measurements of the stellar velocity dispersion find ~ 20 km s^{-1} in the inner disk (Kormendy & McClure, 1993; Gebhardt et al., 2001; Corbelli, 2003), suggesting the stellar disk scale height is larger than the HI and CO disk scale heights. If this result holds true for $R_{\rm gal} < 7$ kpc, the ratio of the CO and HI disk scale heights is the ratio of the line widths: $H_{\rm CO}/H_{\rm HI} \approx \sigma_{\rm CO}/\sigma_{\rm HI}$ (Combes & Becquaert, 1997), for line widths measured at the disk scale height. Based

on our analysis at 80 and 160 pc, the line width ratio is ~ 0.6, suggesting that the CO scale height should be $\sim 60\%$ of the HI scale height[10].

The discrepancy with the scale heights we measure and the similar scale of the break points from Combes et al. (2012) may be a limitation of the data resolution used in their analysis. They use HI and CO data at $12''$ (~ 50 pc) resolution (Gratier et al., 2010) and the constraints on the scale of the break are limited by the beam size, as shown in their Appendix B. Higher resolution observations (~ 20 pc) should determine whether there is a difference in the disk scale heights traced by HI and CO.

5.4.2 Similarities with the flocculent spiral NGC 2403

From the sample of nearby galaxies studied in Caldu-Primo et al. (2013), there are two galaxies that are also dominated by the atomic component throughout most of the disk: NGC 925 and NGC 2403 (see also Schruba et al., 2011). As these are the closest analogs to M33 in their sample, we compare the stacked profile analyses from Caldu-Primo et al. (2013) to our results.

The NGC 925 line width ratios are consistent with unity, however, the signal-to-noise in the CO map limit the analysis to a few radial bins at the galaxy centre. The signal-to-noise of the NGC 2403 data is higher, allowing the analysis to be extended to larger radii (70% of the optical radius) providing a number of radial bins for comparison. NGC 2403 is also the closest galaxy in their sample and the physical scale of the beam is $\sim$200 pc, a factor of about two coarser than the resolution of our M33 data. Interestingly, the line width ratios, outside of the galactic centre ($\sim 0.1 R_{25}$), are consistently smaller than unity, with an average of $\sim$0.8. The increased line widths in the inner disk are likely affected by beam smearing. With the same data, the line-of-sight analysis by Mogotsi et al. (2016) find a smaller line width ratio of $\sim$0.7. Both of these line width ratios are comparable to our results in

[10]The line width ratio from the LOS analysis may be too small, due to the large HI line widths (§5.3.4.2), while the ratio from the stacking analysis is too large due to the CO stacked line width being larger than the LOS analysis (§5.3.4.1). A ratio of ~ 0.6 is in-between the ratios from the two methods.

M33 on 160 pc (38″) scales. Our results are then consistent with the ratios for NGC 2403 from Caldu-Primo et al. (2013) and Mogotsi et al. (2016), suggesting that galaxies with atomic-dominated neutral gas components have at most a moderate contribution from a diffuse molecular disk.

5.5 Summary

We explore the spectral relationship of the atomic and molecular medium in M33 on 80 pc scales by comparing new VLA HI observations (Koch et al., 2018c) with IRAM 30-m CO(2-1) data (Gratier et al., 2010; Druard et al., 2014). We perform three analyses—the difference in the velocity at peak HI and CO(2-1) brightness, spectral stacking, and fits to individual spectra—to explore how the atomic and molecular ISM are related. Each of these analyses demonstrates that the spectral properties between the HI and CO are strongly correlated on 80 pc scales. We also show that relationship between the HI and CO line widths, on 80 pc scales, from individual spectral fits depend critically on identifying the HI most likely associated with CO emission, rather than all HI emission along the line-of-sight.

1. The velocities of the HI and CO peak temperatures are closely related. The standard deviation in the differences of these velocities is 2.7 km s^{-1}, slightly larger than the CO channel widths (2.6 km s^{-1}; Figure 5.2). Significant outliers in the velocity difference ($>$ 10 km s^{-1}) occur where the HI spectrum has multiple components and the CO peak is not associated with the brightest HI peak. These outliers are removed when modelling only the HI component associated with CO (§5.3.3).

2. By stacking HI and CO spectra aligned to the velocity of the peak HI brightness, we find that the width of the CO stacked profile (4.6 ± 0.9 km s^{-1}) is smaller than the HI stacked profile (6.6 ± 0.1 km s^{-1}) on 80 pc scales, unlike similar analyses of other (more massive) nearby galaxies that measure comparable line widths on 500 pc scales (e.g., Caldu-Primo

et al., 2013). The widths of the stacked profiles slowly decrease with galactocentric radius.

3. By repeating the stacking analysis at lower spatial resolutions of 160 pc (38″) and 380 pc (95″), we find that the CO(2-1)-to-HI line width ratio remains constant within uncertainty. We estimate how beam smearing contributes at each resolution and find that resolutions of 80 and 160 pc have a similar contribution to the line width from beam smearing. Beam smearing contributes more at a resolution of 380 pc and can explain the increased line widths relative to those at 160 pc. However, the CO line width remains smaller than the HI on all scales.

4. We perform a spectral decomposition of HI spectra limited to where CO is detected. The CO spectra are fit by a single Gaussian, while the Gaussian fit to the HI is limited to the closest peak in the HI spectrum. We carefully inspect and impose restrictions to remove spectra where this fitting approach is not valid. The average HI and CO line widths of the restricted sample are $7.4^{+1.7}_{-1.3}$ and $4.3^{+1.5}_{-1.0}$ km s^{-1}, where the uncertainties are the 15$^{\rm th}$ and 85$^{\rm th}$ percentiles of the distributions, respectively.

5. The average CO line width from the line-of-sight fits are smaller than those from the stacking analysis. This difference results from aligning the CO spectra to the HI peak velocity, while there is scatter between the HI and CO velocity at peak intensity (Figure 5.2). Recovering larger CO stacked line widths relative to those from individual spectra is a general result that will result whenever CO is aligned with respect to another tracer, such as HI. The amount of line broadening is set by the scatter between the line centres of the two tracers. Thus, line stacking based on a different tracer will bias the line widths to larger values, but is ideal for recovering faint emission (Schruba et al., 2011).

6. The average HI line width from the line-of-sight fits (7.4 ± 1.5 km s^{-1}) is *larger* than the stacked profile width (6.6 ± 0.1 km s^{-1}). The larger line

widths are due to either multiple highly-blended Gaussian components that are not modelled correctly in our analysis, or that the HI associated with CO emission tends to have larger line widths. We favour the latter explanation since our line-of-sight analysis has strong restrictions to remove multi-component spectra (§5.6.3); however, we do not fully decompose the HI spectra and cannot rule out the former explanation.

7. The line-of-sight fits highlight a strong correlation between HI and CO line widths (Figure 5.6). We fit for the line width ratio, accounting for errors in both measurements, and find $\sigma_{\rm CO} = (0.56 \pm 0.01)\ \sigma_{\rm HI}$, smaller than the ratios from the stacked profiles due to the smaller average CO line width and larger average HI line width. There is no trend between the line width ratio and galactocentric radius (Figure 5.8). When repeated at a lower spatial resolution, we find that the HI and CO line widths are increased by the same factor, leading to the same line width ratio, within uncertainties.

8. The scatter in the relation between the HI and CO line widths is larger than the statistical errors and results from regional variations (Figure 5.7). The line widths of HI and CO remain correlated when measured in individual regions, but exhibit systematic offsets with respect to the median HI and CO line widths. These regional variations affect both the HI and CO line widths and suggest that the local environment plays an important role in setting the line widths.

9. We perform stacking tests with the fitted LOS components to constrain sources of the line wing excess. We find that the error beam pick-up from the IRAM 30-m telescope (Druard et al., 2014) and systematics of the stacking procedure can account for $\sim 80\%$ of the line wing excess in the CO(2-1) (§5.3.4.3). Combined with a CO-to-HI line width ratio less than unity, this result implies that M33 has at most a marginal thick molecular disk. We point out that previous analyses of NGC 2403 give

similar results to ours (Caldu-Primo et al., 2013; Mogotsi et al., 2016), suggesting that galaxies where the atomic component dominates the cool ISM may lack a significant thick molecular disk.

Scripts to reproduce the analysis are available at https://github.com/Astroua/m33-hi-co-lwidths[11].

Acknowledgments

We thank the referee for their careful reading of the manuscript and comments. EWK is supported by a Postgraduate Scholarship from the Natural Sciences and Engineering Research Council of Canada (NSERC). EWK and EWR are supported by a Discovery Grant from NSERC (RGPIN-2012-355247; RGPIN-2017-03987). This research was enabled in part by support provided by WestGrid (www.westgrid.ca), Compute Canada (www.computecanada.ca), and CANFAR (www.canfar.net). The work of AKL is partially supported by the National Science Foundation under Grants No. 1615105, 1615109, and 1653300. The National Radio Astronomy Observatory and the Green Bank Observatory are facilities of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

Code Bibliography: CASA (versions 4.4 & 4.7; McMullin et al., 2007) — astropy (Astropy Collaboration et al., 2013) — radio-astro-tools (spectral-cube, radio-beam; radio-astro-tools.github.io) — matplotlib (Hunter, 2007) — seaborn (Waskom et al., 2017) — corner (Foreman-Mackey, 2016) — astroML (Vanderplas et al., 2012)

[11]Code DOI: https://doi.org/10.5281/zenodo.2563161

5.6 Appendix

5.6.1 Line broadening from beam smearing

Spectral line widths can be broadened wherever there is a large gradient in the rotation velocity on scales of the beam size. This line broadening, commonly referred to as beam smearing, tends to have the largest effect near the centres of galaxies, where the rotation curve is steep, and can lead to significant increases in the line width of stacked profiles (e.g., Stilp et al., 2013a; Ianjamasimanana et al., 2015; Caldú-Primo et al., 2015).

We require constraints on beam smearing when comparing line widths measured at different spatial resolutions in our data to distinguish whether broadened line profiles are the result of physical processes. We estimate the maximum broadening from beam smearing by using a rolling tophat filter on the peak HI velocity map to calculate the standard deviation over one beam. This operation measures the beam-to-beam variation in the peak velocity field. We note that these variations may not be entirely due to beam smearing and could arise from local variations in velocity, such as those measured for molecular cloud and envelope rotation in the HI (Imara et al., 2011). Therefore, our estimates are an *upper limit* on line broadening due to beam smearing. This is a similar measurement to the approach used by Caldú-Primo & Schruba (2016a), where they measure the width of the velocity distribution on local scales along the major and minor axes of M31.

We compute the standard deviation in the peak HI velocity surface at the original (80 pc/20″) and degraded resolutions (160 pc/38″ and 380 pc/95″) used in Sections 5.3.2 & 5.3.3. Since large-scale variations in the peak HI velocity describe the circular rotation curve, we calculate the average values of the standard deviation surfaces in 0.5 kpc galactocentric radial bins. If beam smearing significantly broadens the line, we expect the profile of average values to follow the derivative of the circular rotation curve, which is steepest within the inner 2 kpc of M33 (Koch et al., 2018c). Figure 5.10 shows the average of the standard deviation surfaces at the three different beam sizes used for the

analysis. The average radial profiles for beam widths of 80 and 160 pc do not have strong radial trends and show that beam smearing contributes at most ~ 2 km s^{-1} to the line width. We calculate the area-weighted average of the radial profiles in Figure 5.10 and find values of $2.0^{+2.1}_{-1.8}$ km s^{-1} and $1.5^{+1.7}_{-0.8}$ km s^{-1} for beam sizes of 80 and 160 pc, respectively. The uncertainties quoted here are the 15$^{\rm th}$ and 85$^{\rm th}$ percentiles of the radial profiles with the same area-weighted averaging applied. Since the CO(2-1) channel width is 2.6 km s^{-1}, the line width broadening of CO(2-1) from beam smearing is similar to the correction factor for the channel width.

Using the average line-of-sight CO line width of 4.3 km s^{-1} at a resolution of 80 pc (Table 5.2), the correction due to beam smearing gives a $\sim 10\%$ in the line width.

The average standard deviation profile measured at a beam size of 380 pc ($95''$) shows a strong radial trend within the inner 4 kpc, as expected from beam smearing. The broadening from beam smearing is particularly strong within $R_{\rm gal} < 2$ kpc, where the maximum average standard deviation is ~ 8.2 km s^{-1}. The area-weighted average, as applied to the higher resolution measurements, is $2.8^{+1.0}_{-1.0}$ km s^{-1}. Subtracting this mean value in quadrature from the 380 pc stacked line widths (Table 5.1) gives corrected line widths of 8.4 ± 0.9 km s^{-1} for HI and 6.7 ± 3.6 km s^{-1} for CO(2-1). The CO(2-1) is not constraining due to the uncertainty from the channel width, however the HI line width range demonstrates that the 0.9 km s^{-1} increase in the line width between the 160 and 380 pc data can entirely be explained by beam smearing.

5.6.2 Forward-modelling the spectral response function

We forward-model the individual LOS fits to the CO data in §5.3.3 with an approximation for the spectral response function. Here, we briefly describe the fitting process.

We approximate the spectral response function in the CO(2-1) IRAM-30m data based on the nearest neighbour channel correlation found by Sun et al. (2018, $r = 0.26$ for scales > 70 pc, adjusted to a distance of 840 kpc used here).

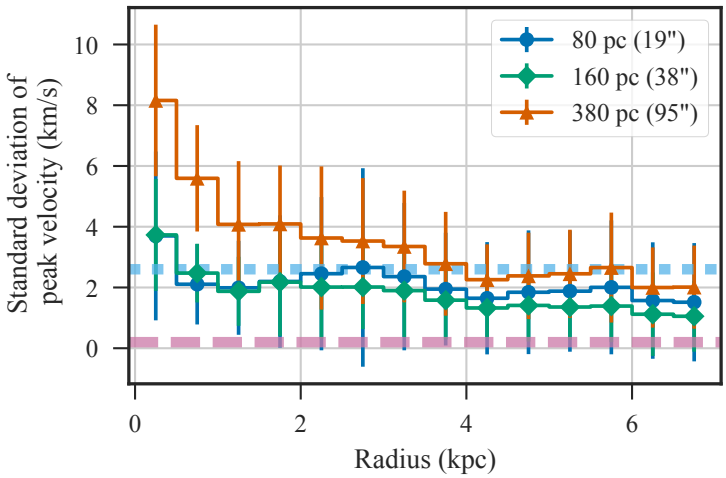

Figure 5.10: Average standard deviations from the peak HI velocity map measured over one beam. Three average curves are shown measured within 0.5 kpc bins at the original beam size (80 pc/20″; blue circles), and at twice (160 pc/38″; green diamonds) and five times (380 pc/95″; orange triangles) the original beam size. Error bars correspond to the standard deviation within each bin, uncorrected for the number of independent samples to demonstrate where the distributions are consistent with 0 km s^{-1}. The thick, horizontal lines correspond to the HI (pink dashed; 0.2 km s^{-1}) and CO(2-1) (cyan dotted; 2.6 km s^{-1}) channel widths. The average values represent the maximum line broadening that could result from beam smearing. Line widths at resolutions of 80 and 160 pc are uniformly broadened by $\sim$ 2 km s^{-1}, while the broadening at a resolution of 380 pc is $\sim$ 3 km s^{-1} and increases to 8 km s^{-1} in the inner kpc.

Using the empirically-derived relation from Leroy et al. (2016), this correlation corresponds to a channel-coupling factor of $k = 0.11$ for a three-element Hann-like kernel ($[k, 1-2k, k]$), which we adopt as the spectral response function.

We forward-model the spectral response in two steps:

1. The Gaussian model is sampled at the spectral channels of the observed spectra. This sampling is equivalent to taking the weighted average of the Gaussian over the spectral channel width (Δv):
$$G(v) = \frac{A}{(\Delta v)^2}\left[\mathrm{erf}\left(\frac{\mu - (v - \Delta v/2)}{\sqrt{2}\sigma}\right) - \mathrm{erf}\left(\frac{\mu - (v + \Delta v/2)}{\sqrt{2}\sigma}\right)\right], \tag{5.5}$$
for a Gaussian centered at velocity μ with an amplitude of A and width of σ that is averaged over channels centered at v. The channel averaging is equivalent to convolving the Gaussian with a top-hat kernel with a width of Δv.

2. The Gaussian sampled over the spectral channels is convolved with the Hann-like kernel described above. This step accounts for the measured channel correlations in the observations.

The sampled and convolved spectrum is then compared to the observed spectrum and the sum of the squared distances is the quantity minimized in the fit. Using this approach removes biases in the fitted line width parameters (Koch et al., 2018b). This approach is similar to the forward-modelling in Rosolowsky et al. (2008).

We note that the non-linear least-squares fit used in this paper assumes that the data uncertainties are independent, which is not true due to the spectral-response of the data. To test whether the parameter uncertainties from the covariance matrix of the fit are underestimated due to being correlated, we repeat the fitting procedure on 1000 simulated spectra sampled with channel widths of $\Delta v = \sigma$. White-noise is added to the spectra then convolved with the Hann-like kernel to give a peak S/N of 5. We find that $\sim 72\%$ of the fitted parameters are within the 1–σ uncertainty interval[12]. This is similar to

[12]An example of this test is available at https://doi.org/10.5281/zenodo.1491796

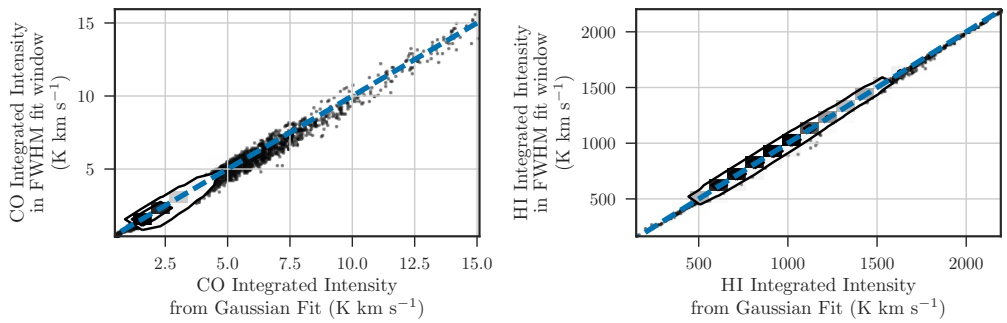

Figure 5.11: The CO (left) and HI (right) integrated intensities from the spectra within the FWHM window used for the fit (y-axis) compared with the integral over the fitted Gaussian models (x-axis). The blue-dashed line is the line of equality. There is little deviation from the line of equality, indicating that the Gaussian models describe the data within the FWHM fitting windows well.

the expected 68.2% expected for a two-tailed p-test, demonstrating that the parameter uncertainties are not underestimated despite the correlated errors.

5.6.3 Validating the gaussian decomposition

We demonstrate our limited Gaussian decomposition method (§5.3.3) and perform two validation checks on the sample used in the analysis.

The first check compares the surface densities from the integral over the fitted Gaussian model to the integrated intensity of the data located within the model's FWHM, scaled by $1/\mathrm{erf}(\sqrt{2})$ to account for emission outside of the mask. Figure 5.11 shows excellent agreement between the two methods for the HI and CO fits. This implies that the peaks are well-described by a Gaussian and validates the choice of model.

The second validation check is a comparison of the integral over the fitted Gaussian model with respect to the integrated intensities over the whole profile. Figure 5.12 shows these quantities for the CO and HI. For the CO integrated intensities, there is no significant variation between the two quantities. This is expected, since we require that the CO profiles be well-fit by a Gaussian in order to be in the sample. The discrepancy between these quanti-

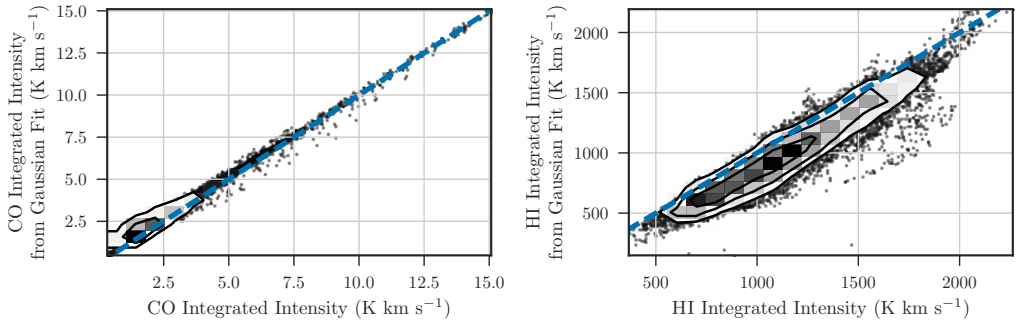

Figure 5.12: The CO (left) and HI (right) integrated intensities from the integral over the fitted Gaussian models (y-axis) compared to the integrated intensity over the whole spectra (x-axis). The blue-dashed line is the line of equality. There remains good agreement for the CO spectra, however the HI integrated intensity from the Gaussian fit is consistently smaller the from the entire spectra due to the additional HI line structure.

ties for the HI is larger. Again, this is expected, since the masking used in the HI fitting is introduced to remove spectral features unlikely to be associated with the CO component.

With these two checks, we are confident that the clean sampled used for the analysis describes only single-peaked CO profiles and HI profiles with a well-defined peak associated with the CO emission.

5.6.3.1 Effect of HI masking on fitting

We investigate how the FWHM mask affects the fitted HI line width by repeating the fitting without the mask. This procedure has been used in other studies relating HI and CO line profiles from individual spectra (Fukui et al., 2009; Mogotsi et al., 2016). For HI profiles with multiple components or prominent line wings, we expect that the fitted profiles without masking will be much wider. Figure 5.13 shows that most HI profiles are indeed wider without the masking, with the median width increasing from 7.4 to 8.3 km s^{-1}. This result highlights the need to carefully disinguish bright HI emission from extended line wings to avoid biasing the HI line widths and reinforces the use of the upper-limit of 12 km s^{-1} in $\sigma_{\rm HI}$ set here to minimize contamination in our

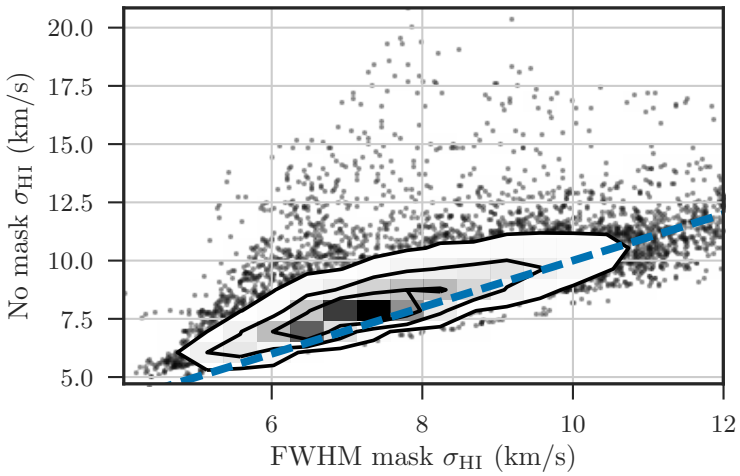

Figure 5.13: HI line widths fit with and without a FWHM mask around the peak. The contours represent the 2- to 4-σ limits of the population, and points outside the contours are outliers beyond 4-σ. The blue-dashed line indicates equality between the line widths.

sample.

5.6.3.2 Examples of fitted spectra

In Figures 5.14 and 5.15, we demonstrate fitted HI and CO spectra and how our criteria for the analysis sample removes clear issues. Figure 5.14 is an example of a valid fit to both HI and CO, while Figure 5.15 demonstrates a poor fit that is excluded from the analysis.

5.6.4 Stacked profile widths

Table 5.4 shows the stacked line widths ($\sigma_{\rm HWHM}$) in radial bins. These line widths are plotted in the top panel of Figure 5.4 and described in §5.3.2.1.

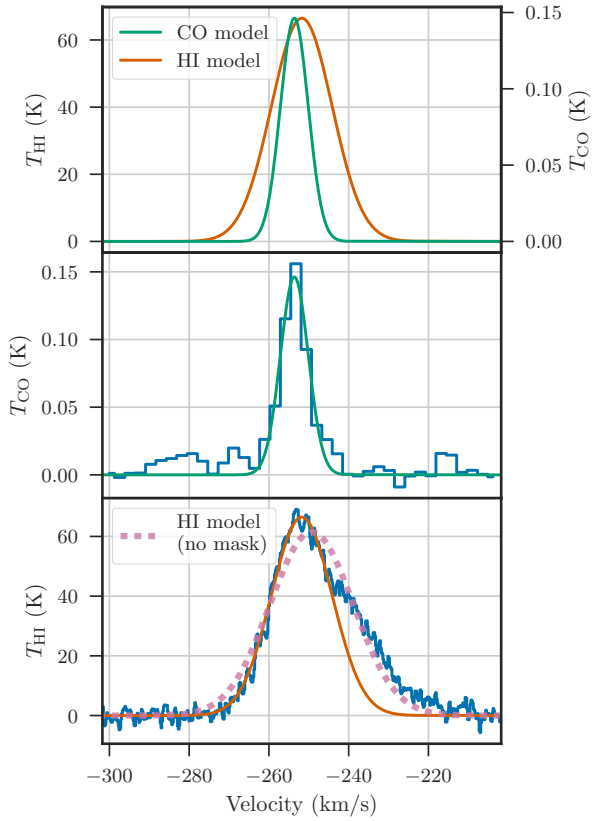

Figure 5.14: Example of the HI and CO Gaussian fitting showing well-fit single components that are included in our clean sample. The top panel shows the fitted HI and CO profiles. The middle panel shows the CO spectrum with the fit overlaid. The bottom panel shows the same for the HI data, and also includes a fit to the HI data if no masking is applied when fitting (thick-dashed line; §5.6.3.1). The HI fit the brightest peak is significantly improved when masking is used.

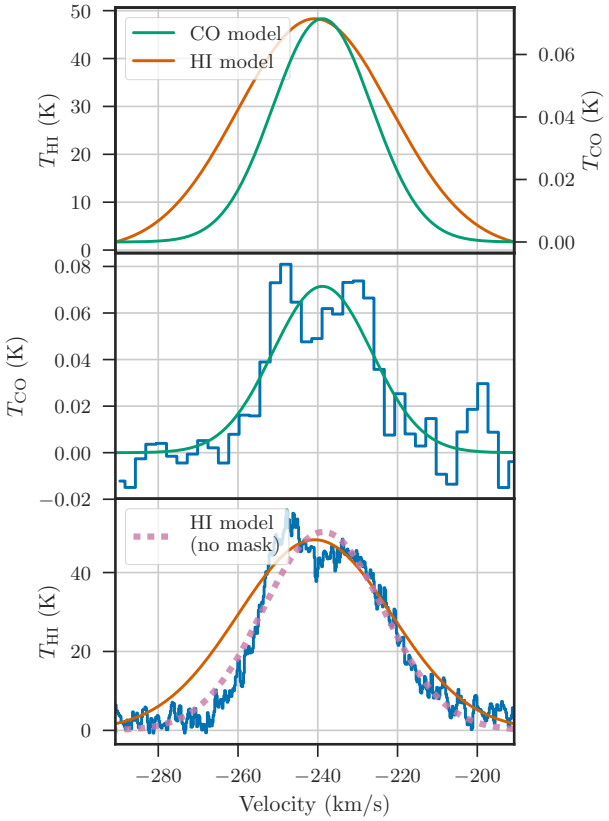

Figure 5.15: Same as Figure 5.14. This example shows failed fits in both tracers due to multiple-components and is rejected from the clean sample. There appears to be two Gaussians in both spectra and fitting a multi-Gaussian model should distinguish between the two. Extending this analysis to multi-component spectra will be the focus of future work.

Table 5.4: HWHM line widths ($\sigma_{\rm HWHM}$) for the HI and CO(2-1) stacked profiles in 500 pc radial bins at 80 pc resolution. The uncertainties are propagated assuming an uncertainty of half the channel width and the uncertainty of each point in the spectrum is the standard deviation of values within that channel scaled by the square-root of the number of beams.

$R_{\rm gal}$ (kpc)	$\sigma_{\rm HI}$ (km s^{-1})	$\sigma_{\rm CO}$ (km s^{-1})
0.0–0.5	8.0 ± 0.1	5.0 ± 0.9
0.5–1.0	7.3 ± 0.1	4.6 ± 1.0
1.0–1.5	6.9 ± 0.1	4.5 ± 1.0
1.5–2.0	7.3 ± 0.1	4.8 ± 0.9
2.0–2.5	7.4 ± 0.1	4.8 ± 0.9
2.5–3.0	7.4 ± 0.1	4.7 ± 0.9
3.0–3.5	6.8 ± 0.1	4.7 ± 0.9
3.5–4.0	7.2 ± 0.1	4.8 ± 1.0
4.0–4.5	6.7 ± 0.1	4.4 ± 1.0
4.5–5.0	6.9 ± 0.1	4.6 ± 1.0
5.0–5.5	6.9 ± 0.1	4.1 ± 1.0
5.5–6.0	6.9 ± 0.1	3.9 ± 0.9
6.0–6.5	6.6 ± 0.1	4.3 ± 1.0
6.5–7.0	6.1 ± 0.1	3.8 ± 1.1

Chapter 6

Spatial Power Spectra of Dust across the Local Group: No Constraint on Disc Scale Height

Koch, E.W et al. (2020). "Spatial power spectra of dust across the Local Group: No constraint on disc scale height," *MNRAS*, 492, 2663–2682.

Preface

This chapter, published as Koch et al. (2020), focusses on modelling one-dimensional power spectra, which are shown in Figures 6.2–6.5 & 6.8 (also see Appendix A). For clarity, I note that these power spectra are plotted with the corresponding physical scale x on the x-axis defined as $x = 1/\nu$, where ν is the *spatial* frequency. This convention differs from how traditional power spectra are shown, with the x-axis in spatial frequency or wavenumber, but is useful for this work where our 4 targets vary in distance. This convention is used elsewhere in the literature where power spectra of observational data are presented (e.g., Combes et al., 2012).

Abstract

We analyze the 1D spatial power spectra of dust surface density and mid to far-infrared emission at 24–500 μm in the LMC, SMC, M31, and M33. By

forward-modelling the point-spread-function (PSF) on the power spectrum, we find that nearly all power spectra have a single power-law and point source component. A broken power-law model is only favoured for the LMC 24 μm MIPS power spectrum and is due to intense dust heating in 30 Doradus. We also test for local power spectrum variations by splitting the LMC and SMC maps into 820 pc boxes. We find significant variations in the power-law index with no strong evidence for breaks. The lack of a ubiquitous break suggests that the spatial power spectrum does not constrain the disc scale height. This contradicts claims of a break where the turbulent motion changes from 3D to 2D. The power spectrum indices in the LMC, SMC, and M31 are similar (2.0–2.5). M33 has a flatter power spectrum (1.3), similar to more distant spiral galaxies with a centrally-concentrated H_2 distribution. We compare the power spectra of HI, CO, and dust in M31 and M33, and find that HI power spectra are consistently flatter than CO power spectra. These results cast doubt on the idea that the spatial power spectrum traces large scale turbulent motion in nearby galaxies. Instead, we find that the spatial power spectrum is influenced by (1) the PSF on scales below ~ 3 times the FWHM, (2) bright compact regions (30 Doradus), and (3) the global morphology of the tracer (an exponential CO disc).

6.1 Introduction

Turbulence is an integral part of the dynamics in the interstellar medium (ISM). Within the inertial range of turbulence, the self-similar structure of the density and velocity fields produce a power-law distribution, which can be measured using statistical techniques like the power spectrum (Elmegreen & Scalo, 2004). Together, the density and velocity fields constrain the energy power spectrum $\mathbf{E}(\mathbf{k})$. This can directly be compared to turbulence models for incompressible (Kolmogorov, 1941) and compressible gas (Burgers, 1948; Fleck, 1996; Galtier & Banerjee, 2011; Federrath, 2013). ISM observations provide usable constraints on 3D turbulent velocity and density fluctuations

from the 2D line-of-sight velocity and column density maps (Federrath et al., 2010). This connection offers a method for constraining the turbulent energy power spectrum from observational data.

Of particular interest for the star formation process and galaxy evolution is distinguishing what mechanism drives turbulence throughout a galaxy. Because turbulence decays quickly (~ 10 Myr), the ubiquity of observed turbulent properties implies the need for a near-continuous source of turbulent energy injection (Mac Low & Klessen, 2004b). Observational constraints on the turbulent driving scale may provide a clean measurement to distinguish between different sources of energy injection. This connection can be difficult to make with Milky Way observations as line-of-sight confusion makes it difficult to distinguish scales at and above the disc scale height (> 100 pc). As a result, high dynamic range extragalactic observations may offer the best way to trace the scale of energy injection.

The spatial power spectrum of a turbulent cascade offers a potential solution to constrain the disc scale height and driving scale in face-on galaxies. The index of the energy power spectrum changes with both the type of turbulence and the number of spatial dimensions. For the latter, the index is expected to steepen by $+1$ as the turbulent motions transition from being confined from three- to two-dimensions (e.g., Lazarian & Pogosyan, 2000). Extragalactic observations that resolve scales below the disc scale height are ideal for testing whether this "break" in the power spectrum indeed occurs, using the column density or the line-of-sight velocity fields. From this break scale, the disc scale height can be measured, constraining quantities like the turbulent energy injection on galactic scales (e.g., Tamburro et al., 2009; Koch et al., 2018c; Utomo et al., 2019a), and the mid-plane pressure (Blitz & Rosolowsky, 2006) that are used in star formation models based on vertical dynamical equilibrium (Ostriker et al., 2010).

Several studies, primarily using column density or intensity maps, investigate the spatial power spectrum in nearby galaxies. Some studies find power spectra well-described by both a single power-law (e.g., Stanimirovic et al.,

2000; Dutta et al., 2013; Zhang et al., 2012). Others find a broken power-law (e.g., Elmegreen et al., 2001; Dutta et al., 2009a; Combes et al., 2012), where the break has been interpreted as the disc scale height. These and other studies also find a large range in the power law index. This is true even when comparing results use a single traced like the 21-cm HI line (e.g., Dutta et al., 2013).

This range in extragalactic power spectrum properties makes it difficult to draw general conclusions about the nearby galaxy population. One reason for the confusion may be that extragalactic power spectrum analyses use heterogeneous data and techniques. In general, extragalactic studes have also not corrected for steepening on small scales due to the PSF response (excepting Muller et al., 2004), though this effect is commonly account for in Galactic power spectrum analyses (Miville-Deschênes et al., 2003; Martin et al., 2015; Blagrave et al., 2017). This issue was also noted by Grisdale et al. (2017) who found that the break points in HI power spectra in a few nearby galaxies are consistently limited by the PSF scale.

A further issue to consider with extragalactic power spectra is how galactic structure not dominated by turbulence (i.e., spiral arms) affects the power spectrum shape. These large-scale distributions are known to contribute additional power on large-scales. Grisdale et al. (2017) show that changes in the mass distribution steepens the column density power spectrum from galaxy-scale simulations. Koch et al. (2019b) show how the clustering of GMC locations in M33's inner disc contributes to an excess in the power spectrum up to scales near the disc scale length ($\sim$ 2 kpc; Druard et al., 2014).

Accurate measurements of the power spectrum are particularly important now because recent advances in galaxy-scale numerical simulation resolve similar scales to current observations of Local Group galaxies (e.g., Grisdale et al., 2017; Dobbs et al., 2018; Garrison-Kimmel et al., 2019). Comparing the power spectra between these observations and simulations can provide a powerful diagnostic for how large-scale galactic structure affects the power spectrum shape (e.g., Grisdale et al., 2017). For example, several simulations show a power

spectrum break that is altered by stellar feedback (Bournaud et al., 2010; Pilkington et al., 2011; Combes et al., 2012; Grisdale et al., 2017), though the prominence of spiral arms also appears to play a role (Renaud et al., 2013).

In this paper, we present a uniform analysis of 1D dust emission power spectra in four Local Group galaxies, the Large and Small Magellanic Clouds, M31, and M33. We use archival *Spitzer* and *Herschel* data, as well as dust surface density maps from Utomo et al. (2019b). We compare power spectrum properties across different galactic environments while resolving scale similar to or below the disc scale height ($\sim$ 100 pc; Kalberla & Kerp, 2009). Our analysis models the point-spread function (PSF) on the power spectrum shape and demonstrates that a single power-law combined with unresolved point sources can reproduce most of the observed power spectra. We present the maps used in §6.2 and the power spectrum model in §6.3. We discuss the implications of our modelling in §6.4, including comparisons between IR bands and galaxies, and how the dust power spectrum relates to power spectra of HI, tracing the atomic ISM, and CO, tracing the molecular ISM. Our uniform power spectrum analysis of multiple phases in multiple galaxies offers a benchmark for simulations of Local Group-like galaxies.

Throughout this paper, we define $P(k)$ as the 1D power spectrum produced from an intensity or surface density maps and the power spectrum index β as $\propto k^{-\beta}$ such that $\beta > 0$.

6.2 Observations

We focus our study on the Magellanic Clouds, M31, and M33. These are the closest targets uniformly observed across the mid- to far-infrared by both *Spitzer* (Werner et al., 2004) and *Herschel* (Pilbratt et al., 2010). Due to their large angular size and proximity ($<$ 1 Mpc), these targets maximize the spatial range that can be studied in their power spectra. The *Spitzer* and *Herschel* maps of the Magellanic Clouds have resolve $\sim$ 10 pc scale, well below the expected scale height of both the atomic and molecular gas discs.

We use existing *Spitzer* MIPS (24, 70, & 160 μm Rieke et al., 2004), and *Herschel* PACS (100 & 160 μm Poglitsch et al., 2010) and SPIRE (250, 350, & 500 μm Griffin et al., 2010) data products from several projects: LMC *Spitzer* MIPS (SAGE; Meixner et al., 2006), *Herschel* PACS & SPIRE (HERITAGE; Meixner et al., 2013); SMC *Spitzer* MIPS (SAGE-SMC & S^{3}MC; Gordon et al., 2006, 2011; Bolatto et al., 2007), *Herschel* PACS & SPIRE (HERITAGE; Meixner et al., 2013); M31 *Spitzer* MIPS (Barmby et al., 2006), *Herschel* PACS & SPIRE (Groves et al., 2012; Draine et al., 2014); M33 *Spitzer* MIPS (Hinz et al., 2004; Tabatabaei et al., 2007), *Herschel* PACS & SPIRE (HerM33es; Kramer et al., 2010).

Altogether, we create 1D power spectra from maps in eight infrared bands in our analysis. We include both the MIPS and PACS 160 μm maps to check for consistency between different instrumental effects and noise levels. As a check, we did rerun our analysis on background-subtracted maps and found that the background remove had little effect on the power spectrum properties. This lack of change in the power spectrum is expected since the background tends to be both low intensity and smooth on large-scales.

A key component in our analysis is the effect of the instrumental PSF response on the power spectrum shape. We use the PSF and convolution kernels from Aniano et al. (2011) to model for PSF effects. We also convolve each map to the "moderate" Gaussian size provided in Table 6 of Aniano et al. (2011), again using their publicly available convolution kernels.

We also analyze the dust surface density maps from Utomo et al. (2019b), which were derived applying a uniform analysis to the *Herschel* data for each of our targets. A modified blackbody model is fit to the spectral energy distribution (SED) from 100-500 μm following the methodology of Chiang et al. (2018) and Gordon et al. (2014). The dust maps are provided at a common set of physical resolutions; here, we use the highest resolution for each galaxy: 13 pc for the LMC and SMC, and 167 pc for M31 and M33.

Figure 6.1 shows the dust surface density maps from Utomo et al. (2019b) for each galaxy. The region displayed in the figure shows the area used for the

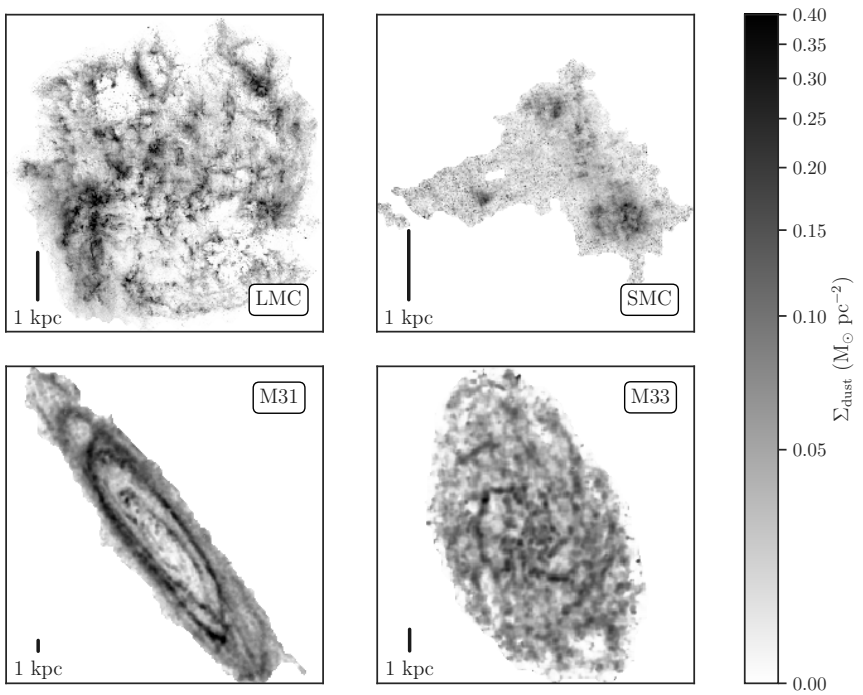

Figure 6.1: Dust surface density maps of the LMC, SMC, M31, and M33 from Utomo et al. (2019b). The bars in the lower left corners show the 1 kpc scale shown for each galaxy. The region shown for each galaxy is used for all images in the power spectrum analysis (§6.3).

analysis of all maps in all bands. Thus, the power spectra can be compared directly.

Throughout this paper, we adopt distances of 62.1 kpc for the SMC (Graczyk et al., 2014), 50.2 kpc to the LMC (Klein et al., 2014), 744 kpc to M31 (Vilardell et al., 2010), and 840 kpc to M33 (Freedman et al., 2001). These are the same distances used by Utomo et al. (2019b) to create the dust maps at common physical resolutions that we use here.

6.3 Power spectrum analysis

We characterize and compare the spatial structure in the LMC, SMC, M31, and M33 with the 1D spatial power spectrum from intensity or dust surface density maps, a commonly-used technique for describing ISM structure from

$\sim$AU to kpc scales (e.g., Elmegreen & Scalo, 2004). We present the power spectrum calculation in §6.3.1, the power-law model and fitted results in §6.3.2, and the model selection criteria in §6.3.3. Except for the MIPS 24 μm results for the LMC, all of the power spectra the we measure are well-fit by a single power-law plus point source component. In §6.3.5, we demonstrate that 30 Doradus is responsible for a power spectrum break in the LMC MIPS 24 μm. Finally, §6.3.6 presents local dust surface density power spectra from 820 $\times$ 820 pc^2 square regions in the LMC and SMC. This analyses allows us to explore variations in the power spectrum index. These local power spectra are also well-fit by a single power-law model.

6.3.1 Calculating power spectra

We use `TurbuStat` (Koch et al., 2019a)[1] to compute the 1D spatial power spectrum. `TurbuStat` implements a common version of many turbulence statistics described in the literature, including the spatial power spectrum. While `TurbuStat` can model the full 2D power spectra of images, we focus this study on modelling 1D power spectra azimuthally-averaged in Fourier space. The 1D power spectrum, $P(k)$, is most commonly used in extra-galactic studies.

When large values are at the edge of the map, the Gibbs phenomenon causes ringing in the Fourier transform, which manifests as a strong cross-shape in the 2D power spectrum. Since the ringing will affect the 1D power spectrum shape, we apply a Tukey function to smoothly taper the edges of the map prior to computing the power spectrum. The maps that require this added step are described in §6.3.2 and §6.6.1.

6.3.2 Modelling the power spectra

We consider two models to describe the 1D power spectrum shape: (1) a single power-law and (2) a broken power-law. In both cases, we allow an optional point source component. Both models account for extended emission with the

[1]Version 1.0; turbustat.readthedocs.io

power-law components and the response of bright, individual point sources on small scales with a constant component.

The single power-law model for a 1D power spectrum $P(k)$ is:

$$P_{\rm single}(k) = P_{\rm ext}(k) + P_{\rm pt}(k) = A\,k^{-\beta} + B. \tag{6.1}$$

This model has three free parameters to fit: the power-law amplitude A, the index β, and the point-source contribution B.

The broken power-law model accounts for a change in the power-law index at some scale. This model has been used in previous extragalactic studies (Block et al., 2010; Combes et al., 2012). We adopt a broken power-law model following the form implemented in Astropy Collaboration et al. (2013):

$$P_{\rm broken}(k) = A \left(\frac{k}{k_b}\right)^{-\beta} \left\{\frac{1}{2}\left[1 + \left(\frac{k}{k_b}\right)^{1/\Delta}\right]\right\}^{(\beta-\beta_2)\Delta} + B. \tag{6.2}$$

This model adds three additional parameters relative to Equation 6.1, two of which are left as free parameters when fitting. The parameters β and β_2 describe the power-law index below and above the break k_b, respectively. This form of a broken power-law smoothly varies between the power-law components, with the "smoothness" set by the Δ parameter. We fix $\Delta = 0.1$ based on visually comparing model solutions. Given that the fitting is done in frequency pixel units and the bin size of 1 is used for all power spectra, we expect this to be an appropriate choice for our analysis. This smooth version of a broken power-law offers a more realistic description of the data, rather than a model with a sharp break at k_b

Equations 6.1 and 6.2 are physically-motivated, idealized models that do not account for any real observational effects. In the simplest interpretation, the power-law component results direction from turbulent density fluctuation while the point source component reflect, e.g., young stellar objects and embedded star forming regions (Seale et al., 2009).

Instrumental systematics affect the observed shape of the power spectrum, causing it to deviate from the idealized models above. Fortunately, most

of these effects can be account for by forward modelling. In this analysis, we forward model the point spread function (PSF) response for each map. Multiplying by the PSF response ($P_{\rm PSF}(k)$), the models from Equations 6.1 & 6.2 become:

$$P_{\rm obs}(k) = P_{\rm PSF}(k) \cdot P_{\rm model}(k). \tag{6.3}$$

Since the PSF response has a fixed form, it does not introduce additional free parameters in the model. For a Gaussian response, $P_{\rm PSF}(k) \propto \exp\left(-4\pi^2\sigma^2_{\rm beam}k^2\right)$, where $\sigma_{\rm beam} = {\rm FWHM}/\sqrt{8{\rm log}2}$ is the Gaussian rms of the beam. Similar models that include the PSF response have been used in several studies (e.g., Miville-Deschênes et al., 2002; Muller et al., 2004; Martin et al., 2015; Blagrave et al., 2017).

We fit the power spectra of the maps at two resolutions: (1) the native resolution, and (2) convolved to a Gaussian beam using the "moderate" kernels listed in Table 6 of Aniano et al. (2011). At the native resolution, we account for the non-Gaussian PSF shape by regridding the PSF map from Aniano et al. (2011) to have the same pixel size as the observed map and using its 1D power spectrum as $P_{\rm PSF}(k)$ in Equation 6.3. For the convolved maps, we use the analytic form for a Gaussian PSF.

We found that fits to the the power spectra of the dust surface density maps were improved by including an uncorrelated white noise term C:

$$P_{\rm obs}(k) = P_{\rm PSF}(k) \cdot P_{\rm model}(k) + C. \tag{6.4}$$

This additional C term is due to fitting the dust SED to individual pixels. The inherent uncertainty of the SED fit adds some noise to the dust surface density map. Since the fits are performed for each pixel, this additional noise is not affected by the PSF and is reasonably approximated as white (uncorrelated) noise.

Since the amplitudes A, B and C in the model vary over several orders of magnitude, we fit the $\log_{10}$ of these parameters to make it easier to sample large variations. Due to the potentially wide range in parameters from map

to map, we adopt uninformative uniform priors on the parameters:

$$\log_{10} A \sim \mathcal{U}(-20, 20) \tag{6.5}$$

$$\log_{10} B \sim \mathcal{U}(-20, 20) \tag{6.6}$$

$$\log_{10} C \sim \mathcal{U}(-20, 20) \quad \text{for Eq. 6.4} \tag{6.7}$$

$$\beta \sim \mathcal{U}(0, 10) \tag{6.8}$$

$$\beta_2 \sim \beta + \mathcal{N}(0, 10) \quad \text{for Eq. 6.2} \tag{6.9}$$

$$k_b \sim \mathcal{U}(k_{\min}, k_{\max}) \quad \text{for Eq. 6.2} \tag{6.10}$$

The chosen parameter ranges are significantly wider than the expected values and none of the fitted parameters converged to the edge of a parameter range. We also note that adopting wide Gaussian priors on the parameters did not affect the fits.

When fitting Equation 6.2, we treat the second power-law component, on scales below k_b, as a perturbation on the large-scale index β. This allows for $\beta \sim \beta_2$, thereby converging to Equation 6.1 when a break is not strongly preferred in the fit. The break point k_b is sampled uniformly over the whole range of spatial frequencies k. The importance of k_b diminishes at large k_b where forward-modelling the PSF response dominates the power spectrum shape. In these cases, β_2 and k_b could be well-constrained but the fit will be indistinguishable from Equation 6.1.

We fit the 1D power spectra and assume that the standard deviation of the azimuthal average is a reasonable uncertainty. Since most of these maps have a high signal to noise ratio, particularly on large scales, the variations in radial bins will be larger than the inherent uncertainty. Thus, we treat the 1D power spectrum values ($P_{\rm 1D}$) as independent samples drawn from a normal distribution with a width inferred from the standard deviation in azimuthal bins. We draw these samples in $\log_{10}$ space to avoid sampling negative values for the power spectrum:

$$\log_{10} P^{\star}_{\rm 1D} \sim \mathcal{N}\left(\log_{10} P_{\rm 1D}, \sigma_{\log_{10} P_{\rm 1D}}\right). \tag{6.11}$$

The model is fit to the sampled values $P_{\rm 1D}^{\star}$, ensuring the data uncertainty is reflected in the parameter posterior distributions.

When fitting maps from the photometric bands to Equation 6.3, we only consider scales above $3\times$ the Gaussian standard deviation of the beam to avoid regions where pixelization or convolution residuals dominate the power spectrum shape (see §A). We also limit the fit to scales less than $k < k_{\rm max}/3$, where $k_{\rm max}$ is the inverse of half the map shape. This removes large deviations in the largest bin that arise from the need to account for the total intensity in the image. These large scale bins are estimated from just a few samples in the 2D power spectrum and thus have a large uncertainty (see Figure 6.2). The dust surface density maps from Utomo et al. (2019b) have pixel sizes 2.5 times smaller than the beam which avoids small scales dominated by convolution residuals. Therefore, we include the smallest scales of the column density power spectra in the fit to the surface density.

We use the PYMC3 package (Salvatier et al., 2016) to fit the models, using Sequential Monte Carlo to sample the parameter space (Del Moral et al., 2006), as we found it rapidly converged for this problem. Sequential Monte Carlo runs a set of parallel Markov chains through a series of stages. At each stage, the sampling progresses from the prior to posterior distribution by tempering, controlled by a tempering parameter[2] $\beta^{\star}$. At each stage, $\beta^{\star}$ is increased according to the samples in the previous step, starting at 0 for the prior distribution and ending at 1 for the posterior distribution. For our fits, we found a good balance between computational cost and convergence using 100 chains that sample over 6000 iterations for each step. For comparison, we fit several power spectra using the Levenberg-Marquardt algorithm, which provides similar parameter values but severely underestimates parameter uncertainties and covariance. Using Markov Chain Monte Carlo (MCMC) and accounting for the data uncertainty (Equation 6.11) provides realistic parameter uncertainties.

[2]Typically β is used for the tempering parameter, but we adopt $\beta^{\star}$ to avoid confusion with the power spectrum index β.

6.3.3 Model selection

We fit each of the power spectra to the single (Equation 6.1) and broken power-law (Equation 6.2) models while forward-modelling the PSF response. To compare the models, we compute the Widely-applicable Information Criterion (WAIC; Watanabe, 2010), as implemented in PYMC3, to determine the preferred model. WAIC estimates the out-of-sample prediction accuracy from a Bayesian model based on the log-likelihood from the MCMC parameter samples, with a correction for the number of variables to account for overfitting. We note that the model comparisons calculated using leave-one-out (LOO) cross-validation are consistent with those from WAIC for our fits (Vehtari et al., 2017).

We choose the preferred models by comparing the WAIC and its uncertainty between the two models. The preferred model should minimize the WAIC. However, we find that the WAIC is similar for many of the fits. In this case, we choose the simpler single power-law model (Equation 6.1) given the lack of clear evidence for a broken power-law. In many of these cases, the break point approaches the PSF FWHM and the broken power-law has a diminishing influence on the fit quality.

6.3.4 Fit results

For all of the fits but one (§6.3.5), we find that the power spectra are well-fit by a single power-law and point source model (Equation 6.1) with no significant requirement for a broken power-law. In our measurements, the PSF response can account for any observed steepening of the power spectra on smaller scales.

Figure 6.2 shows the power spectra and fits for the dust surface density maps from Utomo et al. (2019b). The PSF response, shown separately for each galaxy, has a noticeable effect on the shape of the power spectrum on scales $\sim$ 3–4 times the FWHM. By incorporating the PSF response into our model, the fits shown in Figure 6.2 account for the apparent break point on those scales. Table 6.1 provides the fitted parameters using Equation 6.4.

These results show that the power spectra are all well-fit by a single power-law, plus a point source term for maps with ~ 10 pc resolution (i.e., the LMC and SMC) and do not require a physical break point in the model.

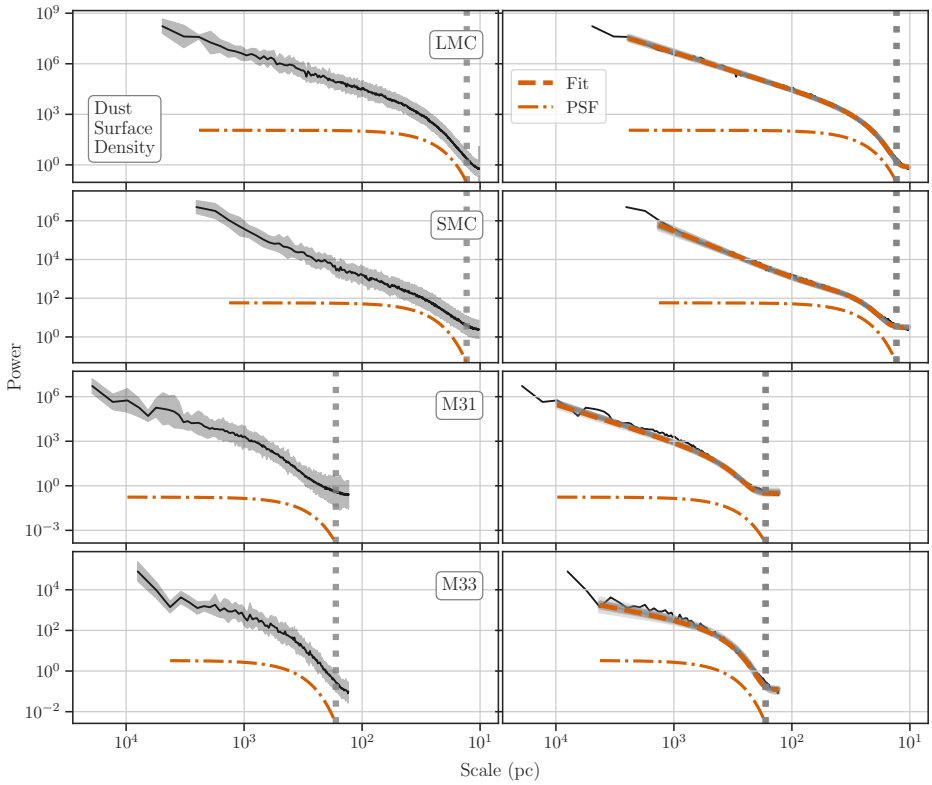

Figure 6.2: 1D power spectra of the dust surface density (left column) and the fitted models (right column; Table 6.1) for the LMC, SMC, M31, and M33 shown on a common physical scale. The orange dashed-dotted line shows the power spectrum of the PSF response, scaled to an order of magnitude below the amplitude from the fit. The vertical thick gray dotted line is the FWHM of the PSF response. The standard deviation on the power spectrum is shown in the shaded gray region in the left panel. The dashed orange line in the right column panels is the best fit model, and the underlying gray lines are the model fits from 10 random draws of the MCMC. The power spectra are well-fit by a single power-law and point-source component, while the PSF response is solely responsible for the break point on small scales.

Table 6.1: Fit parameters for the 1D dust surface density power spectra to Equation 6.4. Uncertainties are the 1-σ interval estimated from the MCMC samples. Missing entries in $\log_{10}B$ are unconstrained in the fit and not used.

Galaxy	Resolution ($''$)	Phys. Resolution (pc)	$\log_{10}$ A	β	$\log_{10}$ B	$\log_{10}$ C
LMC	53.4	13	1.67 ± 0.08	2.18 ± 0.05	2.82 ± 0.32	0.79 ± 0.08
SMC	43.2	13	-0.16 ± 0.25	2.47 ± 0.15	2.52 ± 0.04	2.93 ± 0.18
M31	46.3	167	0.02 ± 0.13	2.46 ± 0.14	–	0.33 ± 0.06
M33	41.0	167	1.32 ± 0.15	1.11 ± 0.14	–	0.13 ± 0.03

Figure 6.3 shows the power spectra and fits for the MIPS 24 μm maps at their original resolution. The MIPS 24 μm PSF has noticeable non-Gaussian features that result in a "step" in the PSF response[3]. Each of the galaxy maps show this step feature in their power spectra, indicating these breaks are solely from the PSF. All of the power spectrum fits are best described by a single power-law model, except for the LMC MIPS 24 μm shown in the top panel of Figure 6.3, which we explore in further detail in §6.3.5. Repeating this analysis with the MIPS 24 μm convolved to a Gaussian PSF, we find consistent power-law indices, demonstrating that the fits are not strongly dependent on the PSF model. Table 6.2 provides the complete fit parameters for all bands and resolutions, with the WAIC used for model selection.

[3]For M33, the step in the PSF is on $\sim$ 80 pc scales.

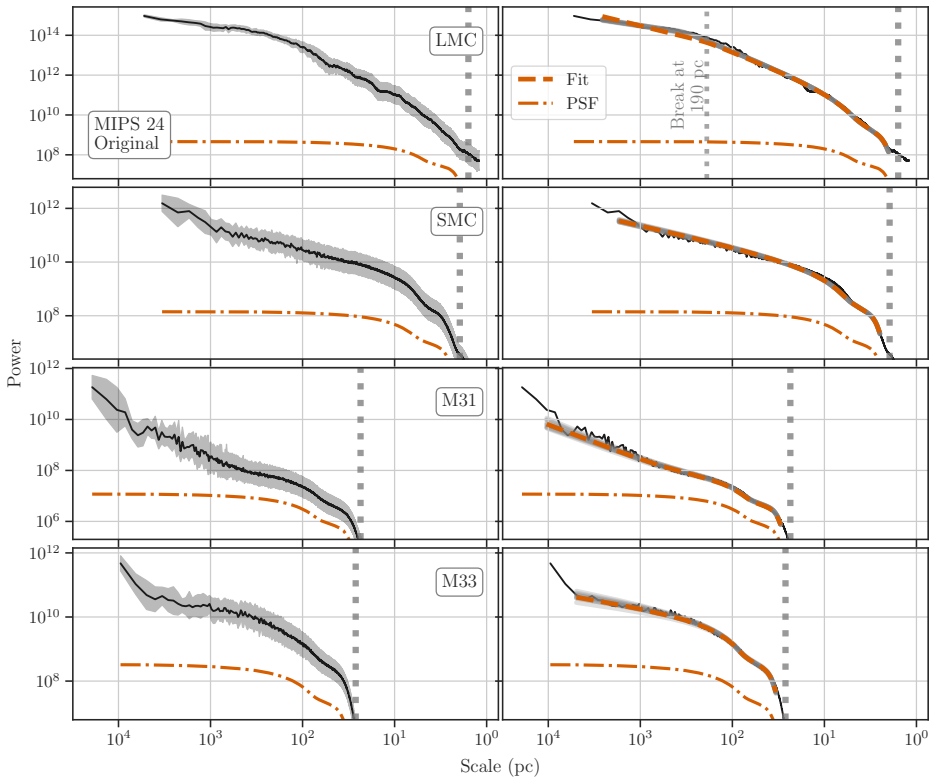

Figure 6.3: 1D power spectra of MIPS 24 μm maps (left column) and the fitted models (right column; Table 6.2), similar to Figure 6.2. A similar figure for each band is included as supplemental information. The MIPS 24 μm PSF is highly non-Gaussian, as is clear from the extra "step" in the response on small scales. Similar to the dust surface density, the power spectra are well fit without a break point. The only exception across all of the bands is the LMC MIPS 24 μm power spectrum in the top panel. A broken power-law model (Eq. 6.2) is marginally preferred with a break point at 190 pc, shown in the labeled gray vertical line. We show in §6.3.5 that 30 Doradus causes this deviation.

Table 6.2: Fit parameters for the individual band power spectra to Equations 6.1 & 6.2 with forward-modelling the PSF response (Equation 6.3. Uncertainties are the 1-σ interval estimated from the MCMC samples. Missing entries in $\log_{10}B$ are unconstrained in the fit and not used, while missing entries in other parameters (β_2 and k_b) are cases where the broken power-law model is not preferred. The model preference is indicated by the **bold-faced** WAIC value and chosen based on a 1σ difference for the broken power-law, or $\leq 1\sigma$ for the single power-law. The latter is chosen based on the lack of evidence for a more complex model. Fits labelled with ★ do not include the PSF shape in the model (see §6.6.1), and those labelled with ♦ use an apodizing kernel to minimize ringing in the FFT from large values at the image edges.

Band	FWHM	FWHM	$\log_{10}$ A	β	$\log_{10}$ B	β_2	$x_b = k_b^{-1}$	BP	SP
							WAIC	WAIC	
	($''$)	(pc)				(pc)	($\times 10^3$)	($\times 10^3$)	
LMC									
MIPS 24	6.5	2	13.65 ± 0.09	1.13 ± 0.03	–	1.79 ± 0.07	190 ± 25	$\mathbf{241.0 \pm 0.4}$	241.5 ± 0.4
	11.0	3	14.28 ± 0.07	0.50 ± 0.09	–	1.77 ± 0.18	439 ± 46	$\mathbf{152.9 \pm 0.3}$	153.2 ± 0.3
MIPS 70	18.7	5	10.14 ± 0.04	1.70 ± 0.02	10.96 ± 1.40	–	–	102.2 ± 0.2	$\mathbf{102.2 \pm 0.2}$
	30.0	7	10.22 ± 0.03	1.67 ± 0.02	–	–	–	67.1 ± 0.2	$\mathbf{67.1 \pm 0.2}$
PACS 100	7.1	2	11.17 ± 0.04	1.73 ± 0.02	13.13 ± 0.02	–	–	289.6 ± 0.3	$\mathbf{289.7 \pm 0.3}$
	9.0	2	11.24 ± 0.04	1.70 ± 0.02	13.02 ± 0.02	–	–	233.7 ± 0.3	$\mathbf{233.7 \pm 0.3}$
MIPS 160	38.8	9	7.84 ± 0.07	2.13 ± 0.04	9.05 ± 0.71	–	–	40.7 ± 0.2	$\mathbf{40.8 \pm 0.02}$
	64.0	16	8.07 ± 0.05	2.00 ± 0.04	–	–	–	26.3 ± 0.1	$\mathbf{26.3 \pm 0.1}$

PACS 160	11.2	3	10.07 ± 0.04	1.88 ± 0.02	11.61 ± 0.04	–	–	166.3 ± 0.3	$\mathbf{166.3 \pm 0.3}$
	14.0	3	10.14 ± 0.05	1.84 ± 0.03	11.45 ± 0.08	–	–	137.5 ± 0.3	$\mathbf{137.5 \pm 0.3}$
SPIRE 250	18.2	4	8.59 ± 0.06	1.89 ± 0.04	–	–	–	89.7 ± 0.2	$\mathbf{89.7 \pm 0.2}$
	21.0	5	8.64 ± 0.04	1.87 ± 0.03	–	–	–	79.2 ± 0.2	$\mathbf{79.2 \pm 0.02}$
SPIRE 350	25	6	7.31 ± 0.03	1.95 ± 0.02	–	–	–	56.5 ± 0.2	$\mathbf{56.5 \pm 0.2}$
	28.0	7	7.33 ± 0.03	1.95 ± 0.03	–	–	–	51.3 ± 0.2	$\mathbf{51.3 \pm 0.2}$
SPIRE 500	36.4	9	6.16 ± 0.03	1.99 ± 0.03	–	–	–	33.9 ± 0.2	$\mathbf{33.9 \pm 0.2}$
	41.0	10	6.18 ± 0.04	1.97 ± 0.03	–	–	–	30.7 ± 0.2	$\mathbf{30.7 \pm 0.2}$
				SMC					
MIPS 24	6.5	2	8.91 ± 0.02	0.78 ± 0.01	–	–	–	116.2 ± 0.2	$\mathbf{116.2 \pm 0.2}$
	11.0	3	9.09 ± 0.03	0.67 ± 0.02	–	–	–	72.7 ± 0.2	$\mathbf{72.7 \pm 0.2}$
MIPS 70★	18.7	6	7.17 ± 0.33	1.98 ± 0.14	10.56 ± 0.38	–	–	144.5 ± 0.4	$\mathbf{144.5 \pm 0.4}$
	30.0	9	7.51 ± 0.37	1.83 ± 0.17	10.67 ± 0.23	–	–	145.9 ± 0.4	$\mathbf{145.9 \pm 0.4}$
PACS 100	7.1	2	8.47 ± 0.16	1.93 ± 0.07	11.78 ± 0.01	–	–	131.1 ± 0.2	$\mathbf{131.1 \pm 0.2}$
	9.0	3	8.68 ± 0.16	1.85 ± 0.07	11.74 ± 0.01	–	–	105.5 ± 0.2	$\mathbf{105.5 \pm 0.2}$
MIPS 160♦	38.8	12	5.76 ± 0.14	2.33 ± 0.10	7.51 ± 0.08	–	–	16.4 ± 0.1	$\mathbf{16.4 \pm 0.1}$
	64.0	19	6.03 ± 0.23	2.15 ± 0.15	–	–	–	10.7 ± 0.1	$\mathbf{10.7 \pm 0.1}$
PACS 160★,♦	11.2	3	7.40 ± 0.06	2.16 ± 0.04	–	–	–	37.4 ± 0.1	$\mathbf{37.4 \pm 0.1}$
	14.0	4	7.74 ± 0.05	2.04 ± 0.03	–	–	–	38.0 ± 0.1	$\mathbf{38.0 \pm 0.1}$
SPIRE 250	18.2	5	6.24 ± 0.11	2.17 ± 0.07	8.19 ± 0.04	–	–	37.2 ± 0.2	$\mathbf{37.2 \pm 0.2}$
	21.0	6	6.27 ± 0.13	2.15 ± 0.08	8.17 ± 0.06	–	–	32.9 ± 0.2	$\mathbf{32.9 \pm 0.2}$

SPIRE 350	25	8	5.05 ± 0.12	2.28 ± 0.08	6.88 ± 0.05	–	–	23.1 ± 0.1	$\mathbf{23.1 \pm 0.1}$
	28.0	8	5.11 ± 0.11	2.24 ± 0.07	6.87 ± 0.06	–	–	20.9 ± 0.1	$\mathbf{20.9 \pm 0.1}$
SPIRE 500	36.4	11	3.88 ± 0.15	2.39 ± 0.09	5.93 ± 0.06	–	–	13.6 ± 0.1	$\mathbf{13.6 \pm 0.1}$
	41.0	12	3.98 ± 0.15	2.32 ± 0.10	5.90 ± 0.07	–	–	12.3 ± 0.1	$\mathbf{12.3 \pm 0.1}$
					M31				
MIPS 24	6.5	23	5.55 ± 0.34	1.39 ± 0.18	7.82 ± 0.68	–	–	50.9 ± 0.1	$\mathbf{50.9 \pm 0.1}$
	11.0	40	5.08 ± 0.35	1.59 ± 0.16	7.90 ± 0.50	–	–	31.7 ± 0.1	$\mathbf{31.7 \pm 0.1}$
MIPS 70	18.7	67	4.74 ± 0.42	2.11 ± 0.20	7.52 ± 0.08	–	–	16.5 ± 0.1	$\mathbf{16.5 \pm 0.1}$
	30.0	108	4.59 ± 0.44	2.16 ± 0.22	7.53 ± 0.13	–	–	10.9 ± 0.1	$\mathbf{11.0 \pm 0.1}$
PACS 100★	7.1	26	8.34 ± 0.13	1.91 ± 0.07	–	–	–	19.9 ± 0.1	$\mathbf{19.9 \pm 0.1}$
	9.0	32	8.77 ± 0.14	1.77 ± 0.07	–	–	–	20.2 ± 0.1	$\mathbf{20.2 \pm 0.1}$
MIPS 160♦	38.8	140	5.79 ± 0.23	1.87 ± 0.19	–	–	–	8.05 ± 0.08	$\mathbf{8.05 \pm 0.08}$
	64.0	231	5.76 ± 0.16	1.89 ± 0.16	–	–	–	5.21 ± 0.08	$\mathbf{5.21 \pm 0.08}$
PACS 160♦	11.2	40	7.95 ± 0.18	2.20 ± 0.08	11.57 ± 0.04	–	–	43.3 ± 0.1	$\mathbf{43.3 \pm 0.1}$
	14.0	50	8.18 ± 0.21	2.11 ± 0.09	11.44 ± 0.07	–	–	35.5 ± 0.1	$\mathbf{35.5 \pm 0.1}$
SPIRE 250♦	18.2	66	6.10 ± 0.07	2.16 ± 0.07	–	–	–	18.4 ± 0.1	$\mathbf{18.4 \pm 0.1}$
	21.0	76	6.10 ± 0.08	2.16 ± 0.07	–	–	–	16.4 ± 0.1	$\mathbf{16.4 \pm 0.1}$
SPIRE 350♦	25	90	5.01 ± 0.07	2.21 ± 0.08	–	–	–	11.5 ± 0.1	$\mathbf{11.5 \pm 0.1}$
	28.0	101	5.03 ± 0.08	2.21 ± 0.08	–	–	–	10.5 ± 0.1	$\mathbf{10.5 \pm 0.1}$
SPIRE 500♦	36.4	131	4.04 ± 0.11	2.18 ± 0.11	–	–	–	6.8 ± 0.1	$\mathbf{6.8 \pm 0.1}$
	41.0	148	4.07 ± 0.12	2.17 ± 0.12	–	–	–	6.2 ± 0.1	$\mathbf{6.2 \pm 0.1}$

M33									
MIPS 24	6.5	26	9.25 ± 0.09	0.46 ± 0.05	–	–	–	25.7 ± 0.1	$\mathbf{25.7 \pm 0.1}$
	11.0	45	9.26 ± 0.09	0.46 ± 0.05	–	–	–	15.9 ± 0.1	$\mathbf{15.9 \pm 0.1}$
MIPS 70	18.7	76	8.21 ± 0.08	0.99 ± 0.06	–	–	–	8.38 ± 0.06	$\mathbf{8.38 \pm 0.06}$
	30.0	122	8.30 ± 0.11	0.94 ± 0.08	–	–	–	5.51 ± 0.05	$\mathbf{5.51 \pm 0.05}$
PACS 100	7.1	29	9.24 ± 0.25	1.30 ± 0.13	–	–	–	26.7 ± 0.1	$\mathbf{26.7 \pm 0.1}$
	9.0	37	9.47 ± 0.10	1.20 ± 0.06	–	–	–	21.7 ± 0.1	$\mathbf{21.7 \pm 0.1}$
MIPS 160	38.8	158	6.18 ± 0.11	1.50 ± 0.11	–	–	–	3.19 ± 0.05	$\mathbf{3.19 \pm 0.05}$
	64.0	261	6.43 ± 0.28	1.30 ± 0.21	–	–	–	2.06 ± 0.03	$\mathbf{2.06 \pm 0.03}$
PACS 160	11.2	46	8.61 ± 0.09	1.31 ± 0.07	–	–	–	15.6 ± 0.1	$\mathbf{15.6 \pm 0.1}$
	14.0	57	8.52 ± 0.09	1.37 ± 0.06	–	–	–	12.9 ± 0.1	$\mathbf{12.9 \pm 0.1}$
SPIRE 250	18.2	74	6.68 ± 0.07	1.51 ± 0.06	–	–	–	7.83 ± 0.07	$\mathbf{7.83 \pm 0.07}$
	21.0	86	6.73 ± 0.08	1.47 ± 0.07	–	–	–	6.94 ± 0.07	$\mathbf{6.94 \pm 0.07}$
SPIRE 350	25	102	5.54 ± 0.07	1.42 ± 0.07	–	–	–	4.75 ± 0.06	$\mathbf{4.75 \pm 0.06}$
	28.0	114	5.60 ± 0.08	1.38 ± 0.07	–	–	–	4.32 ± 0.07	$\mathbf{4.32 \pm 0.07}$
SPIRE 500	36.4	148	4.66 ± 0.09	1.17 ± 0.08	–	–	–	2.72 ± 0.05	$\mathbf{2.72 \pm 0.05}$
	41.0	167	4.72 ± 0.10	1.15 ± 0.08	–	–	–	2.46 ± 0.05	$\mathbf{2.46 \pm 0.05}$

In a few cases, we found that the power spectra did not follow the PSF response on small scales. In each case, other systematic effects not included in the model dominate the power spectrum shape. These cases are indicated by a ♦ in Table 6.2 and a longer explanation is provided in §6.6.1. In most cases, we found that cross-hatching of telescope scans near bright emission enhances the power on scales near to or smaller than the FWHM of the PSF. It is difficult to account for this effect in our model, so we instead fit a model without the PSF response (Equation 6.1) and limit the scales fit to several times the PSF FWHM.

6.3.5 A break due to 30 Doradus in the LMC MIPS 24 μm power spectrum

We find that all power spectra in our sample are well-fit by a single power-law plus point source model with the exception of the LMC MIPS 24 μm map. For the LMC at 24 μm, Figure 6.4 shows a distinct bump in the power spectrum on scales of > 80 pc, making the broken power-law model (Equation 6.2) preferred based on the WAIC of each fit. This feature is also noted by Block et al. (2010). However, the break scale from broken power-law model is not well-constrained. We derive somewhat different values when we fit the data at their original resolution and when we fit the map after smoothing to a Gaussian kernel (Table 6.2). In this section, we identify 30 Doradus (30 Dor) as the source for this break.

A significant fraction ($\sim 30\%$) of the LMC's emission at 24 μm is solely from this giant HII region. We investigate the effect that a prominent single source has on the power spectrum by calculating the power spectrum in 3 kpc boxes with and without 30 Dor, where the edges of both boxes have the same apodizing kernel applied to suppress Gibbs ringing. Figure 6.4 shows the distinct difference in these power spectra. The box containing 30 Dor has a power spectrum that closely matches the LMC's power spectrum and contains a similar bump at the same scales. In contrast, the box without 30 Dor follows a power-law to the scale of the box.

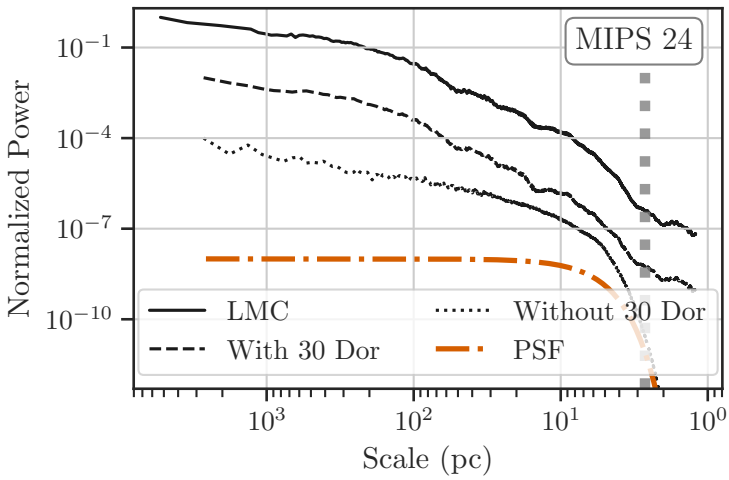

Figure 6.4: 1D power spectra of the MIPS 24 μm image for the LMC convolved to an 11″ (3 pc) Gaussian beam, and equal area regions (4×3.6 kpc^2 area) that include and exclude 30 Doradus. The power spectra are normalized by their maximum and offset by 10^2 to show the relative shapes. For the entire LMC and the region including 30 Doradus, there is a deviation and break point from a single power-law at ~ 80 pc. In the region without 30 Doradus, however, the power spectrum is well-described by a single power-law. In this case, the flux from 30 Doradus relative to the entire LMC is sufficient to produce a break in the power spectrum.

Figure 6.4 shows the power spectrum from two 3 kpc boxes, one with and one without 30 Dor. The region with 30 Dor clearly shows a similar deviation from a single power-law matching the power spectrum of the whole galaxy. The region without 30 Dor does not show this deviation is an noticably shallower, similar to the MIPS 24 μm power spectra of the other galaxies (Table 6.2). We fit a single power-law to the power spectrum without 30 Dor and find an index of 1.29 ± 0.02, steeper than the large-scale index of 1.13 ± 0.03 from the power spectrum of the entire image.

We initially fit a broken power-law (Eq. 6.2) to the power spectrum with 30 Dor, however, the break point from the fit converges to the beam size. Restricting the fit to larger scales did not lead to an improved fit. We therefore only fit a single power-law model to the power spectrum from this region. The single power-law fit gives an index of 1.76 ± 0.02, which is consistent with the index below the break point from the power spectrum of the entire galaxy. Since this fit agrees with the entire image power spectrum, the broken power-law fit likely did not converge due to the lack of data points at the largest scales within the 3 kpc region. This is similar to the change in the break point between the original and convolved LMC MIPS 24 μm power spectra. The break point is very sensitive to the data.

This example demonstrates how the power spectrum shape can be significantly altered by a small number of regions with large intensity relative to the whole image. Willett et al. (2005) find a similar result in the power spectrum of NGC 2366 in Hα, where a giant HII region causes an additional power spectrum "bump." Images with power spectrum breaks should be tested for whether the break is due to a limited number of bright discrete features.

6.3.6 Power spectrum variations within local 820 pc regions

We test whether the power spectrum varies across the LMC and SMC by computing the dust column density power spectra in local ($\sim$ 820 pc) regions. We choose 820 pc to balance between measuring local variations and retaining

sufficient information to constrain the power spectrum. The high physical resolution ($\sim$ 10 pc) allows for a large spatial range to be studied in the local power spectra. Due to the lower physical resolution in M31 and M33, we cannot access such a large spatial range in those galaxies.

Previous power spectrum studies of the Magellanic Clouds find spatial variations in the power-law index (Muller et al., 2004), potentially tracing variations in the turbulence. Furthermore, Padoan et al. (2001) identify a spatially varying scale height across the LMC studying the HI emission from 180 pc regions using the Spectral Correlation Function.

We test for a varying break point by splitting the LMC and SMC dust surface density maps into $\sim$ 820 pc overlapping squares and fit the power spectra in each region with Equation 6.4. These square regions tend to have bright emission at their edges, so we apply a Tukey apodizing kernel to remove ringing in the FFT, as explained in §6.3.1. We focus only on the Magellanic Clouds for this analysis since the resolution of 13 pc is an order of magnitude below the ~few 100 pc disc scale height of atomic gas in dwarf galaxies (Walter & Brinks, 1999).

There are large signal-to-noise (S/N) variations among the regions. The vast majority of regions have sufficient signal to measure the power law component, though a few regions in the SMC are clearly dominated by noise. When fitting the local power spectra, we found that the S/N variation leads to strong correlations between the B and C parameters from Equation 6.4. Since this analysis is primarily concerned with variation in the power law properties (index and break point), we limit the fits in this section to scales of $4 \times$ FWHM $\approx$ 52 pc and only fit the power law (i.e., we fix $B = 10^{-20}$, the lower limit on the prior). This spatial limit on the fit still captures the "break" due to the PSF response and so is adequate for this analysis.

We further attempted to model the local power spectra with a broken power-law model (Eq. 6.2) to test for the presence of strong breaks. However, we do not find any cases where the broken power-law model is strongly preferred. This is in part due to the smaller spatial range, where there is a lack of data

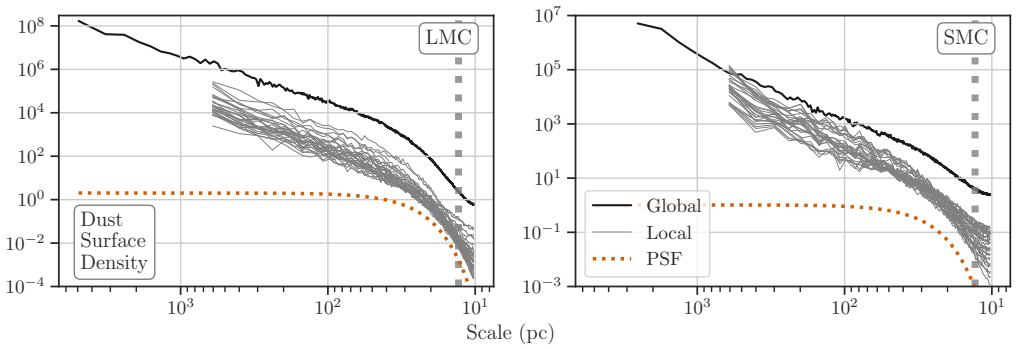

Figure 6.5: Dust surface density power spectra of the entire LMC and SMC maps (black lines) and thirty randomly-selected power spectra (gray lines) from $\sim$ 820 pc local regions. The dashed line is the PSF scaled to compare with the power spectra shape. Local power spectra steepen at small scales according to the PSF response and do not show a distinct break point on small scales. There is substantial spatial variation in the local power spectrum index, which is shown in Figure 6.6.

points on $>$ 400 pc to constrain a break. We focus the remainder of this analysis on variations in the power spectrum index.

Figure 6.5 shows LMC and SMC power spectra from the whole map and from randomly-selected 820 pc boxes. In all cases, we find that the PSF accounts for the power spectrum shape on small scales and that a break point is not required in the model. The lack of a break point on local scales rule out a varying disc scale height as an explanation for finding no break point in the whole galaxy power spectra.

The local power spectra in Figure 6.5 show variation on scales unaffected by the PSF response, implying that the fitted amplitude and power law index vary with position in both galaxies. While the amplitude is set by the total emission in the box, variations in the index imply changes in the emission morphology. Figure 6.6 shows the power spectrum index (β) overlaid on the dust surface density maps. The power spectrum indices vary from 1.31 ± 0.23 to 2.88 ± 0.22 and 0.82 ± 0.38 to 2.50 ± 0.22 in the LMC and SMC, respectively, after removing all regions near the edges of the column density maps where the noise increases. The range in local power spectrum indices is drastically larger than the uncertainty on the global power spectrum index for both galaxies

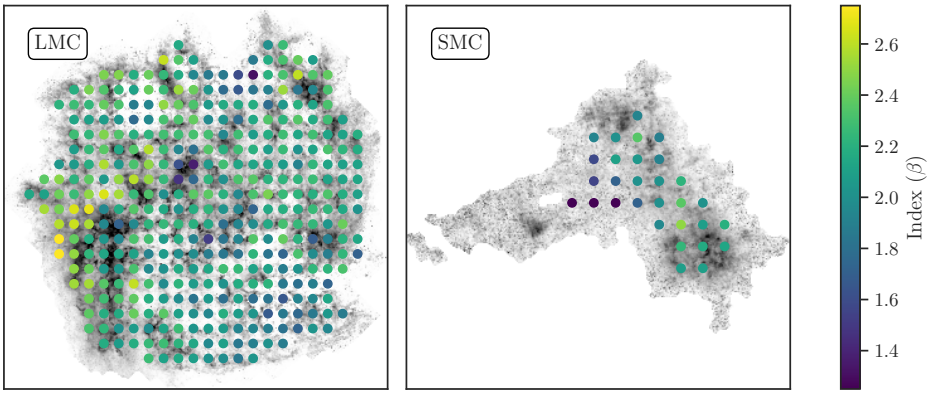

Figure 6.6: The LMC and SMC dust surface density maps overlaid with the power spectrum index measured in 820 pc regions. There is a 400 pc overlap between regions so the points shown in the figure are correlated with their neighbours. We restrict the analysis to high S/N regions, which removes regions at the edge of the LMC and a significant portion of the SMC. In both galaxies, there is significant variation in the power spectrum index (see uncertainty map in Figure 6.10) from the index over the whole galaxy (2.18 ± 0.05 and 2.47 ± 0.15 for the LMC and SMC, respectively).

(2.18 ± 0.05 and 2.47 ± 0.15; Table 6.1). This then implies that the local power spectrum variations are real and not due to noise fluctuations.

In the higher surface density regions to the north and west, the SMC power spectrum varies in index by ± 1. The steepest power spectra have indices of 2.5 and are offset from the highest surface density regions, suggesting steeper power spectra are sensitive to the gradients in surface density. In these regions, there is an excess of emission on larger scales and a deficit of small scales—due to the offset from the peak in the emission—leading to a steep power spectrum. A similar result is found by Burkhart et al. (2010), who find extremes in the skewness and kurtosis of the local HI surface density distribution near large gradients. The LMC index map shows a similar trend, with steep power spectra offset from 30 Dor and south along the Molecular Ridge along the eastern edge of the LMC.

Spatial variations in the HI power spectrum or structure function, the real-space analog, have been noted in the Magellanic Bridge (Muller et al., 2004)

and the SMC (Nestingen-Palm et al., 2017). In the latter work, Nestingen-Palm et al. (2017) split the SMC into regions based on the star formation rate. They find no change in the index with star formation rate, which is somewhat different than the variations correlated with IR brighntess that we observe. In addition to using a different tracer (they use HI, we use dust), the regions they use are significantly larger than the 820 pc boxes than from this analysis. In particular, the bright north and west regions are included in the same high SFR region, which is the area we find moderate variations in the power spectrum index.

More recently, Szotkowski et al. (2019) use the "rolling power spectrum" to explore changes in the HI power spectrum with spatial position in the SMC and LMC. They find evidence of power spectrum breaks only in the LMC, where the power spectra flatten above the break. Since we do not find this behaviour over the same regions using the dust column or IR bands, this suggests that the HI may be better coupled to stellar feedback than the total gas column traced by the dust. While Szotkowski et al. (2019) do not forward model the PSF response, the power spectra are cut-off at the beam scale. This implies that breaks on scales much larger than the beam ($3 \times \mathrm{FWHM} \sim 90$ pc) are robust against the PSF shape.

6.4 Discussion

We show that the IR and dust surface density power spectra for the LMC, SMC, M31 and M33 are well-modeled by a single power-law with point-source term, when the PSF response is accounted for. A broken power-law model is only preferred for the LMC MIPS 24 μm image and results from 30 Doradus (§6.4).

Here we discuss trends in the power spectrum properties across bands and galaxies and compare with previous studies, some of which have found evidence for breaks in the power spectrum. We also compare the dust, CO and HI power spectra in M31 and M33. We find discrepancies in the fitted index of

these three tracers. This strongly suggests that a comprehensive spatial power spectrum analysis requires a multi-tracer approach.

6.4.1 Comparisons with literature power spectra

This paper uses a large suite of archival observations, many of which have been previously analyzed using the spatial power spectrum. Here, we present an overview of spatial power spectra in the four galaxies analyzed here, including tracers not explored in this work. Where appropriate, we compare our results to these previous works, highlighting discrepancies in fit values that occur due to different fitting procedures. Accounting for differences in methodology, our power spectrum fits agree with previous analyses using similar data sets. We then compare how the power spectra from the IR and dust surface density compares to literature values at other bands. We note here that our definition of the power-law index ($\propto k^{-\beta}$) is defined so measured indices should have $\beta > 0$. Where appropriate, we alter the sign of literature values to follow this convention.

6.4.1.1 LMC

Block et al. (2010) use the LMC MIPS maps (Meixner et al., 2006) fit to a two-component power law model, where the component on large scales should have a similar index to our fits. On large scales, they find indices of 0.78 ± 0.19, 1.83 ± 0.36, and 2.15 ± 0.48 at 24, 70, and 160 μm, respectively. The latter two agree with our fitted indices, while the 24 μm is flatter due to the influence of 30 Dor (§6.4).

The power spectrum of HI in the LMC is presented in Elmegreen et al. (2001) and Elmegreen et al. (2003a), where they find that the large-scale index is around 5/3 on larger-scales (> 100 pc), as would be expected for Kolmogorov turbulence.

6.4.1.2 SMC

In the SMC, Stanimirovic et al. (2000) present power spectra from 60, and 100 μm *IRAS* bands. When fit to a single power law model, the 60 and 100 μm power spectra have indices of 2.4 ± 0.2 and 3.2 ± 0.3, respectively. Small scales that appear to be affected by the PSF response are included in the fit, leading to steeper power spectra than if only the large scales were fit; the 100 μm with its larger PSF is more affected by the decrease in power on small scales. Consistent with this difference in the power spectrum models, we find much shallower power spectra of 1.98 ± 0.14 and 1.93 ± 0.07 in the MIPS 70 and 100 μm bands, respectively. Accounting for uncertainty, the *IRAS* 60 μm index from Stanimirovic et al. (2000) is consistent with our fit to the MIPS 70 μm.

In HI, Stanimirovic et al. (1999a) find a power spectrum index of 3.04 ± 0.02 (see also Stanimirovic et al., 2000; Nestingen-Palm et al., 2017), steeper than the power spectra we find for the dust surface density (2.47 ± 0.15).

6.4.1.3 M33

Combes et al. (2012) present power spectra from a large number of tracers for M33, including the MIPS, PACS, and SPIRE bands. Like the LMC analysis by Block et al. (2010), they fit a two-component power law model to the power spectra; we compare the large-scale indices with ours. A second difference between our analyses is the fits from Combes et al. (2012) include the two smallest spatial frequency bins, which we do not include in our fits as the values in these bins are significantly larger than what would be inferred from the power law model. The indices we find tend to be shallower than those found reported by Combes et al. (2012), consistent with excluding the smallest frequency bins. The discrepancies are the largest for the MIPS bands, where our fitted indices are ~ 0.5 smaller. The discrepancy is smaller and roughly within the index uncertainty for the PACS and SPIRE maps.

Combes et al. (2012) include power spectra of the molecular and atomic

neutral ISM traced through HI and CO(2-1), respectively (Gratier et al., 2010; Druard et al., 2014). On large-scales (unaffected by the PSF shape), they find indices of 2.4 for the HI and 1.5 for CO(2-1). The latter does not flatten on large-scales. Koch et al. (2019b) shows that an excess of power on large scales is due to the clustering of GMCs in the inner few kpc. We also explore these tracers in M33 in §6.4.4 and compare the recovered indices there.

Combes et al. (2012) also include additional tracers that we do not explore, including Hα, and *GALEX* NUV and FUV. They find indices of 1.2 for the NUV and FUV bands, similar to what we find in the IR bands. The Hα index they find is 0.77, similar to the MIPS 24 μm power spectrum and consistent with an independent study by Elmegreen et al. (2003b). These similarities are expected since these are tracers of star formation that inherits some of the galactic ISM structure.

Finally, Elmegreen et al. (2003b) present power spectra of M33 in the B, V, and R bands. These optical bands are dominated by the stellar component and are not expected to match the ISM-dominated maps that we explore here. By calculating a power spectrum from 1D azimuthal strips, they find indices of 0.66 ± 0.66. The large uncertainty is due to contamination from foreground stars.

6.4.1.4 M31

There is little previous work on spatial power spectra in M31, likely due to its high inclination. We discuss the similarity of M31's power spectra to those of the Magellanic Clouds in §6.4.3.

6.4.2 Variations in the power spectrum index across bands

Large-scale variations in the shape of the dust spectral energy density, e.g., due to temperature variations, could alter the shape of the power spectrum measured at IR different bands. Figure 6.7 summarizes our fitted power spectrum indices from Tables 6.1 & 6.2 for each galaxy. The MIPS 24 μm index

is consistently shallower than those at longer wavelengths by 0.5–1.2. As we have shown in §6.3.5, bright emission regions at 24 μm, like 30 Doradus in the LMC, can alter the power spectrum shape over a large range of scales. Bright concentrated sources could result in the shallow power spectra we find in all four galaxies. The 24 μm power spectra have a similar index to studies using Hα in nearby galaxies, consistent with bright emission in both tracers arising from compact star-forming regions (Elmegreen et al., 2003b; Combes et al., 2012).

At longer wavelengths, there is less variation in the index. In the LMC and SMC, there is a mild trend of steeper power spectra at longer wavelength. The power spectrum index in M31 is relatively constant across the longer wavelength bands, while M33's power spectrum in the three SPIRE bands becomes shallower. With these differences, and the limitations of a small sample size, we find no consistent trend in the power spectrum index from 70–500 μm.

6.4.3 Variation in the power spectrum index between galaxies

The LMC, SMC, and M31 differ significantly in their large-scale morphology (Figure 6.1), yet they have a similar power spectrum index. Figure 6.7 shows that the LMC, SMC, and M31 have similar power spectra ranging from 2–2.5, excluding the MIPS 24 μm band. In all cases, the power spectra are shallower than the 8/3 predicted for 2D Kolmogorov turbulence (e.g., Elmegreen & Scalo, 2004).

The small range in power spectrum indices of the LMC, SMC, and M31 could suggest the dust emission shares a similar morphology when resolving < 167 pc scales, though this similarity is not apparent from the maps in Figure 6.1. For the LMC and SMC, where < 13 pc scales are resolved, bright features in the dust surface density maps correspond to GMCs, and, particularly for the LMC, there are large voids from supershells (Kim et al., 1999). The 167 pc resolution of the M31 dust surface density map is not sufficient to resolve

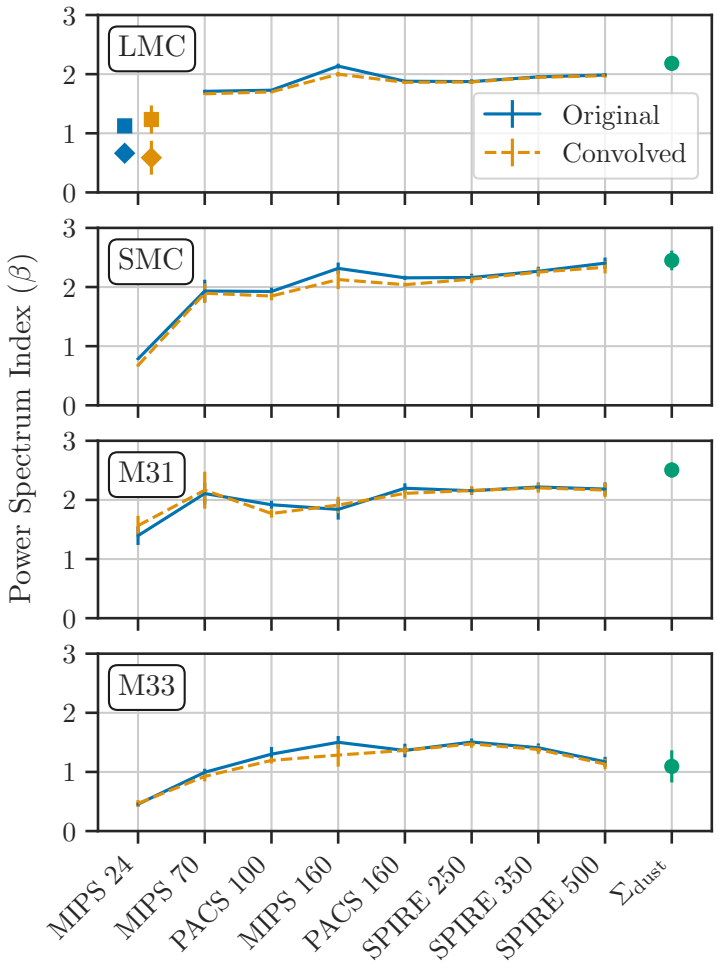

Figure 6.7: 1D power spectrum index with 1σ uncertainty error-bars for each band and the dust surface density map across all four galaxies (see Tables 6.1 & 6.2). We show the indices for the original images and the images convolved to a Gaussian, using the kernels described in Aniano et al. (2011), to highlight the agreement between the two results, which is expected if the powers-spectra are well-fit without a break point. The LMC MIPS 24 μm has both indices from the broken power-law fit shown, where squares indicate the index on small scales (above the break) and diamonds indicates the index on large-scales (below the break). The MIPS 24 μm indices are consistently flatter than the other bands, while all other have slopes of ~ 2, in broad agreement with literature values. M33 has a consistently flatter power spectrum relative to the other galaxies.

individual GMCs, and the bright regions in the map primarily highlight M31's ring structure at $R_{\rm gal} \sim 10$ kpc. This discrepancy in the spatial morphology at different scales demonstrates that different spatial morphologies can produce similar power spectra.

M33 has a consistently flatter power spectrum compared to the other three galaxies, excluding the MIPS 24 μm band, with indices ranging from 1–1.5. This difference in the power spectrum index results from M33's flocculent spiral morphology with a central enhancement of molecular gas, which differs from the distributed molecular gas morphology in the LMC and SMC, and the predominant molecular rings in M31. M33's molecular gas, and thus the highest dust surface density, is centrally concentrated into the inner few kpc; Druard et al. (2014) show that the azimuthally-averaged molecular gas surface density, from CO(2-1), is well-fit by an exponential disc with a scale length of 2.1 kpc. This is in contrast with the HI distribution, which has a roughly constant average surface density of ~ 8 M$_\odot$ pc^{-2} in the inner 8 kpc (Druard et al., 2014; Koch et al., 2018c). Approximating the disc as a uniform exponential disc, the power spectrum should have a break near the disc scale length[4] and a flat power spectrum on larger scales. While this is a plausible explanation for the CO(2-1) power spectrum from Combes et al. (2012) (see Koch et al., 2019b), this simple model does not explain the lack of a power spectrum break near the disc scale length nor the shallow power spectrum measured for the dust surface density and IR bands. This implies that the more diffuse and predominantly atomic gas plays an important role in setting the large scale power spectrum shape. We compare the HI properties to the dust in M33 in §6.4.4.

To test this hypothesis of shallow power spectra from a centralized H_2 distribution, we require other nearby face-on galaxies where similar physical scales to the M33 observations can be resolved, of which there are few. We choose to compare with the face-on spiral galaxy IC 342, which has a distance of 3.4 kpc (Tully et al., 2013). Using the Herschel bands (Kennicutt et al., 2011), the

[4]The inclination would broaden the power spectrum break.

PACS 160 μm resolution resolves $11.2'' \approx 180$ pc scales, similar to the physical resolution of the SPIRE 500 μm and dust surface density maps of M31 and M33 used here[5].

We perform the same analysis on IC 342 that is described in §6.3. We find that IC 342 has a flat power spectrum similar to M33, with typical indices around 1 (see Table 6.4), deviating from the index from the LMC, SMC, and M31. The similarity between M33 and IC 342 suggests that galaxies with a centrally concentrated H_2-distribution tend to have flatter power spectra, consistent with our expectation above for an exponential (molecular) disc plus a constant (mostly atomic) component. Grisdale et al. (2017) find a similar result in their analysis of power spectra from galaxy-scale simulations and HI data from THINGS (Walter et al., 2008). They demonstrate this dependence on the gas mass distribution by including an extended uniform gas component with different surface densities and find that this added component steepens the power spectrum on large scales. The connection with 2D turbulence on large scales is then tenuous for these galaxies.

Though the methodology and resolution of the data differ, in general the power spectrum indices that we find agree with previous work on other galaxies. Indices from power spectra measured in various optical bands range from 0.6 to 1.8 (Elmegreen et al., 2003b; Willett et al., 2005; Elmegreen et al., 2006), while those from 3.6–8.0 μm range from 0.8–2.8 (Block et al., 2009). The most studied tracer, and with the widest range in indices, is the 21-cm HI line. Previous studies find indices that range from 0.3 (Dutta et al., 2013) to 4.3 (Zhang et al., 2012), though most indices range from 1.5–3.0 (Begum et al., 2006; Dutta et al., 2008, 2009a,b; Zhang et al., 2012; Dutta et al., 2013; Dutta & Bharadwaj, 2013). Low-inclination spiral galaxies tend to have flatter power spectra in previous studies (Dutta & Bharadwaj, 2013), broadly consistent with our findings for M33 and IC 342.

Spatial power spectra within the Milky Way tend to be steeper than in

[5]There are discrepancies in the expected PSF shape for the IC 342 PACS 100 μm map and so we exclude it for this comparison.

extragalactic systems, though they around found in a similar range. The power spectra we find here are consistently shallower than power spectra from Milky Way studies. Galactic HI power spectra typically have indices from 2.5–4 (Deshpande et al., 2000; Dickey et al., 2001; Miville-Deschênes et al., 2003; Pingel et al., 2013; Martin et al., 2015; Blagrave et al., 2017; Pingel et al., 2018), with extreme values of 2.2 (Green, 1993) to 4.9 (Kalberla et al., 2017). Values from dust include 2.7 from extinction over the Perseus molecular cloud (Pingel et al., 2018), 2.9 from diffuse galactic light in optical bands (Miville-Deschênes et al., 2016), and 2.7 from Herschel SPIRE maps of the Polaris flare (Miville-Deschênes et al., 2010). The latter example is ~ 0.5 steeper than the indices we find for SPIRE maps of the LMC, SMC, and M31 (Table 6.2). We note that Milky Way studies using the spatial power spectrum do not find strong evidence for power spectrum breaks.

Even if multiple spatial distributions yield the same power spectrum index, our results still a key benchmark for simulations that aim to reproduce Local Group-like galaxies. Several recent works aim to simulate galaxies with properties closely matching the LMC, SMC, M31, M33, or the Milky Way (Combes et al., 2012; Wetzel et al., 2016; Grisdale et al., 2017; Dobbs et al., 2018; Garrison-Kimmel et al., 2019) with many producing "synthetic" observations to compare with properties found in the actual observations (e.g., Dobbs et al., 2019), a key step for directly comparing simulations and observations (Haworth et al., 2018). For any simulation the produces dust maps or synthetic IR observations, matching our measured power spectrum represents an important check.

6.4.4 Comparisons with HI and CO power spectra

The dust surface density closely traces the total neutral gas surface density, related only through the dust-to-gas ratio. In contrast, 21-cm HI or CO emission traces only a particular phase of the neutral ISM. This makes the dust surface density a potentially useful tool to compare how the power spectrum changes in the atomic and molecular ISM phases.

We found in the previous section (§6.4.3) that the large-scale galactic distribution affects the power spectrum shape. In this section, we compare the dust power spectrum with those from HI and CO, which cleanly separate the atomic and molecular components of the neutral ISM, to measure how the power spectrum changes in the different neutral ISM components.

Due to the different conditions in each neutral ISM phase, the turbulent properties in the HI and H_2 may differ. Romeo et al. (2010) demonstrate how the transsonic or subsonic conditions in warm HI alter the stability conditions relative to the supersonic turbulence from H_2 in molecular clouds. In particular, if warm HI traces a transsonic or subsonic density field, density fluctuations from the mean will be small. This means that the 1D power spectrum from the HI surface density should be flatter than the H_2 (traced by CO). These differences in the turbulent properties of the atomic and molecular ISM have important consequences for setting the local stability of the galactic disc (Hoffmann & Romeo, 2012; Romeo & Agertz, 2014).

Furthermore, dust may be a passive tracer in ISM turbulence, meaning that it may not actively contribute to the turbulence and may have different properties from the gas (e.g., Goldman, 2000). Dust may further be subject to additional drag instabilities (Hopkins, 2014; Hopkins & Squire, 2018). The comparisons between the dust, HI and CO we show here may results from any of these sources. We focus our analysis on looking for consistent differences between the power spectra of these different tracers.

Previous work on the SMC and LMC shows that the HI is steeper than the dust power spectra we find here. In the SMC, the HI power spectrum is well-described by a single power-law (over the entire galaxy) with an index of 3.04 ± 0.02, and between 5/3 and 8/3 (Elmegreen et al., 2001), respectively. Combes et al. (2012) similarly find a steeper HI power spectrum (2.4 ± 0.2) relative to the IR bands and CO in M33.

To further this comparison, we calculate the HI and CO integrated intensity power spectra of M33 and M31[6]. We use the Karl G. Jansky Very Large Array

[6]We assume optically-thin HI emission, but see Braun (2012).

(VLA) M33 HI map from Koch et al. (2018c) and IRAM 30-m CO(2-1) map from Druard et al. (2014), both of which are convolved to the 167 pc (46.3″) resolution of the dust surface density map. For M31, we use a new VLA HI map (Koch et al. in prep) with 58″ resolution and the CO(1-0) map from Nieten et al. (2006). The M31 HI map has a lower resolution than the dust and CO maps, so we convolve these maps to match the HI. This gives a resolution of 201 pc (58″).

We fit the HI and CO power spectra to Equation 6.3. Table 6.3 presents the fit parameters for the HI, CO, and dust power spectra. The power spectra are all well fit by a single power-law model (Eq. 6.1). Figure 6.8 shows the consistent differences between the HI, dust, and CO power spectra, where HI is the steepest and CO is the shallowest. Our fits to the M33 HI and CO are flatter than those found by Combes et al. (2012), though they use a resolution of 48 pc for both and do not account for the PSF shape; fitting the HI power spectrum at its native 80 pc resolution gives an index of 2.30 ± 0.13, consistent with Combes et al. (2012). Combes et al. (2012) also include the smallest frequency bins from the power spectrum, which strongly deviate above the power-law relation. We exclude these data in our analysis.

A steeper HI power spectrum relative to the dust implies a lack of power on small scales in the HI. In terms of the molecular and atomic column density, there are four sources for this discrepancy: (1) the H_2 distribution differs from the HI on galactic scales, (2) saturated HI on small scales, (3) optically-thick HI dominates on small scales, and (4) radial changes in the dust-to-gas ratio due to radial metallicity gradient. The first three will remove structure on small-scales, smoothing the spatial distribution of HI relative to the total gas distribution traced by dust. The first two sources arise from the conversion of HI to H_2 (e.g., Bigiel et al., 2008; Krumholz, 2013; Sternberg et al., 2014). In M33, the H_2 distribution is centrally-concentrated in the galaxy and primarily from GMC scale emission (Rosolowsky et al., 2003). Koch et al. (2019b) show that the distribution of GMCs in M33 can provide an excess in power on scales up to ~ 2 kpc, and can therefore affect the CO and dust power spectra shape

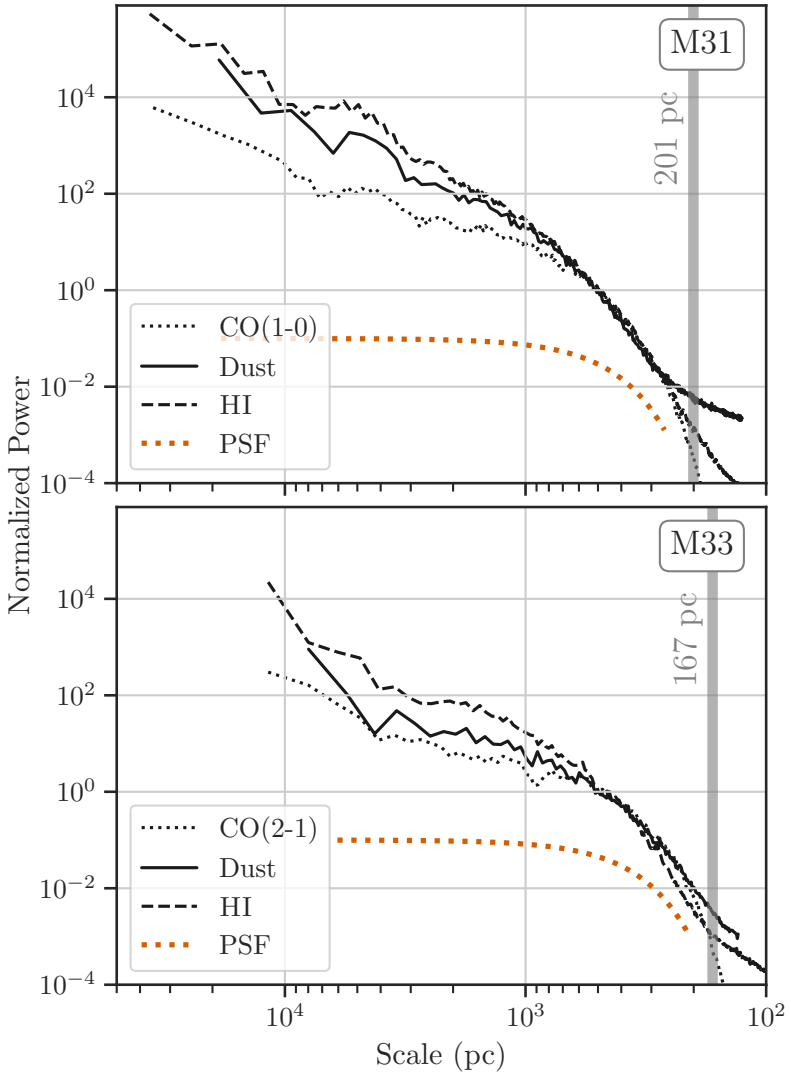

Figure 6.8: 1D power spectra from CO, dust and HI in M31 (201 pc;58″) and M33 (167 pc; 41″) convolved to a common beam size and normalized to the power on 500 pc scales. The power spectrum of the beam is shown with the orange dotted curve. All of the power spectra are well fit by a single power law attenuated by the beam. Indices are given in Table 6.3. In both galaxies, the CO power spectrum is the flattest and the HI is the steepest, with the dust somewhere in between. In M33, the centralized H_2 emission significantly affects the dust power spectrum shape by comparison to the CO power spectrum. This comparison shows that the power spectrum shape is affected by the large-scale structure of the emission on all measurable scales, making connections to 2D turbulence tenuous.

on similar scales. The final fourth point will tend to flatten the dust power spectrum on large scales, as the dust abundance decreases with metallicity.

The H_2 distribution in M31 is dominated by the ring-structures and does not show a strong concentration in the inner disc. However, with bright CO still clustered into a large-galactic structure (i.e., the rings), the morphology may still provide excess power on the scale of the rings, affecting the large-scale power spectrum.

We test the influence of the H_2 distribution on the dust surface density power spectrum by combining the HI and CO maps to get the neutral gas surface density power spectrum. We assume constant Milky-Way α_{CO} factor of 4.8 $M_\odot\,pc^{-2}$ and 6.7 $M_\odot\,pc^{-2}$ for the 1-0 and 2-1 lines, respectively, and optically-thin HI to convert the integrated intensities to the molecular and atomic surface densities. All of the neutral gas surface density power spectra are well fit by a single power-law (Eq. 6.1) with indices between the HI and CO (Table 6.3). We expect this result from the relative differences in the power spectra shown in Figure 6.8.

The power-law indices from the neutral gas power spectra are similar to the dust index, though the index is highly sensitive to the choice of α_{CO}, as shown in Table 6.3. For example, doubling α_{CO} changes the index by 0.23 and 0.30 in M31 and M33, respectively. We note that this is an overly-simplified treatment of α_{CO} that only demonstrates the power spectrum's sensitivity to these variations. The variations in the dust-to-gas ratio, tied primarily to metallicity (e.g., Bolatto et al., 2013a), will show a similar sensitivity to changes in α_{CO}. Both quantities also vary within galaxies (e.g., Sandstrom et al., 2013), and a more sophisticated handling for these variations may explain the moderate discrepancies in the dust and gas power spectra. Though the turbulent properties may change in the atomic and molecular ISM (Romeo et al., 2010) and dust may have different dynamics than the gas (Hopkins, 2014), these variations in the netural gas power spectrum index suggest that conversion factors alone can explain the difference between dust and the gas phases.

In all cases, we do not correct for optically-thick HI when computing the

Table 6.3: Power spectrum indices of HI, CO, dust, and total neutral gas surface density in M31 and M33, with different assumed values of $\alpha_{\rm CO}$. In all cases, the power spectra are well-fit by a single power-law (Figure 6.8). The M31 maps are convolved to the 197 pc (57.5″) beam from the HI data (Koch et al. in prep.) and the M33 maps are convolved to the 167 pc (41″) beam of the dust surface density map. Despite differences in the M31 and M33 indices, there is a consistent trend for steeper atomic gas power spectra and flatter molecular gas power spectra, with the dust power spectra somewhere in the middle. The neutral gas surface density maps have power spectra similar to the dust, but are sensitive to the choice of $\alpha_{\rm CO}$. Differences between power spectra of different tracers emphasises how the large-scale structure affects the power spectrum shape on all measurable scales. [a] CO(1-0) (Nieten et al., 2006). [b] CO(2-1) (Gardan et al., 2007; Gratier et al., 2010; Druard et al., 2014).

	M31	M33
	201 pc scales	167 pc scales
$I_{\rm HI}$	2.66 ± 0.12	2.18 ± 0.10
$I_{\rm CO}$	1.59 ± 0.08[a]	0.91 ± 0.14[b]
$\Sigma_{\rm dust}$	2.44 ± 0.15	1.11 ± 0.14
$\Sigma_{\rm HI} + \alpha_{\rm CO} I_{\rm CO}$	2.34 ± 0.11	1.66 ± 0.12
	$\alpha_{\rm CO10} = 4.8$ $\rm M_\odot\,pc^{-2}/\,K\,km\,s^{-1}$	$\alpha_{\rm CO21} = 6.7$ $\rm M_\odot\,pc^{-2}/\,K\,km\,s^{-1}$
$\Sigma_{\rm HI} + \alpha_{\rm CO} I_{\rm CO}$	2.11 ± 0.10	1.36 ± 0.07
	$\alpha_{\rm CO10} = 9.6$ $\rm M_\odot\,pc^{-2}/\,K\,km\,s^{-1}$	$\alpha_{\rm CO21} = 13.4$ $\rm M_\odot\,pc^{-2}/\,K\,km\,s^{-1}$

HI surface density as, for M33, Koch et al. (2018c) do not find evidence for flattened HI velocity spectra indicative of bright optically-thick HI emission (Braun, 1997; Braun et al., 2009), though Braun (2012) argue there is a 30% correction factor to the atomic gas mass in M31 and M33. If optically-thick HI emission contributes to the lack of power in the HI power spectrum at small scales, we expect it to arise from spatial regions < 167 pc in size. We note, however, that Nestingen-Palm et al. (2017) find no change in the SMC HI power spectrum when correcting for optically-thick HI (Stanimirovic et al., 1999a), and Pingel et al. (2018) also find no change in the index for HI of the Perseus molecular cloud.

These results demonstrate that the power spectrum of the dust, and IR bands dominated by dust emission, is strongly influenced by the location of H_2, in this case traced by CO, leading to a significantly different slope relative to only the atomic component traced by the HI. These differences in the power spectra of the atomic and molecular power spectra, and between different galaxies, strongly suggests that the properties of the large-scale power spectrum are dominated by the galactic distribution of the tracer. This makes comparisons to 2D turbulent properties on $>$ kpc scales tenuous without accounting for these differences.

6.4.5 Power spectrum breaks are not ubiquitous

A key result from previous studies using the spatial power spectrum of dust and gas in nearby galaxies is a break in the power spectrum on scales similar to the expected disc scale height, which is otherwise difficult to constrain from observations of low or moderately inclined galaxies. Here we find that previous claims of a power spectrum break for M33 (Combes et al., 2012) and the LMC (Block et al., 2010) can be entirely accounted for by the shape of the PSF. The only exception we find is the 24 μm LMC map, where 30 Doradus is sufficiently bright to cause an excess on ~ 200 pc scales (§6.4). In §6.3.6, we find that splitting the map does not make a power spectrum break evident, which may occur if the disc scale height changes substantially over the maps,

thereby smearing out a single clear break-point. With the lack of a power spectrum break, the spatial power spectrum does not constrain the disc scale height.

Most studies that explore the power spectrum break find that it is located on scales a few times the PSF FWHM (Elmegreen et al., 2001; Dutta et al., 2009a; Block et al., 2010; Combes et al., 2012). Based on our results, this suggests that the scale of the break could be influenced by the PSF shape of the observation. A similar suggestion is made by Grisdale et al. (2017) based on HI power spectra of 6 galaxies from THINGS (Walter et al., 2008). There are some exceptions where breaks are found on scales many times the PSF FWHM, however, these tend to be measured at either 24 μm (Block et al., 2010) or the Hα line (Willett et al., 2005; Combes et al., 2012) where a small number of giant HII regions provide a significant fraction of the total flux of the galaxy.

Szotkowski et al. (2019) have recently found power spectrum breaks in the HI in the LMC when measured over local scales. Several areas in their analysis show a break on scales significantly larger than the beam size (30 pc), with variations around near regions with strong stellar feedback (e.g., giant HII regions). This is in apparent disagreement with the lack of break points we find in the local LMC dust surface density power spectra (§6.3.6), yet we show in §6.4.4 that dust and HI power spectra are different when measured over the entirety of M31 and M33. These differences could indicate that the HI, which saturates above some surface density (e.g., Krumholz, 2013), better traces the influence of stellar feedback on the surrounding atomic ISM. There is significant precedent for feedback affecting galaxy-scale spatial power spectra from numerical studies (Bournaud et al., 2010; Pilkington et al., 2011; Combes et al., 2012; Grisdale et al., 2017), including those that do not find a power spectrum break (Renaud et al., 2013).

These results point to multiple factors that influence the power spectrum shape and the presence of a break, rather than a ubiquitous break related to the disc scale height. These factors include the large-scale distribution of gas

in the galaxy, especially the presence of high column density, H_2-dominated regions (§6.4.3), and the gas tracer used (§6.4.4). The relative influence of each factor can be explored using local power spectra (§6.3.6) of multiple tracers. We plan to explore this in future work.

We also note that power spectrum breaks are not commonly found on smaller scales within nearby Milky Way molecular clouds (< 20 pc). Power spectrum studies of the Perseus molecular cloud include scales where stellar feedback provides sufficient energy to drive turbulence (Padoan et al., 2009; Arce et al., 2011) but do not find a power spectrum break, despite results from alternative methods, like the probability distribution function (PDF), that suggest small scale driving should be dominant (Bialy et al., 2017).

The multiple factors influencing the power spectrum shape do not rule out a break at the disc scale height, tracing the transition from 3D to 2D turbulence. It may be possible to account for each of these factors with a more sophisticated model to search the uniform presence of a break. However, the current quality of data does not support the need for a more complex model.

Finally, we note that this analysis is limited to information from the projected density field of these galaxies. When using a spectral-line, the line-of-sight velocity offers additional information useful for this type of analysis. Velocity information be incorporated into the power spectrum or structure function by using the line-of-sight velocity centroid (e.g., Bertram et al., 2015) or different on power spectra of the whole spectral-line data cube, such as the Velocity Channel Analysis or Velocity Coordinate Spectrum (e.g., Stanimirović & Lazarian, 2001; Lazarian & Pogosyan, 2006; Chepurnov et al., 2015). Alternatively, empirically-based methods like the Spectral Correlation Function (SCF; Rosolowsky et al., 1999) provide a complementary measure of structure with spatial scale. Padoan et al. (2001) found deviations in the SCF relation applied to HI data of the LMC and attributed the deviations to the LMC disc scale height. In future work, we will utilize velocity information in our analysis of HI of M33 (Koch et al., 2018c) and M31 (Koch et al. in prep.).

6.5 Summary

We present a unified analysis of the 1D power spectra of mid- to far-IR emission and the dust surface density in the LMC, SMC, M31, and M33. A key result of our work is that previous claims of a power spectrum break can be explained by the instrumental PSF response and are not a measurement of the disc scale height. This result has important consequences for simulations of Local Group-like galaxies, which have also found break points in power spectra (e.g., Bournaud et al., 2010; Combes et al., 2012; Grisdale et al., 2017).

1. We model the PSF response on the 1D spatial power spectra and find that the power spectra of all the galaxies is well-modeled by a single power-law plus point source components. We demonstrate that previous studies that find a break point in the power spectra is entirely due to the PSF response. We also note that, comparing to both Galactic and extra-galactic power spectrum studies, there are few cases a power spectrum break from intensity maps is unambiguously found over all spatial scales across several wavebands.

2. M31, the LMC, and SMC have similar power spectra indices ranging from 2 to 2.5. The indices in these three galaxies are broadly consistent in the dust surface density and individual infrared bands despite the difference in their morphology. This similarity demonstrates that different spatial morphologies can produce similar power spectra, showing the need to carefully consider multiple sources that can alter the power spectrum shape.

3. Compared to the other three galaxies, M33 has a significantly flatter power spectrum with an index of ~ 1.3. We calculate the power spectrum of IC 342, a nearby face-on spiral, in the *Herschel* bands and find a similarly flat power-law index. This similarity suggests that spiral galaxies with a central H_2 concentration tend to have flatter power spectra, which can be explained by the shape of an exponential (molecular)

disc with a flat (mostly atomic) component, rather than large-scale 2D turbulence.

4. We compare the dust, HI, and CO power spectra of M31 and M33 at a common scale. The HI and CO power spectra are well-fit by a single power-law. We find a consistent trend in the indices, with HI being the steepest and CO being the shallowest. This is consistent with HI having more structure on large scales and CO having more structure on small scales. We create total neutral gas surface density maps by combining the HI and CO, and find their power spectrum index is intermediate between the HI and CO, and is similar to the dust. The neutral gas power spectrum is sensitive to $\alpha_{\rm CO}$, leading to variations that can account for the difference in the power spectra index of the dust. The dust and gas are further related by the dust-to-gas ratio, which is known to vary on large scales (Sandstrom et al., 2013). This result provides further evidence that the power spectrum is sensitive to the large-scale distribution of a tracer, making it difficult to connect to 2D turbulence without accounting for this effect.

5. We compute the dust surface density power spectra over local ($\sim$ 820 pc) regions within the LMC and SMC and find they are also well-fit without a power spectrum break. This result rules out local variations in the disc scale height as an explanation for the lack of a break measured from the whole galaxy's power spectrum. The difference between the dust and HI power spectra that we find, and the recent identification of local HI breaks in the LMC by Szotkowski et al. (2019) shows that the HI may better trace feedback relative to the total neutral gas column traced by the dust.

6. The local power spectra in the LMC and SMC show substantial variation across the galaxies. We find that steeper power spectra occur near large intensity gradients, similar to what Burkhart et al. (2010) find using the skewness and kurtosis of the HI column density in the SMC.

7. Simulations of Local Group-like galaxies reflect some of the results we find here. Grisdale et al. (2017) find that the column density power spectrum on ∼few kpc scales is sensitive to the mass distribution in galaxies but is insensitive to other effects like stellar feedback. They also find that, when comparing to HI intensity power spectra of nearby galaxies, the small-scale power spectra are dominated by the PSF shape. These results demonstrate the need to produce synthetic observations when comparing power spectra of simulations and observations (Haworth et al., 2018). Our results provide a benchmark for comparing observations and simulations of Local Group-like galaxies.

Our results demonstrate that power spectra are sensitive to systematic effects that significantly effect how they are modeled and interpreted. Where applicable, we recommend forward-modelling the instrument PSF when fitting a model to power spectra.

Previous work has focused on the source of the power spectrum break, using it as a measure of the disc scale height to constrain galactic structure. With these results, we show that an alternative explanation is required to understand the ubiquity of galactic power spectra on scales well below the disc scale height. Further work requires investigating whether a break is measurable from the velocity field from spectral lines (e.g., Padoan et al., 2001).

Scripts to reproduce the analysis are available at
github.com/e-koch/DustyPowerSpectra[7].

Acknowledgments

We thank the referee, Alessandro Romeo, for helpful comments that improved manuscript. EWK acknowledges helpful discussions with Bruce Elmegreen and Snežana Stanimirović. EWK is supported by a Canada Graduate Scholarship and Michael Smith Foreign Study Supplement from the Natural Sciences and Engineering Research Council of Canada. EWK and EWR are supported

[7]Code DOI: https://doi.org/10.5281/zenodo.3583220

by a Discovery Grant from Natural Sciences and Engineering Research Council of Canada (RGPIN-2017-03987). The work of DU and AKL is partially supported by the National Science Foundation under Grants No. 1615105, 1615109, and 1653300. *Herschel* is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA. This work is based, in part, on observations made with the *Spitzer Space Telescope*, which is operated by the Jet Propulsion Laboratory, California Institute of Technology under a contract with NASA. The National Radio Astronomy Observatory and the Green Bank Observatory are facilities of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

Code Bibliography: astropy (Astropy Collaboration et al., 2013, 2018) — radio-astro-tools (spectral-cube, radio-beam; radio-astro-tools.github.io) — matplotlib (Hunter, 2007) — seaborn (Waskom et al., 2017) — corner (Foreman-Mackey, 2016) — pymc3 (Salvatier et al., 2016) — numpy & scipy (Oliphant, 2006)

6.6 Appendix

6.6.1 Additional systematics affecting fits

Some of the fits to the power spectra presented in Table 6.2 required an altered model or additional step applied to the data to find a valid fit. These cases are shown in Table 6.1 with an additional symbol. We provide further details of these special cases here.

- SMC MIPS 70 μm – There is cross-hatching in the map on scales below the PSF's FWHM. Thus, the power spectrum does not follow the PSF on small scales. We do not use the PSF in the model and restrict the fitting to scales above 25 pc.

- SMC MIPS 160 μm – There are large, noisy values at the edge of these

images that cause ringing in the FFT. We apply a Tukey apodizing kernel to taper these values in the power spectrum.

- SMC PACS 160 μm – We do not use the PSF for the fit for the same reason as the MIPS 70 μm. The fitting is restricted to scales larger than 10 pc. We also apply a Tukey apodizing kernel to this map, similar to the MIPS 160 μm map.

- M31 PACS 100 μm – We do not use the PSF in the fit as the small-scales in the power spectrum are dominated by the scan pattern. We restrict the fitting to scales larger than 150 pc.

- M31 MIPS 160 μm, PACS 160 μm, SPIRE 250, 350, 500 μm – We apply a Tukey apodizing kernel to taper large values at the map edge, avoiding ringing in the FFT.

We also test whether bright foreground point sources affect the power spectrum shape or contribute to the additional systematics described above. Using the MIPS point-source subtracted maps from the SAGE (Meixner et al., 2006) and SAGE-SMC (Gordon et al., 2006, 2011) data releases, we find no difference in the power spectrum index. We do not expect point source contamination to affect our results.

6.6.2 Deprojection does not change the large-scale power spectrum

Some previous works presenting extragalactic spatial power spectra have deprojected the image into the galaxy frame prior to computing the power spectrum (Block et al., 2010; Combes et al., 2012). We show an example of deprojection using the SPIRE 500 μm maps of M31 and M33 to demonstrate that deprojection does not significantly affect the large-scales of the power spectrum. We use these maps for this example because the SPIRE 500 μm map have well-behaved PSFs with near-uniform noise, which allows for the deprojected PSF to be well-described by an elliptical Gaussian.

We use ths position angle and inclination for M31 (i = 77.7;PA=38) and M33 (i = 55.1;PA=201) from Corbelli et al. (2010) and Koch et al. (2018c), respectively. We deproject each map in three steps: (1) the galaxy centre is shifted to the central image pixel, (2) the image is rotated to have the semi-major axis aligned along the y-axis of the image, and (3) the image is warped along the minor axis to match the major axis. Each of these steps are applied with interpolation methods in the SCIPY.NDIMAGE package[8]. For a Gaussian PSF, step (3) can be applied to the PSF to approximate the PSF in the deprojected frame; thus the deprojected images have a larger effective beam shape set by the inclination.

Figure 6.9 shows the original and deprojected power spectra for M31 and M33. In both cases, the shape of the power spectrum on large scales (small frequencies) is not affected. The fitted indices to the deprojected power spectra are 2.17 ± 0.12 and 1.08 ± 0.08 for M31 and M33, respectively; these indices are consistent with the original power spectra indices of 2.17 ± 0.12 and 1.15 ± 0.08 (Table 6.2).

The power spectrum index after deprojection is consistent with Grisdale et al. (2017), who compare power spectra of simulated galaxy discs at $i = 0 and i=40$ and find that only scales of order the disc diameter are affected by inclination. Block et al. (2010) also note no significant difference in indices from deprojecting the MIPS maps of the LMC.

Finally, we note the difficulty in simultaneously modelling for instrumentation effects and the projection effects from the observed frame of the galaxy. Deprojection of a non-axisymmetric PSF is complicated by the rotation step. Fully modelling for both of these effects would require forward-modelling the 2D power spectrum through a deprojection step, followed by applying the PSF. The computational requirements to model the 2D power spectrum for large images would be prohibitive in practice.

[8] docs.scipy.org/doc/scipy/reference/ndimage.html

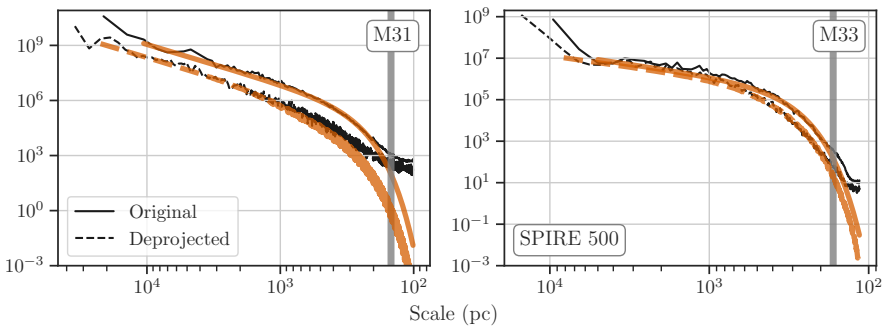

Figure 6.9: Comparison of the M31 and M33 SPIRE 500 μm power spectra with and without deprojecting the maps. The translucent orange lines show the best fit to Equation 6.3 and the vertical lines are the FWHM of the original image. The warping step for deprojection effectively adds a "white noise" due to the interpolation that is evident on scales below the FWHM. This is more evident for M31 due to its large inclination. We limit the fit to scales larger than 3 $\times$ FWHM to avoid these regions. For M31, the deprojected power spectrum index is 2.20 ± 0.19, consistent with the original power spectrum index of 2.17 ± 0.12. Similarly, the M33 deprojected power spectrum index 1.08 ± 0.08 is consistent with the original index of 1.15 ± 0.08. Thus, we expect that deprojection will not significantly alter the fitted indices in this paper.

6.6.3 Local LMC and SMC power spectrum uncertainty

Figure 6.10 shows the power spectrum index uncertainties for the values shown in Figure 6.6. The uncertainties are small relative to the change in the indices, indicating that spatial variation in Figure 6.6 represent real variations in the power spectrum shape.

6.6.4 Dust power spectra of IC342

In §6.4.3, we perform a similar analysis to §6.3 on the *Herschel* maps of IC342 to compare with M33's power spectra. Despite being 4$\times$ the distance of M33, IC342 is one of the nearest face-on spiral galaxies. Critically for this comparison, the molecular gas fraction increases towards the inner disc, similar to M33 and other spiral galaxies but unlike the LMC, SMC, and M31.

Similar to the Local Group galaxies, we calculate the power spectrum centered on IC342 and exclude regions far from the galaxy. This step is more

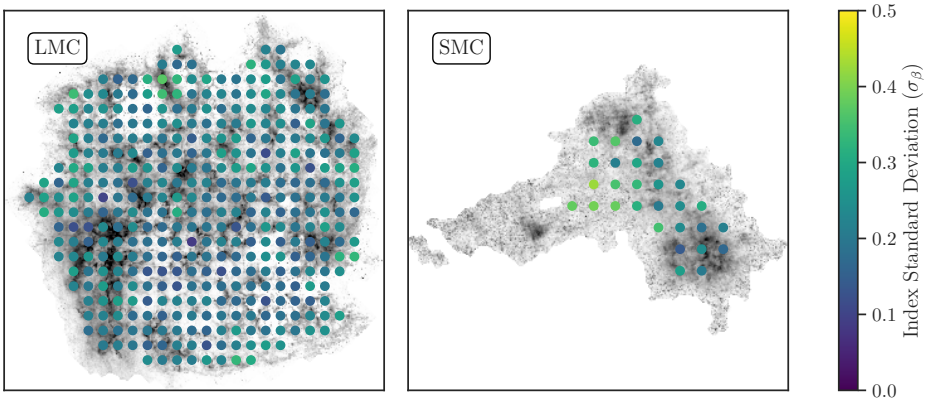

Figure 6.10: The power spectrum index uncertainties in 820 pc regions overlaid on the LMC and SMC dust surface density. The power spectrum indices are shown in Figure 6.6.

critical for IC 342 than the other galaxies because of its low Galactic latitude; Galactic emission in the *Herschel* bands is substantial over most of the maps. We do not find substantial contamination from Galactic emission over the regions used for the power spectrum, which would be indicated by emission near the edges of the region causing the Gibbs phenomenon in the 2D power spectrum.

Table 6.4 provides fit results to Equation 6.3 for the *Herschel* bands, excluding the PACS 100 μm map due to variations from the expected PSF shape that appear to be systematics (§6.6.1). In §6.4.3, we compare these results with the other galaxies.

Table 6.4: Fit parameters to Equation 6.3 for the *Herschel* bands of IC 342. Uncertainties are the 1-σ interval estimated from the MCMC samples. None of the fits constrain the unresolved point-source term B. We exclude the PACS 100 μm map as the small scale structure does not follow the expected shape of the PSF (see §6.6.1). IC 342 has a flat power spectrum similar to M33, suggesting that a flatter power spectrum is associated with typical spirals where the molecular gas fraction increases in the inner disc, unlike the LMC, SMC, and M31.

Band	Resolution ($''$)	Phys. Resolution (pc)	$\log_{10} A$	β
PACS 160	11.2	181	2.49 ± 0.04	0.99 ± 0.05
SPIRE 250	18.2	295	8.03 ± 0.07	1.04 ± 0.06
SPIRE 350	25	405	6.61 ± 0.10	0.98 ± 0.10
SPIRE 500	36.4	589	5.30 ± 0.16	0.89 ± 0.17

Chapter 7

Conclusions & Future Work

7.1 Summary of Results

In this thesis, I explore three topics on the neutral ISM that bridge between extragalactic and Galactic scales. Connecting these scales is crucial for understanding the baryonic cycle and secular galaxy evolution through the star formation process. Furthermore, measuring the ISM properties across these missing scales will bridge the discrepant "bottom-up" and "top-down" views of star formation and the ISM (Krumholz, 2014).

To explore both large and small physical scales in the neutral ISM, I use new HI VLA observations of the Local Group galaxies M31 and M33. As the nearest major star-forming disc galaxies to the Milky Way, M31 and M33 are the ideal targets to combine an external galaxy view with the proximity for current instrumentation to resolve down to $\sim$ 10s pc scales. These observations are the first part of a larger and on-going L-band Local Group VLA Legacy survey.

By combining these new VLA HI data with existing observations of CO (tracing H_2) and infrared thermal dust emission, I address three questions linking the neutral ISM properties in M31, M33, and the Magellanic Clouds:

How do models of the 21-cm HI line shape affect our interpretation of the neutral ISM? Approximate methods used to measure the HI spectral properties (e.g., line width) bias measurements by up to $\sim$ 50%, affecting

derived physical properties. Further, spectral models based on physical assumptions about the atomic ISM (e.g., CNM/WNM phases) must be tested whether they accurately describe the observations. Without this testing, basic measurements like the atomic ISM mass are affected by 30–120% from instrinsic model assumptions. Accurate spectral modeling is the key to progress studies of the atomic ISM.

In Chapter 3, I introduce our M33 HI VLA observations at 80 pc scales and techniques for imaging massive inteferometric spectral-line observations (Chapter 2). With a 0.21 km s^{-1} spectral resolution, these observations can resolve components from the CNM ($>$ 100 K thermal line widths). I find that nearly all HI spectra in M33 are *not* well-described by a single Gaussian component. Most spectra show multiple Gaussian components and/or the presence of extended line wings. My co-authors and I explore this complexity using common measurement methods found in the HI extragalactic literature and conclude that:

1. **Asymmetric HI line wings trace a rotationally-lagging thick HI disc in M33.** We find that HI line wings are asymmetrically skewed towards the systemic velocity, indicating that the line wings trace a rotationally-lagging HI disc, previously detected at $> 1'$ scales (Kam et al., 2017). The rotational lag results from a large disc scale height of the emitting gas and likely traces a combination of outflows due to stellar feedback and cold HI accretion onto M33.

2. **Extragalactic HI line widths measured with approximation methods are biased and tend to be overestimated.** We compare typical methods of measuring the HI line width in extragalactic HI observations and find they produce $\sim 50\%$ discrepancies in the measured line widths (7 km s^{-1} versus 12 km s^{-1}). This discrepancy arises from two sources. First, the line width from the second moment of a spectrum is strongly biased by multiple Gaussian components and asymmetric line wings. We find an average line width of $\sim$ 12 km s^{-1}, much larger than the WNM

thermal line width of $\sim$ 6 km s^{-1} (for 6000 K gas). Second, HI line widths from spectral stacking are affected by the choice of line centre. Assuming a circular rotation model or the centroid velocity as the line centre produces $\sigma_{\rm HI} \sim 10$ km s^{-1} line widths, while stacking based on the velocity at peak intensity give an average line width of $\sigma_{\rm HI} \sim 7$ km s^{-1}. Similarly, the fraction of emission in the stacked spectrum line wings varies from 9% to 27%, demonstrating that previous work identifying the line wings as a galaxy's turbulent WNM emission is unlikely to measure the true WNM fraction.

3. **Supernovae can drive observed levels of turbulence but is not strongly constrained due to uncertainty in the coupling of SN energy into the ISM.** We compare the range of line widths we measure (7–12 km s^{-1}) to the energy injection rates of SN and MRI. Despite the factor of $\sim$ 2 in line widths, the SN energy injection reasonably matches with an assumed efficiency of 10%, while MRI is unlikely to be a dominant source of turbulent motion anywhere in M33. Our result is consistent with previous work (Tamburro et al., 2009) but suggests that stronger constraints on the energy coupling efficiency into the ISM are needed for more detailed comparisons (Mac Low & Klessen, 2004b).

4. **Cold HI produces the sharp peak in stacked spectra, but the universal stacked spectrum shape is due to stacking itself, not common properties in the HI between nearby galaxies.** Several HI spectral stacking studies find a "universal" shape which some studies suggests reflects common HI properties in neaby galaxies (Petric & Rupen, 2007). Using a Gaussian Mixture Model, we show that the universal shape results from the stacking technique itself, and is therefore not strongly related to the actual HI properties. From the mixture model, however, we demonstrate that the narrow peak of the stacked HI spectra in M33 can *only* be produced from HI emission with line widths < 5 km s^{-1} located at the velocity at peak intensity. Our HI ob-

servations are therefore detecting cold HI emission enabled by the fine spectral resolution.

5. **Spectral stacking aligned to the velocity at peak intensity provides a rough measure of individual HI spectral components.** Stacking based on the velocity at peak intensity minimizes the stacked spectrum width compared to using the centroid or circular rotation velocities. However, the stacking method removes the ability to distinguish variations in individual HI component properties. Where bright emission is detected in individual spectra, it is preferable to fit individual spectra to better recover information about individual HI components.

Finally, our observations detect and resolve a population of HVCs, including an 8 kpc HI "filament" in the southern half of M33 and a ~ 600 pc HI cloud in the northern half which overlaps with the main disc in velocity. These HVCs and related HI structures have a combined mass of $1.3 \pm 0.5 \times 10^7$ $M_\odot$, about 1% of the total atomic ISM mass in M33. We suggest that these HI features map interactions between M33 and the warped HI disc on larger scales (below our detection threshold and outside of the area mapped with the VLA; Putman et al., 2009). These structures provide a detailed look into how HI from the CGM and galactic halo is accreted onto galactic discs.

Chapter 4 builds on this work by directly fitting spectral models to the HI spectra from the M33 observations in Chapter 3 and additional HI VLA observations that map the northern half of M31. Specifically, we compare two spectral models: a multi-Gaussian model and a single opaque component model. The latter is explored in archival VLA observations of M31, M33, and the LMC in Braun et al. (2009) and Braun (2012). They suggest the presence of HI spectra with flat-tops in all three galaxies, which they interpret as opaque HI emission. A strong flattened top from opaque emission follows from basic radiative transfer, however, to observe such features would require the opaque emission to have a large beam filling factor on the 100 pc scales used in Braun et al. (2009) and Braun (2012). From this model, they suggest that the HI

mass assuming optically-thin emission underestimates the total atomic ISM mass by $\sim 35\%$ without accounting for opaque HI.

From visual inspection of our new HI observations, many spectra clearly show multiple Gaussian components. To distinguish which model best represents the HI spectra, I fit the opaque model from Braun et al. (2009) and a multi-Gaussian model from an automated multicomponent fitter adapted from Lindner et al. (2015) and Riener et al. (2020). I then quantitatively determine the preferred model using a model selection test based on the Bayesian Information Criterion (BIC), which strongly penalizes for additional free model parameters. From the model selection test, I find that:

6. 99% **of the** HI **spectra show a preference for the multi-Gaussian model fit (**$\sim 80\%$ **with a strong preference).** The preference becomes stronger with lines-of-sight with large HI integrated intensity, opposite of the expectation for opaque HI at high column densities (Braun, 1997).

7. **Spectra where the opaque model is preferred account for an additional 1% in opaque, dark** HI **mass.** I compare the inferred opaque HI mass to the optically-thin cases to estimate the correction factor relative to the optically-thin HI masses. I find that the correction factor depends strongly on the sample selection criteria for where the opaque model is considered to be a good fit. Using the BIC model comparison for the 1% spectra where the opaque model is preferred, the mass correction factor is just 1%. This does not imply that lack of opaque cold HI in M31 and M33. Instead, it suggests that the opaque model assumptions of beam-filling cold HI on 100 pc is not supported by our observations.

8. **The opaque** HI **mass using the Braun et al. (2009) model is uncertain due to a strong dependence on the goodness-of-fit and choice of allowed parameter space.** To directly compare with the Braun (2012) mass correction factors, I adopt their valid fit criteria,

a reduced $\chi^2 < 25$, and find that this excludes no opaque HI model fits, even when there are large residuals in the opaque model fit. Over the entire sample, disregarding the BIC model selection, the opaque HI mass correction factor is 118^{+119}_{-40}% and 37^{+83}_{-12}% for M31 and M33, respectively (1σ uncertainties given). This estimate is equivalent to the Braun (2012) mass correction factor for M33, but much larger for M31. However, the uncertainties are large due to the non-linear response in the opaque model. I find that the larger mass correction factor for M31 is from a set of fits with extremely large inferred peak optical depths ($\tau_{\rm peak} > 5$), which fall within the allowed parameter space also from Braun et al. (2009) and Braun (2012), but are much larger than what is found in HI absorption studies (e.g., Murray et al., 2018). The large $\tau_{\rm peak}$ fits are due to lines-of-sight in M31 that trace a longer path along the inclined disc, and so have distinct peaks in the spectrum over a wide velocity range. Together, these results suggest that the inferred dark HI mass is primarily sensitive to the chosen parameter space and is poorly constrained due to the large uncertainties.

I ensure the BIC model selection test is sensitive to differences between the two models by drawing a sample of $20,000$ synthetic spectra from the opaque model, adding a noise level to match the observations and sampling at the same spectral resolution. Using the same fitting procedure for both models, I confirm the BIC model selection test correctly prefers the opaque model, demonstrating that the model comparisons of the observed spectra has sufficient power to distinguish between the models.

Together, these works demonstrate (i) the potential of Local Group HI observations to trace complex HI kinematics, and (ii) the *need* for accurate HI line modeling to guide our interpretation of the atomic ISM. The multi-Gaussian HI decomposition in Chapter 4 enables a more detailed look into HI kinematics on molecular clouds scales. I expand on this point further in §7.2.

Our new HI maps of the Local Group show that extragalactic HI studies

have progressed beyond the approximate methods used in archival observations (e.g., second moment, line stacking). Previously, uncertainties on the line width and spectral shape were limited by the data uncertainty. The vast differences I find in HI line widths and spectral modeling suggests that these approximate methods no longer adequately describe the data and that the uncertainties on the spectral shape are now dominated by systematics (bias) of these methods. To advance HI studies in the Local Group, more sophisticated modeling, including multi-Gaussian decomposition, is required.

Are atomic (HI) and molecular (H_2) ISM kinematics coupled on 80 pc scales? Yes, there is strong evidence for coupled atomic and molecular ISM kinematics on ~ 80 pc scales in M33. However, higher resolution observations are required to minimally-resolve the GMCs and provide links to molecular cloud evolution.

In Chapter 5, I compare the spectral properties of HI and CO(2-1) in M33[1] using two methods: (i) spectral stacking, typically applied to data with coarse resolution; and (ii) individual spectral modeling, typically used at higher physical resolution. These analyses measure the relationship of atomic and molecular clouds averaged over galactic scales (stacking) and across individual molecular clouds (spectral modeling). My co-authors and I find three key results:

9. **The velocity at peak intensity is strongly correlated between HI and CO**, with a difference that is typically consistent within the coarse 2.6 km s^{-1} spectral resolution of the CO data. Large outliers result from HI spectra with multiple distinct components where the CO peaks centered on a fainter HI component.

10. **M33 has a minimal thick molecular disc compared to more massive nearby spiral galaxies.** Following Chapter 3, we stack the HI and CO aligned to the velocity at peak HI intensity. We find that the

[1]I make use the IRAM 30-m M33 map, a large project mapping the inner 7 kpc of M33 (Gardan et al., 2007; Gratier et al., 2010; Druard et al., 2014).

CO stacked line width is consistently 70% of the HI stacked line width, unlike other nearby galaxies that find equivalent stacked line widths. We further show that CO line wings in the stacked spectrum are consistent with (i) error beam pick-up from the IRAM 30-m telescope (which is not corrected for in Druard et al., 2014), and (ii) systematics from the stacking technique. Previous work on other nearby galaxies use the equivalent line widths and line wings to infer the presence of diffuse CO emission. M33 lacks this significant thick molecular disc based on our findings.

11. HI **and CO line widths are strongly correlated in individual spectra at 80 pc scales.** However, this correlation is apparent when only the HI peak most closely associated with CO is modeled. Similar to Chapters 3 & 4, our observations demonstrate the need for more complex spectral modeling of HI and for a small number of CO spectra that also show multiple components along a single line-of-sight. Our results follow those of Fukui et al. (2009) in the LMC, where they also find a correlation between HI and CO line widths and suggest that most HI along a line-of-sight is not associated with CO (Wong et al., 2009). The spectral modeling from Chapter 4 will enable further comparisons of HI and CO spectral properties.

For comparisons between spectral lines from two different tracers we recommend that: (i) Stacking is best used to with low S/N data to achieve a detection. Due to the significant systematics we find, more information can be recovered from the emission detected in individual spectra. (ii) Spectral-line observations that do not resolve the spectral line shape should fit models while forward-modelling the the spectral channel sampling and, if required, the spectral response function of the telescopes (also see Sun et al., 2018). Many strong CO detections in the IRAM 30-m data are measured in ~ 3 channels where the coarse spectral channel sampling will underestimate the peak temperature and overestimate the line width. Forward modeling accounts for this bias (Koch et al., 2018b).

The correlation we find in the HI and CO line widths warrants further examination at higher resolutions ($\sim$ 25 pc). At the 80 pc we investigate, GMCs are distinct peaks in emission but remain unresolved. However, we find intrinsic scatter in the HI-CO line width relation that results from variations between GMCs. By comparing line widths at a higher spatial resolution, this relationship can be tested to determine whether the scatter arises from evolutionary variations in the cloud lifecycle, or from local spatial variations across clouds. I discuss this further in §7.2.

Is the source of large-scale turbulent driving imprinted on the ISM density structure? Using the spatial power spectrum, current observations do not provide a clear answer.

In Chapter 6, I model the spatial power spectrum of dust surface density, mid- to far-IR emission in the LMC, SMC, M31, and M33 from Utomo et al. (2019a). Additionally, I model the HI and CO spatial power spectra in M31 and M33. I compute the spatial power spectrum using the `TurbuStat` python package (Koch et al., 2019a) and forward model the power spectra with the PSF response function. The spatial power spectrum is a common technique to estimate the turbulent power spectrum. In this analysis, the density turbulent power spectrum is the quantity of interest. My co-authors and I find the following:

12. **Power spectrum breaks in nearly all of our spatial power spectra are due to the instrument PSF response and are not physical.** Several previous observational studies find a break in the spatial power spectrum and interpret this as tracing the disc scale height in nearby galaxies, which results from the turbulent power spectrum index changing by +1 as motions transition from two-dimensions on large scales to three-dimensions on small scales below the scale height. By forward modeling with the PSF, our results demonstrate this is not supported by the data for these four galaxies and, more generally, call this interpretation into question.

13. **Bright compact sources can produce power spectrum breaks.** We find a real power spectrum break at 24 μm wavelengths in the LMC, which we demonstrate is from the giant HII region, 30 Doradus. The break corresponds to the scale of 30 Doradus in the image and demonstrates that a real break in the power spectrum may not trace the disc scale height.

14. **The large-scale emission distribution in galaxies alters the power spectrum index, making it difficult to distinguish whether the index is related to 2D turbulence on large scales.** We find a consistent trend from steep to shallow indices from HI, dust surface density, and CO power spectra, respectively. This trend is consistent with HI having more power on large scales and CO having more structure on small scales (i.e., GMCs). Since the dust surface density traces both ISM components, the dust power spectrum index falls between these two extremes. We further find, in Chapter 5, that the concentration of GMCs (traced by CO) in the central region of M33 produces additional power on 2 kpc, matching the CO disc scale length (Druard et al., 2014). These variations on large scales demonstrates that spatial power spectra are sensitive to the large scale distribution of a particular tracer and is not uniquely defined by large-scale turbulence.

15. **Spatial power spectra computed over ~ 820 pc regions show real variations that may trace localized variations in turbulence,** as is expect, for example, near regions with recent strong stellar feedback. We compute local power spectra of the dust surface density in the LMC and SMC, which have a physical resolution of ~ 13 pc, well below the disc scale height. We find that (i) steeper power spectrum indices surround regions of bright emission and (ii) no strong evidence for localized power spectrum breaks. The former result is due to a large intensity gradient across the local region, similar to measure of skewness and kurtosis of the HI intensity distribution (Burkhart et al., 2010). The latter rules

out a spatially-varying disc scale height as the reason for no breaks in the galactic power spectra. Furthermore the lack of power spectrum breaks disagrees with a similar HI study of the LMC and SMC by Szotkowski et al. (2019), who find breaks near massive star-forming regions (e.g., 30 Doradus) Since the dust surface density measures both atomic and molecular ISM, we suggest that breaks in the HI power spectrum may better trace the impact of stellar feedback.

Our findings demonstrate how additional non-turbulent sources affect the spatial power spectrum. We suggest that measurements of large-scale 2D turbulence from the spatial power spectrum must correctly account for other sources that alter the power spectrum index. However, the power spectra we model can be fully explained with a simple power-law model. Other methods may be required to distinguish the large-scale source of turbulence.

We suggest that local measures of turbulence statistics offer a useful path forward to constrain the source of turbulence. Spatial variations in turbulent statistics can be compared with signatures of different driving mechanisms, for example, tracers of the star formation rate as a proxy for energy injection of stellar feedback from massive stars.

7.2 On-going & Future Work

The VLA L-band Local Group Legacy Survey My work is the first component of a complete high-resolution survey of the 21-cm HI, OH, and 1–2 GHz polarized continuum in M31, M33, NGC 6822, IC10, IC1613, and WLM. These observations will enable comprehensive study of atomic gas physics, turbulence, supernova remnants and HII regions across the Local Group. Combined with existing large large surveys at other wavelengths, these observations will provide the most comprehensive resolved studies of the baryonic cycle to date.

In addition to the observations I present in this thesis, our collaboration has aleady taken an additional 300 hr observations of these galaxies. In particular,

our additional M33 observations provide a 20 pc HI map of M33, and 8 pc resolution towards two regions within M33. The 20 pc HI observations have sufficient sensitivity to resolve the HI complexities highlighted in Chapters 3 & 4 on 80 pc scales.

When fully observed, the complete survey will include an additional 1800 hr of observations from the VLA and will form an unparalleled legacy product available for the community.

Cloud formation and HI envelopes from detailed HI kinematics The atomic ISM is expected to play a role throughout the molecular cloud lifecycle, from forming H_2 (Dobbs et al., 2014), to shielding in an envelope around molecular clouds (Krumholz et al., 2009), and tracing stellar feedback as clouds are destroyed (Kruijssen et al., 2019).

In Chapter 5, I show that there is indeed a relation between the HI and CO (tracing H_2) kinematics. However, the 80 pc resolution of the HI observations preclude testing for variations between and within molecular clouds. To overcome this, our 20 pc HI map of M33 will be combined with a new CO(2-1) mosaic recently observed with ALMA. This new ALMA map resolves 20 pc ($\sim 6''$) scales in the CO(2-1) line with a spectral resolution of 0.4 km s^{-1}, far exceeding the IRAM 30-m observations with 2.6 km s^{-1} spectral channels. My preliminary analysis of these observations suggests that the strong correlation between HI and CO persists on 20 pc scales.

With these new HI and CO observations of M33, my on-going work uses the multi-Gaussian modeling from Chapter 4 to explore:

1. Common resolved kinematics between HI and CO in early non-star forming molecular clouds to observationally constrain cloud formation mechanisms.

2. Typical HI line widths within and surrounding molecular clouds to determine (i) if the atomic HI envelope is gravitationally-bound, and (ii) if cool HI ($\sigma < 6$ km s^{-1}) is preferentially detected towards molecular

clouds.

3. How the total cloud mass, combining HI and CO, compares with the CO-only derived cloud mass (e.g., Elmegreen & Elmegreen, 1987) to determine (i) whether the total cloud mass varies with environment (Colombo et al., 2014), and (ii) the fraction of CO-dark H_2 in individual molecular clouds (Wolfire et al., 2010).

This future and on-going work builds on Chapter 5 to provide more detailed comparisons relating the atomic and molecular ISM.

Resolving stellar feedback coupling to the ISM Stellar feedback produces obvious effects on the ISM, most notably HI bubbles or shells surround giant HII and supernova remnants (Tenorio-Tagle & Bodenheimer, 1988). What remains unclear is how efficient feedback couples with the ISM. Large-scale averaging suggests that supernova can drive observed levels of turbulence if $\sim 10\%$ of the energy transferred to the ISM (Tamburro et al., 2009; Stilp et al., 2013b; Koch et al., 2018c; Utomo et al., 2019b). If this is indeed the case, resolved HI observations should show enhanced turbulent properties near recent sources of stellar feedback.

Our HI observations can explicitly test whether enhanced HI turbulence is observed by combining the HI multi-Gaussian models from Chapter 4 with the supernova remnants and HII regions detected in our L-band continuum data, and the locations of all O and B stars from *HST* photometry (e.g., from the PHAT project; Dalcanton et al., 2012) and young stellar clusters (Johnson et al., 2015). This measurement will test recent simulation (Seifried et al., 2018) and theory (Krumholz & Burkhart, 2016) results that challenge supernovae as the large-scale driver of turbulence.

7.3 Final Thoughts

Over the next few years, observations from the VLA L-band Local Group Legacy Survey will have a transformative effect on atomic gas physics and

stellar feedback. My work is a first demonstration of how high spectral resolution HI observations will alter many of our views of the extragalactic ISM. Building on my work and our collaboration's on-going efforts, there are two clear impacts for how this survey will affect our understanding of the ISM.

The first is accessibility. Koch et al. (2018c, Chapter 3) provides the *first* publicly available M33 HI data cube (along with derived products), despite decades of Local Group HI observations and large time investments from various observatories (Chapter 2). The full suite of observations I use in this thesis will form an initial data release for the VLA L-band Local Group Legacy Survey.

The availability of Local Group observations is crucial for combining Galactic and nearby galaxy surveys, where major surveys are already released to the community (for example, Walter et al., 2008; Hunter et al., 2012; Winkel et al., 2016, and many others). The accessibility of these observations in a science-ready format is enormous for fully exploring the data and is a major reason why some surveys have guided our interpretation of the ISM.

The second impact is defining our view of the neutral ISM leading into the era of the next generation of cm-interferometers, namely the Square Kilometre Array (SKA) and next-generation VLA (ngVLA). The restriction on studying HI on molecular cloud scales (< 20 pc) beyond the Local Group is the lack of a cm-interferometer that combines long baselines (> 11 km) with high surface brightness sensitivity for imaging extended emission. This is the primary reason why the Local Group galaxies are such valuable targets for the topics I explore in this thesis, as we can currently only resolve ~ 20 pc scales in HI within $D < 1$ Mpc.

The SKA and ngVLA will dramatically overcome this limitation. In particular, these telescopes will enable the detailed HI kinematic studies that I present in this thesis to ~ 100 nearby galaxies, matching the $\sim 1''$ imaging that ALMA achieves at mm-wavelengths. Expanding the sample size for high-resolution HI studies is essential for robust tests of how density increases in the neutral ISM (i.e., cloud formation, H_2 formation) leading to star formation.

These constraints play an important role for refining the ISM models used in cosmological simulations, and for comparisons of the ISM and star formation properties measured at high redshift. My work in this thesis, and continuing with the large Local Group survey, will provide the analysis and modeling techniques for observations with the next generation of cm-inteferometers.

Despite already being one of the best-studied tracers, the 21-cm HI spectral line continues to answer fundamental questions on ISM physics and the baryonic cycle.

Bibliography

Adamo A., Kruijssen J. M. D., Bastian N., Silva-Villa E., Ryon J., 2015, MNRAS, 452, 246

André P., Di Francesco J., Ward-Thompson D., Inutsuka S. I., Pudritz R. E., Pineda J. E., 2014, Protostars and Planets VI, pp 27–51

Aniano G., Draine B. T., Gordon K. D., Sandstrom K., 2011, PASP, 123, 1218

Araya E., Baan W. A., Hofner P., 2004, ApJS, 154, 541

Arce H. G., Borkin M. A., Goodman A. A., Pineda J. E., Halle M. W., 2010, ApJ, 715, 1170

Arce H. G., Borkin M. A., Goodman A. A., Pineda J. E., Beaumont C. N., 2011, ApJ, 742, 105

Argyle E., 1965, ApJ, 141, 750

Astropy Collaboration et al., 2013, Astronomy & Astrophysics, 558, A33

Astropy Collaboration et al., 2018, AJ, 156, 123

Bagetakos I., Brinks E., Walter F., de Blok W. J. G., Usero A., Leroy A. K., Rich J. W., Kennicutt R. C. J., 2011, AJ, 141, 23

Bajaja E., Shane W. W., 1982, A&AS, 49, 745

Barbieri C. V., Fraternali F., Oosterloo T., Bertin G., Boomsma R., Sancisi R., 2005, A&A, 439, 947

Barmby P., et al., 2006, ApJ, 650, L45

Barnabè M., Ciotti L., Fraternali F., Sancisi R., 2006, A&A, 446, 61

Begum A., Chengalur J. N., Bhardwaj S., 2006, MNRAS, 372, L33

Bertram E., Konstandin L., Shetty R., Glover S. C. O., Klessen R. S., 2015, MNRAS, 446, 3777

Bialy S., Sternberg A., 2019, ApJ, 881, 160

Bialy S., Burkhart B., Sternberg A., 2017, ApJ, 843, 92

Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2846

Bigiel F., Bolatto A. D., Leroy A. K., Blitz L., Walter F., Rosolowsky E. W., Lopez L. A., Plambeck R. L., 2010, ApJ, 725, 1159

Bihr S., et al., 2015, A&A, 580, A112

Blagrave K., Martin P. G., Joncas G., Kothes R., Stil J. M., Miville-Deschênes M. A., Lockman F. J., Taylor A. R., 2017, ApJ, 834, 126

Blitz L., Rosolowsky E., 2006, ApJ, 650, 933

Block D. L., Puerari I., Elmegreen B. G., Elmegreen D. M., Fazio G. G., Gehrz R. D., 2009, ApJ, 694, 115

Block D. L., Puerari I., Elmegreen B. G., Bournaud F., 2010, ApJ, 718, L1

Bolatto A. D., et al., 2007, ApJ, 655, 212

Bolatto A. D., et al., 2011, ApJ, 741, 12

Bolatto A. D., Wolfire M., Leroy A. K., 2013a, Annual Review of A&A, 51, 207

Bolatto A. D., et al., 2013b, Nature, 499, 450

Boothroyd A. I., Blagrave K., Lockman F. J., Martin P. G., Pinheiro Gonçalves D., Srikanth S., 2011, A&A, 536, A81

Boquien M., et al., 2015, A&A, 578, A8

Boulanger F., Viallefond F., 1992, A&A, 266, 37

Bournaud F., Elmegreen B. G., Teyssier R., Block D. L., Puerari I., 2010, MNRAS, 409, 1088

Boyden R. D., Koch E. W., Rosolowsky E. W., Offner S. S. R., 2016, ApJ, 833, 233

Braine J., Rosolowsky E., Gratier P., Corbelli E., Schuster K. F., 2018, A&A, 612, A51

Brandt J. C., 1960, ApJ, 131, 293

Braun R., 1990, ApJS, 72, 755

Braun R., 1997, ApJ, 484, 637

Braun R., 2012, ApJ, 749, 87

Braun R., Walterbos R. A. M., 1992, ApJ, 386, 120

Braun R., Thilker D. A., Walterbos R. A. M., Corbelli E., 2009, ApJ, 695, 937

Bresolin F., 2011, ApJ, 730, 129

Brinks E., Shane W. W., 1984, A&AS, 55, 179

Brundage W. D., Kraus J. D., 1966, Science, 153, 411

Burgers J., 1948, Elsevier, pp 171 – 199, doi:https://doi.org/10.1016/S0065-2156(08)70100-5, http://www.sciencedirect.com/science/article/pii/S0065215608701005

Burke B. F., Ecklund E. T., Joenson P. A., Turner K. C., Tuve M. A., 1963, AJ, 68, 70

Burkhart B., Stanimirović S., Lazarian A., Kowal G., 2010, ApJ, 708, 1204

Caldú-Primo A., Schruba A., 2016a, AJ, 151, 34

Caldú-Primo A., Schruba A., 2016b, AJ, 151, 34

Caldu-Primo A., Schruba A., Walter F., Leroy A., Sandstrom K., de Blok W. J. G., Ianjamasimanana R., Mogotsi K. M., 2013, AJ, 146, 150

Caldú-Primo A., Schruba A., Walter F., Leroy A., Bolatto A. D., Vogel S., 2015, AJ, 149, 76

Carignan C., Chemin L., Huchtmeier W. K., Lockman F. J., 2006, ApJ, 641, L109

Carroll B. W., Ostlie D. A., 2007, An introduction to modern astrophysics; 2nd ed.. Addison-Wesley, San Francisco, CA, https://cds.cern.ch/record/1009754

Chemin L., Carignan C., Foster T., 2009, ApJ, 705, 1395

Chepurnov A., Burkhart B., Lazarian A., Stanimirovic S., 2015, ApJ, 810, 33

Chevance M., et al., 2020a, Space Sci. Rev., 216, 50

Chevance M., et al., 2020b, MNRAS, 493, 2872

Chiang I. D., Sandstrom K. M., Chastenet J., Johnson L. C., Leroy A. K., Utomo D., 2018, ApJ, 865, 117

Colombo D., et al., 2014, ApJ, 784, 3

Colombo D., et al., 2019, MNRAS, 483, 4291

Combes F., Becquaert J. F., 1997, A&A, 326, 554

Combes F., et al., 2012, A&A, 539, A67

Corbelli E., 2003, MNRAS, 342, 199

Corbelli E., Schneider S. E., 1997, ApJ, 479, 244

Corbelli E., Lorenzoni S., Walterbos R., Braun R., Thilker D., 2010, A&A, 511, A89

Corbelli E., Thilker D., Zibetti S., Giovanardi C., Salucci P., 2014, A&A, 572, A23

Corbelli E., et al., 2017, A&A, 601, A146

Cornwell T. J., 2008, IEEE Journal of Selected Topics in Signal Processing, 2, 793

Cram T. R., Roberts M. S., Whitehurst R. N., 1980, A&AS, 40, 215

Dalcanton J. J., et al., 2012, ApJS, 200, 18

Dale J. E., 2017, MNRAS, 467, 1067

Dale J. E., Ngoumou J., Ercolano B., Bonnell I. A., 2014, MNRAS, 442, 694

Dame T. M., Thaddeus P., 1994, ApJ, 436, L173

Davies R. D., Gottesman S. T., 1970, MNRAS, 149, 237

Dawson J. R., McClure-Griffiths N. M., Dickey J. M., Fukui Y., 2011, ApJ, 741, 85

de Blok W. J. G., Walter F., Brinks E., Trachternach C., Oh S. H., Kennicutt R. C. J., 2008, AJ, 136

de Vaucouleurs G., de Vaucouleurs A., Corwin Jr. H. G., Buta R. J., Paturel G., Fouqué P., 1991, Third Reference Catalogue of Bright Galaxies. Volume I: Explanations and references. Volume II: Data for galaxies between 0^h and 12^h. Volume III: Data for galaxies between 12^h and 24^h.

Del Moral P., Doucet A., Jasra A., 2006, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 68, 411

Deshpande A. A., Dwarakanath K. S., Goss W. M., 2000, ApJ, 543, 227

Deul E. R., van der Hulst J. M., 1987, A&AS, 67, 509

Di Teodoro E. M., et al., 2019, MNRAS, 483, 392

Dickey J. M., Brinks E., 1993, ApJ, 405, 153

Dickey J. M., Lockman F. J., 1990, ARA&A, 28, 215

Dickey J. M., Hanson M. M., Helou G., 1990, ApJ, 352, 522

Dickey J. M., Mebold U., Marx M., Amy S., Haynes R. F., Wilson W., 1994, A&A, 289, 357

Dickey J. M., Mebold U., Stanimirovic S., Staveley-Smith L., 2000, ApJ, 536, 756

Dickey J. M., McClure-Griffiths N. M., Stanimirović S., Gaensler B. M., Green A. J., 2001, ApJ, 561, 264

Dickey J. M., McClure-Griffiths N. M., Gaensler B. M., Green A. J., 2003, ApJ, 585, 801

Dieter N. H., 1962, AJ, 67, 217

Dobbs C. L., 2008, MNRAS, 391, 844

Dobbs C. L., et al., 2014, Protostars and Planets VI, pp 3–26

Dobbs C. L., Pettitt A. R., Corbelli E., Pringle J. E., 2018, MNRAS, 478, 3793

Dobbs C. L., Rosolowsky E., Pettitt A. R., Braine J., Corbelli E., Sun J., 2019, MNRAS, 485, 4997

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

Draine B. T., Lee H. M., 1984, ApJ, 285, 89

Draine B. T., et al., 2014, ApJ, 780, 172

Druard C., et al., 2014, A&A, 567, A118

Dutta P., Bharadwaj S., 2013, MNRAS, 436, L49

Dutta P., Begum A., Bharadwaj S., Chengalur J. N., 2008, MNRAS, 384, L34

Dutta P., Begum A., Bharadwaj S., Chengalur J. N., 2009a, MNRAS, 397, L60

Dutta P., Begum A., Bharadwaj S., Chengalur J. N., 2009b, MNRAS, 398, 887

Dutta P., Begum A., Bharadwaj S., Chengalur J. N., 2013, New Astron., 19, 89

Elmegreen B. G., Elmegreen D. M., 1987, ApJ, 320, 182

Elmegreen B. G., Scalo J., 2004, Annual Review of Astronomy &Astrophysics, 42, 211

Elmegreen B. G., Kim S., Staveley-Smith L., 2001, ApJ, 548, 749

Elmegreen B. G., Elmegreen D. M., Leitner S. N., 2003a, ApJ, 590, 271

Elmegreen B. G., Leitner S. N., Elmegreen D. M., Cuillandre J.-C., 2003b, ApJ, 593, 333

Elmegreen B. G., Elmegreen D. M., Chand ar R., Whitmore B., Regan M., 2006, ApJ, 644, 879

Emerson D. T., 1974, MNRAS, 169, 607

Federrath C., 2013, MNRAS, 436, 1245

Federrath C., Roman-Duval J., Klessen R. S., Schmidt W., Mac Low M. M., 2010, A&A, 512, A81

Ferrière K. M., 2001, Reviews of Modern Physics, 73, 1031

Field G. B., Goldsmith D. W., Habing H. J., 1969, ApJ, 155, L149

Fleck Robert C. J., 1996, ApJ, 458, 739

Foreman-Mackey D., 2016, The Journal of Open Source Software, 24

Fraternali F., Oosterloo T., Sancisi R., van Moorsel G., 2001, ApJ, 562, L47

Fraternali F., van Moorsel G., Sancisi R., Oosterloo T., 2002a, AJ, 123, 3124

Fraternali F., Cappi M., Sancisi R., Oosterloo T., 2002b, ApJ, 578, 109

Fraternali F., Oosterloo T., Sancisi R., 2004, A&A, 424, 485

Freedman W. L., et al., 2001, ApJ, 553, 47

Fukui Y., et al., 2009, ApJ, 705, 144

Galtier S., Banerjee S., 2011, Phys. Rev. Lett., 107, 134501

Garcia-Burillo S., Guelin M., Cernicharo J., Dahlem M., 1992, A&A, 266, 21

Gardan E., Braine J., Schuster K. F., Brouillet N., Sievers A., 2007, A&A, 473, 91

Garrison-Kimmel S., et al., 2019, MNRAS, 487, 1380

Gebhardt K., et al., 2001, AJ, 122, 2469

Gibson S. J., Taylor A. R., Higgs L. A., Dewdney P. E., 2000, ApJ, 540, 851

Gibson S. J., Taylor A. R., Higgs L. A., Brunt C. M., Dewdney P. E., 2005, ApJ, 626, 195

Goldman I., 2000, ApJ, 541, 701

Goldsmith P. F., Li D., 2005, ApJ, 622, 938

Gordon K. J., 1971, ApJ, 169, 235

Gordon K. D., et al., 2006, ApJ, 638, L87

Gordon K. D., et al., 2011, AJ, 142, 102

Gordon K. D., et al., 2014, ApJ, 797, 85

Gottesman S. T., Davies R. D., Reddish V. C., 1966, MNRAS, 133, 359

Grabelsky D. A., Cohen R. S., Bronfman L., Thaddeus P., May J., 1987, ApJ, 315, 122

Graczyk D., et al., 2014, ApJ, 780, 59

Gratier P., et al., 2010, A&A, 522, A3

Gratier P., et al., 2012, A&A, 542, A108

Gray M., 2012, Maser Sources in Astrophysics

Green D. A., 1993, MNRAS, 262, 327

Griffin M. J., et al., 2010, A&A, 518, L3

Grisdale K., Agertz O., Romeo A. B., Renaud F., Read J. I., 2017, MNRAS, 466, 1093

Grossi M., Giovanardi C., Corbelli E., Giovanelli R., Haynes M. P., Martin A. M., Saintonge A., Dowell J. D., 2008, A&A, 487, 161

Groves B., et al., 2012, MNRAS, 426, 892

Guibert J., 1973, A&AS, 12, 263

Haffner L. M., et al., 2009, Reviews of Modern Physics, 81, 969

Haud U., 2000, A&A, 364, 83

Haworth T. J., Glover S. C. O., Koepferl C. M., Bisbas T. G., Dale J. E., 2018, New Astron. Rev., 82, 1

Heald G., et al., 2011, A&A, 526, A118

Heiles C., 1979, Publications of the Astronomical Society of the Pacific, 91, 611

Heiles C., Troland T. H., 2003, ApJ, 586, 1067

Henshaw J. D., et al., 2016, MNRAS, 457, 2675

Heyer M. H., Carpenter J. M., Snell R. L., 2001, ApJ, 551, 852

Heyer M. H., Corbelli E., Schneider S. E., Young J. S., 2004, ApJ, 602, 723

Hinz J. L., et al., 2004, ApJS, 154, 259

Hoffmann V., Romeo A. B., 2012, MNRAS, 425, 1511

Högbom J. A., 1974, A&AS, 15, 417

Hogg D. W., Bovy J., Lang D., 2010, preprint, p. arXiv:1008.4686 (`arXiv:1008.4686`)

Hopkins P. F., 2014, ApJ, 797, 59

Hopkins P. F., Squire J., 2018, MNRAS, 479, 4681

Huchtmeier W., 1972, A&AS, 7, 397

Huchtmeier W. K., 1978, in Berkhuijsen E. M., Wielebinski R., eds, IAU Symposium Vol. 77, Structure and Properties of Nearby Galaxies. p. 197

Hunter J. D., 2007, Computing In Science & Engineering, 9, 90

Hunter D. A., et al., 2012, AJ, 144, 134

Ianjamasimanana R., de Blok W. J. G., Walter F., Heald G. H., 2012, AJ, 144, 96

Ianjamasimanana R., de Blok W. J. G., Walter F., Heald G. H., Caldu-Primo A., Jarrett T. H., 2015, AJ, 150, 47

Ianjamasimanana R., de Blok W. J. G., Heald G. H., 2017, AJ, 153, 213

Imara N., Bigiel F., Blitz L., 2011, ApJ, 732, 79

Indebetouw R., et al., 2013, ApJ, 774, 73

Jameson K. E., et al., 2016, ApJ, 825, 12

Jameson K. E., et al., 2019, ApJS, 244, 7

Jeffreson S. M. R., Kruijssen J. M. D., 2018, MNRAS, 476, 3688

Johnson L. C., et al., 2015, ApJ, 802, 127

Kalberla P. M. W., Kerp J., 2009, ARA&A, 47, 27

Kalberla P. M. W., Kerp J., Haud U., Haverkorn M., 2017, A&A, 607, A15

Kam S. Z., Carignan C., Chemin L., Foster T., Elson E., Jarrett T. H., 2017, AJ, 154, 41

Karachentsev I. D., Karachentseva V. E., Huchtmeier W. K., Makarov D. I., 2004, AJ, 127, 2031

Kass R. E., Raftery A. E., 1995, Journal of the American Statistical Association, 90, 773

Kavars D. W., Dickey J. M., McClure-Griffiths N. M., Gaensler B. M., Green A. J., 2005, ApJ, 626, 887

Kawamura A., et al., 2009, ApJS, 184, 1

Keenan O. C., Davies J. I., Taylor R., Minchin R. F., 2016, MNRAS, 456, 951

Kennicutt R. C. J., 1989, ApJ, 344, 685

Kennicutt R. C. J., 1998, ApJ, 498, 541

Kennicutt R. C., Evans N. J., 2012, ARA&A, 50, 531

Kennicutt R. C. J., et al., 2003, The Publications of the Astronomical Society of the Pacific, 115, 928

Kennicutt R. C., et al., 2011, Publications of the Astronomical Society of Pacific, 123, 1347

Keown J., Di Francesco J., Teimoorinia H., Rosolowsky E., Chen M. C.-Y., 2019, ApJ, 885, 32

Kepley A. A., Tsutsumi T., Brogan C. L., Indebetouw R., Yoon I., Mason B., Donovan Meyer J., 2020, PASP, 132, 024505

Kim S., Dopita M. A., Staveley-Smith L., Bessell M. S., 1999, AJ, 118, 2797

Kim S., Staveley-Smith L., Dopita M. A., Sault R. J., Freeman K. C., Lee Y., Chu Y.-H., 2003, ApJS, 148, 473

Klein C. R., Cenko S. B., Miller A. A., Norman D. J., Bloom J. S., 2014, arXiv e-prints, p. arXiv:1405.1035

Koch E., Rosolowsky E., Johnson M. C., Kepley A. A., Leroy A., 2018a, RNAAS, 2, 24

Koch E., Rosolowsky E., Leroy A. K., 2018b, RNAAS, 2, 220

Koch E. W., et al., 2018c, MNRAS, 479, 2505

Koch E. W., Rosolowsky E. W., Boyden R. D., Burkhart B., Ginsburg A., Loeppky J. L., Offner S. S. R., 2019a, AJ, 158, 1

Koch E. W., Rosolowsky E. W., Schruba A., Leroy A., Kepley A., Braine J., Dalcanton J., Johnson M. C., 2019b, MNRAS, 485, 2324

Koch E. W., Chiang I. D., Utomo D., Chastenet J., Leroy A. K., Rosolowsky E. W., Sandstrom K. M., 2020, MNRAS, 492, 2663

Koda J., Teuben P., Sawada T., Plunkett A., Fomalont E., 2019, PASP, 131, 054505

Kolmogorov A. N., 1941, Akademiia Nauk SSSR Doklady, 32, 16

Kormendy J., McClure R. D., 1993, AJ, 105, 1793

Kramer C., et al., 2010, A&A, 518, L67

Krieger N., et al., 2019, ApJ, 881, 43

Kruijssen J. M. D., et al., 2019, Nature, 569, 519

Krumholz M. R., 2013, MNRAS, 436, 2747

Krumholz M. R., 2014, Physics Reports, 539, 49

Krumholz M. R., 2017, Star Formation, doi:10.1142/10091.

Krumholz M. R., Burkhart B., 2016, MNRAS, 458, 1671

Krumholz M. R., McKee C. F., Tumlinson J., 2009, ApJ, 693, 216

Krumholz M. R., et al., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. p. 243 (arXiv:1401.2473), doi:10.2458/azu˙uapress˙9780816531240-ch011

Krumholz M. R., Burkhart B., Forbes J. C., Crocker R. M., 2018, MNRAS, 477, 2716

Krčo M., Goldsmith P. F., 2010, ApJ, 724, 1402

Krčo M., Goldsmith P. F., Brown R. L., Li D., 2008, ApJ, 689, 276

Kulkarni S. R., Heiles C., 1987, in , Interstellar Processes. Springer Netherlands, Dordrecht, pp 87–122

Kurono Y., Morita K.-I., Kamazaki T., 2009, Publications of the Astronomical Society of Japan, 61, 873

Kuzio de Naray R., Arsenault C. A., Spekkens K., Sellwood J. A., McDonald M., Simon J. D., Teuben P., 2012, MNRAS, 427, 2523

Lada C. J., Lada E. A., 2003, ARA&A, 41, 57

Larson R. B., 1981, MNRAS, 194, 809

Lazarian A., Pogosyan D., 2000, ApJ, 537, 720

Lazarian A., Pogosyan D., 2006, ApJ, 652, 1348

Lee M.-Y., Stanimirović S., Murray C. E., Heiles C., Miller J., 2015, ApJ, 809, 56

Lee E. J., Miville-Deschênes M.-A., Murray N. W., 2016, ApJ, 833, 229

Leroy A. K., Walter F., Brinks E., Bigiel F., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2782

Leroy A. K., et al., 2009, AJ, 137, 4670

Leroy A. K., et al., 2013, AJ, 146, 19

Leroy A. K., et al., 2015, ApJ, 814, 83

Leroy A. K., et al., 2016, ApJ, 831

Lewis A. R., et al., 2015, ApJ, 805, 183

Li D., Goldsmith P. F., 2003, ApJ, 585, 823

Lindner R. R., et al., 2015, AJ, 149, 138

Liu B., Li D., Staveley-Smith L., Qian L., Wong T., Goldsmith P., 2019, ApJ, 887, 242

Lockman F. J., 2017, in Fox A., Davé R., eds, Astrophysics and Space Science Library Vol. 430, Astrophysics and Space Science Library. p. 49 (arXiv:1612.00774), doi:10.1007/978-3-319-52512-9˙3

Lockman F. J., Murphy E. M., Petty-Powell S., Urick V. J., 2002, ApJS, 140, 331

Lockman F. J., Free N. L., Shields J. C., 2012, AJ, 144, 52

Mac Low M.-M., Klessen R. S., 2004a, Reviews of Modern Physics, 76, 125

Mac Low M.-M., Klessen R. S., 2004b, Reviews of Modern Physics, 76, 125

Madau P., Dickinson M., 2014, ARA&A, 52, 415

Mangum J. G., Emerson D. T., Greisen E. W., 2007, A&A, 474, 679

Marchal A., Miville-Deschênes M.-A., Orieux F., Gac N., Soussen C., Lesot M.-J., d'Allonnes A. R., Salomé Q., 2019, A&A, 626, A101

Martin P. G., Blagrave K. P. M., Lockman F. J., Pinheiro Gonçalves D., Boothroyd A. I., Joncas G., Miville-Deschênes M. A., Stephan G., 2015, ApJ, 809, 153

McClure-Griffiths N. M., et al., 2018, Nature Astronomy, 2, 901

McKee C. F., Ostriker J. P., 1977, ApJ, 218, 148

McKee C. F., Ostriker E. C., 2007, Annual Review of A&A, 45, 565

McMullin J. P., Waters B., Schiebel D., Young W., Golap K., 2007, in Shaw R. A., Hill F., Bell D. J., eds, Astronomical Society of the Pacific Conference Series Vol. 376, Astronomical Data Analysis Software and Systems XVI. p. 127

Meidt S. E., Rand R. J., Merrifield M. R., Shetty R., Vogel S. N., 2008, ApJ, 688, 224

Meixner M., et al., 2006, AJ, 132, 2268

Meixner M., et al., 2013, AJ, 146, 62

Meng S. Y., Kraus J. D., 1966, AJ, 71, 170

Miesch M. S., Bally J., 1994, ApJ, 429, 645

Miville-Deschênes M. A., Lagache G., Puget J. L., 2002, A&A, 393, 749

Miville-Deschênes M. A., Joncas G., Falgarone E., Boulanger F., 2003, A&A, 411, 109

Miville-Deschênes M. A., et al., 2010, A&A, 518, L104

Miville-Deschênes M. A., Duc P. A., Marleau F., Cuillandre J. C., Didelon P., Gwyn S., Karabal E., 2016, A&A, 593, A4

Mogotsi K. M., de Blok W. J. G., Caldu-Primo A., Walter F., Ianjamasimanana R., Leroy A. K., 2016, AJ, 151, 15

Muller E., Stanimirović S., Rosolowsky E., Staveley-Smith L., 2004, ApJ, 616, 845

Murphy E. J., Dong D., Momjian E., Linden S., Kennicutt R. C. J., Meier D. S., Schinnerer E., Turner J. L., 2018, ApJS, 234, 24

Murray N., Rahman M., 2010, ApJ, 709, 424

Murray C. E., et al., 2015, ApJ, 804, 89

Murray C. E., Stanimirović S., Goss W. M., Heiles C., Dickey J. M., Babler B., Kim C.-G., 2018, ApJS, 238, 14

Nestingen-Palm D., Stanimirović S., González-Casanova D. F., Babler B., Jameson K., Bolatto A., 2017, ApJ, 845, 53

Newton K., 1980, MNRAS, 190, 689

Newton K., Emerson D. T., 1977, MNRAS, 181, 573

Newville M., et al., 2020, lmfit/lmfit-py 1.0.1, doi:10.5281/zenodo.3814709, https://doi.org/10.5281/zenodo.3814709

Nieten C., Neininger N., Guélin M., Ungerechts H., Lucas R., Berkhuijsen E. M., Beck R., Wielebinski R., 2006, Astronomy and Astrophysics, 453, 459

Offner S. S. R., Clark P. C., Hennebelle P., Bastian N., Bate M. R., Hopkins P. F., Moraux E., Whitworth A. P., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. p. 53 (arXiv:1312.5326), doi:10.2458/azu˙uapress˙9780816531240-ch003

Oh S.-H., de Blok W. J. G., Brinks E., Walter F., Kennicutt Jr. R. C., 2011, AJ, 141, 193

Oliphant T. E., 2006, A guide to NumPy. Vol. 1, Trelgol Publishing USA

Ostriker E. C., McKee C. F., Leroy A. K., 2010, ApJ, 721, 975

Padoan P., Kim S., Goodman A., Staveley-Smith L., 2001, ApJ, 555, L33

Padoan P., Juvela M., Kritsuk A., Norman M. L., 2009, ApJ, 707, L153

Peek J. E. G., et al., 2011, ApJS, 194, 20

Petric A. O., Rupen M. P., 2007, AJ, 134, 1952

Pety J., et al., 2013, ApJ, 779, 43

Pilbratt G. L., et al., 2010, A&A, 518, L1

Pilkington K., et al., 2011, MNRAS, 417, 2891

Pingel N. M., et al., 2013, ApJ, 779, 36

Pingel N. M., Lee M.-Y., Burkhart B., Stanimirović S., 2018, ApJ, 856, 136

Poglitsch A., et al., 2010, A&A, 518, L2

Putman M. E., et al., 2009, ApJ, 703, 1486

Putman M. E., Peek J. E. G., Joung M. R., 2012, ARA&A, 50, 491

Rahmani S., Lianou S., Barmby P., 2016, MNRAS, 456, 4128

Reakes M. L., Newton K., 1978, MNRAS, 185, 277

Regan M. W., Vogel S. N., 1994, ApJ, 434, 536

Regan M. W., Thornley M. D., Helfer T. T., Sheth K., Wong T., Vogel S. N., Blitz L., Bock D. C. J., 2001, ApJ, 561, 218

Renaud F., et al., 2013, MNRAS, 436, 1836

Rieke G. H., et al., 2004, ApJS, 154, 25

Riener M., Kainulainen J., Beuther H., Henshaw J. D., Orkisz J. H., Wang Y., 2020, A&A, 633, A14

Roberts M. S., 1966, ApJ, 144, 639

Roberts M. S., Whitehurst R. N., 1975, ApJ, 201, 327

Rogstad D. H., Wright M. C. H., Lockhart I. A., 1976, ApJ, 204, 703

Rohlfs K., Wilson T. L., 2004, Tools of radio astronomy

Rohlfs K., Braunsfurth E., Mebold U., 1972, AJ, 77, 711

Roman-Duval J., Heyer M., Brunt C. M., Clark P., Klessen R., Shetty R., 2016, ApJ, 818, 144

Romeo A. B., Agertz O., 2014, MNRAS, 442, 1230

Romeo A. B., Burkert A., Agertz O., 2010, MNRAS, 407, 1223

Rosolowsky E., Simon J. D., 2008, ApJ, 675, 1213

Rosolowsky E. W., Goodman A. A., Wilner D. J., Williams J. P., 1999, ApJ, 524, 887

Rosolowsky E., Engargiola G., Plambeck R., Blitz L., 2003, ApJ, 599, 258

Rosolowsky E., Keto E., Matsushita S., Willner S. P., 2007, ApJ, 661, 830

Rosolowsky E. W., Pineda J. E., Foster J. B., Borkin M. A., Kauffmann J., Caselli P., Myers P. C., Goodman A. A., 2008, The Astrophysical Journal Supplement Series, 175, 509

Salvatier J., Wiecki T. V., Fonnesbeck C., 2016, PeerJ Computer Science, 2, e55

Sancisi R., Fraternali F., Oosterloo T., van der Hulst T., 2008, A&ARv, 15, 189

Sandstrom K. M., et al., 2013, ApJ, 777, 5

Schmidt M., 1959, ApJ, 129, 243

Schmidt T. M., Bigiel F., Klessen R. S., de Blok W. J. G., 2016, MNRAS, 457, 2642

Schruba A., et al., 2011, AJ, 142, 37

Schruba A., Bialy S., Sternberg A., 2018, ApJ, 862, 110

Schwarz G., 1978, Ann. Statist., 6, 461

Seale J. P., Looney L. W., Chu Y.-H., Gruendl R. A., Brandl B., Chen C. H. R., Brandner W., Blake G. A., 2009, ApJ, 699, 150

Seifried D., Walch S., Haid S., Girichidis P., Naab T., 2018, ApJ, 855, 81

Sellwood J. A., Sánchez R. Z., 2010, MNRAS, 404, 1733

Sellwood J. A., Spekkens K., 2015, preprint, (`arXiv:1509.07120`)

Shetty R., Ostriker E. C., 2006, ApJ, 647, 997

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

Simonetti J. H., Cordes J. M., Spangler S. R., 1984, ApJ, 284, 126

Skrutskie M. F., et al., 2006, AJ, 131, 1163

Sokolov V., Pineda J. E., Buchner J., Caselli P., 2020, ApJ, 892, L32

Soler J. D., et al., 2019, A&A, 622, A166

Solomon P. M., Rivolo A. R., Barrett J., Yahil A., 1987, ApJ, 319, 730

Spekkens K., Sellwood J. A., 2007, ApJ, 664, 204

Stahler S. W., Palla F., 2005, The Formation of Stars

Stanimirovic S., 1999, PhD thesis, Univ. Western Sydney Nepean

Stanimirovic S., 2002, in Stanimirovic S., Altschuler D., Goldsmith P., Salter C., eds, Astronomical Society of the Pacific Conference Series Vol. 278, Single-Dish Radio Astronomy: Techniques and Applications. pp 375–396 (arXiv:astro-ph/0205329)

Stanimirović S., Lazarian A., 2001, ApJ, 551, L53

Stanimirovic S., Staveley-Smith L., Dickey J. M., Sault R. J., Snowden S. L., 1999a, MNRAS, 302, 417

Stanimirovic S., Staveley-Smith L., Dickey J. M., Sault R. J., Snowden S. L., 1999b, MNRAS, 302, 417

Stanimirovic S., Staveley-Smith L., van der Hulst J. M., Bontekoe T. R., Kester D. J. M., Jones P. A., 2000, MNRAS, 315, 791

Stanimirović S., et al., 2006, ApJ, 653, 1210

Sternberg A., Le Petit F., Roueff E., Le Bourlot J., 2014, ApJ, 790, 10

Stil J. M., et al., 2006, AJ, 132, 1158

Stilp A. M., Dalcanton J. J., Warren S. R., Skillman E., Ott J., Koribalski B., 2013a, ApJ, 765, 136

Stilp A. M., Dalcanton J. J., Skillman E., Warren S. R., Ott J., Koribalski B., 2013b, ApJ, 773, 88

Stutzki J., Bensch F., Heithausen A., Ossenkopf V., Zielinsky M., 1998, A&A, 336, 697

Sun J., et al., 2018, ApJ, 860, 172

Szotkowski S., et al., 2019, ApJ, 887, 111

Tabatabaei F. S., et al., 2007, A&A, 466, 509

Tabatabaei F. S., Krause M., Fletcher A., Beck R., 2008, A&A, 490, 1005

Tamburro D., Rix H. W., Leroy A. K., Mac Low M. M., Walter F., Kennicutt R. C., Brinks E., de Blok W. J. G., 2009, AJ, 137, 4424

Tenorio-Tagle G., Bodenheimer P., 1988, IN: Annual review of A&A. Volume 26 (A89-14601 03-90). Palo Alto, 26, 145

Thilker D. A., Braun R., Walterbos R. A. M., 2002, Seeing Through the Dust: The Detection of HI and the Exploration of the ISM in Galaxies, 276, 370

Thilker D. A., Braun R., Walterbos R. A. M., Corbelli E., Lockman F. J., Murphy E., Maddalena R., 2004, ApJ, 601, L39

Thompson A. R., Moran J. M., Swenson George W. J., 2017, Interferometry and Synthesis in Radio Astronomy, 3rd Edition, doi:10.1007/978-3-319-44431-4.

Tikhonov A., Goncharsky A., Stepanov V., Yagola A. G., 1995, Numerical Methods for the Solution of Ill-Posed Problems. Springer Netherlands, doi:10.1007/978-94-015-8480-7

Tokuda K., et al., 2020, ApJ, 896, 36

Tully R. B., et al., 2013, AJ, 146, 86

Unwin S. C., 1980a, MNRAS, 190, 551

Unwin S. C., 1980b, MNRAS, 192, 243

Utomo D., Blitz L., Falgarone E., 2019a, ApJ, 871, 17

Utomo D., Chiang I. D., Leroy A. K., Sandstrom K. M., Chastenet J., 2019b, ApJ, 874, 141

Vanderplas J., Connolly A., Ivezić Ž., Gray A., 2012, in Conference on Intelligent Data Understanding (CIDU). pp 47 –54, doi:10.1109/CIDU.2012.6382200

Vargas C. J., et al., 2017, ApJ, 839, 118

Vehtari A., Gelman A., Gabry J., 2017, Statistics and Computing, 27, 1413

Veilleux S., Cecil G., Bland-Hawthorn J., 2005, ARA&A, 43, 769

Verley S., Corbelli E., Giovanardi C., Hunt L. K., 2009, A&A, 493, 453

Vilardell F., Ribas I., Jordi C., Fitzpatrick E. L., Guinan E. F., 2010, A&A, 509, A70

Virtanen P., et al., 2020, Nature Methods, 17, 261

Volders L. M. J. S., 1959, Bull. Astron. Inst. Netherlands, 14, 323

Walch S., et al., 2015, MNRAS, 454, 238

Walter F., Brinks E., 1999, AJ, 118, 273

Walter F., Brinks E., de Blok W. J. G., Bigiel F., Kennicutt R. C. J., Thornley M. D., Leroy A., 2008, AJ, 136, 2563

Wang J., TaafMixturefe M., 2015, INFORMS Journal on Computing, 27, 193

Wang Y., et al., 2020a, A&A, 634, A83

Wang Y., et al., 2020b, A&A, 634, A139

Warren S. R., et al., 2012, ApJ, 757, 84

Waskom M., et al., 2017, mwaskom/seaborn: v0.8.1 (September 2017), doi:10.5281/zenodo.883859, https://doi.org/10.5281/zenodo.883859

Watanabe S., 2010, J. Mach. Learn. Res., 11, 3571

Werner M. W., et al., 2004, ApJS, 154, 1

Westmeier T., Braun R., Thilker D., 2005, A&A, 436, 101

Wetzel A. R., Hopkins P. F., Kim J.-h., Faucher-Giguère C.-A., Kereš D., Quataert E., 2016, ApJ, 827, L23

Whitehurst R. N., Roberts M. S., 1972, ApJ, 175, 347

Wilcox R., 2010, Fundamentals of Modern Statistical Methods (2nd Edition). Springer, 233 Spring Street, New York, NY, USA

Willett K. W., Elmegreen B. G., Hunter D. A., 2005, AJ, 129, 2186

Williams B. F., Dalcanton J. J., Stilp A., Dolphin A., Skillman E. D., Radburn-Smith D., 2013, ApJ, 765, 120

Williams B. F., et al., 2015, ApJ, 806, 48

Wilson C. D., Scoville N., 1990, ApJ, 363, 435

Wilson C. D., Walker C. E., 1994, ApJ, 432, 148

Wilson C. D., et al., 2011, MNRAS, 410, 1409

Winkel B., Kerp J., Flöer L., Kalberla P. M. W., Ben Bekhti N., Keller R., Lenz D., 2016, A&A, 585, A41

Wolfire M. G., Hollenbach D., McKee C. F., Tielens A. G. G. M., Bakes E. L. O., 1995, ApJ, 443, 152

Wolfire M. G., McKee C. F., Hollenbach D., Tielens A. G. G. M., 2003, ApJ, 587, 278

Wolfire M. G., Hollenbach D., McKee C. F., 2010, ApJ, 716, 1191

Wong T., et al., 2009, ApJ, 696, 370

Wright M. C. H., Warner P. J., Baldwin J. E., 1972, MNRAS, 155, 337

Yim K., Wong T., Xue R., Rand R. J., Rosolowsky E., van der Hulst J. M., Benjamin R., Murphy E. J., 2014, AJ, 148, 127

Young L. M., Lo K. Y., 1996, ApJ, 462, 203

Young L. M., van Zee L., Lo K. Y., Dohm-Palmer R. C., Beierle M. E., 2003, ApJ, 592, 111

Zhang H.-X., Hunter D. A., Elmegreen B. G., 2012, ApJ, 754, 29

Zheng Y., Peek J. E. G., Werk J. K., Putman M. E., 2017a, ApJ, 834, 179

Zheng Y., Werk J. K., Peek J. E. G., Putman M. E., 2017b, ApJ, 840, 65

de Blok W. J. G., Walter F., 2006, AJ, 131, 363

de Jager G., Davies R. D., 1971, MNRAS, 153, 9

de Vaucouleurs G., de Vaucouleurs A., Corwin Herold G. J., Buta R. J., Paturel G., Fouque P., 1991, Third Reference Catalogue of Bright Galaxies

van de Hulst H. C., Raimond E., van Woerden H., 1957, Bull. Astron. Inst. Netherlands, 14, 1

van der Marel R. P., Alves D. R., Hardy E., Suntzeff N. B., 2002, AJ, 124, 2639

Appendix A: Supplemental power spectrum figures

This appendix includes the complete set of power spectrum fits from Chapter 6. For each band, the fits to the power spectra are shown with the data at its original PSF and convolved to a moderately-sized Gaussian, as defined by Aniano et al. (2011). The fitted parameter values are given in Table 6.2.

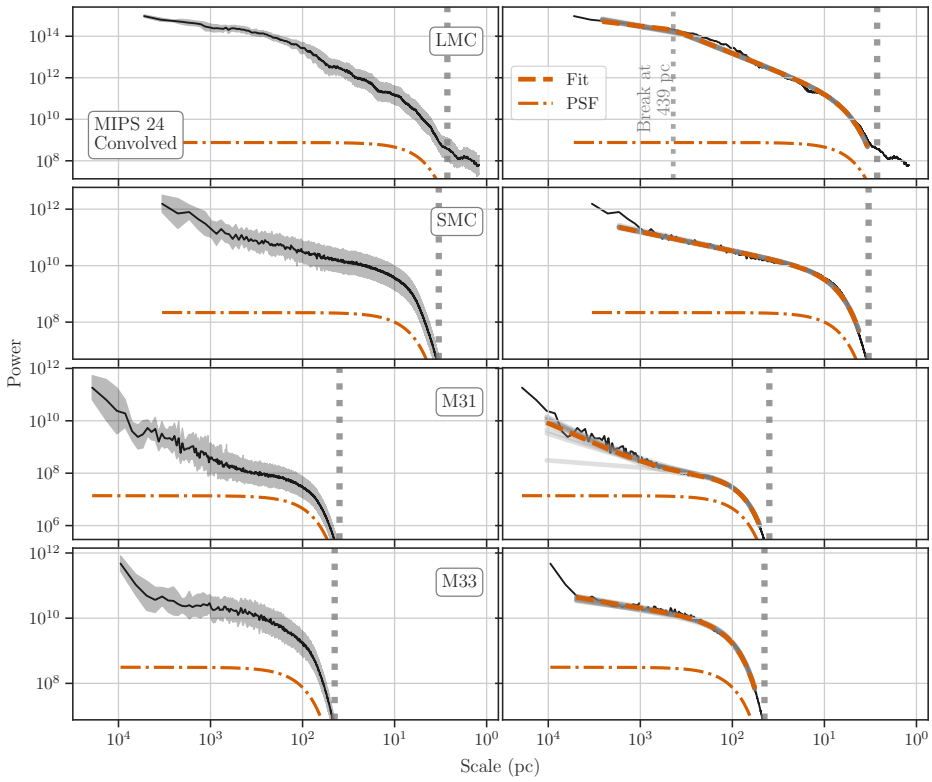

Figure A.1: 1D power spectra of MIPS 24 μm maps convolved to a Gaussian PSF (left column) and the fitted models (right column; Table 6.2).

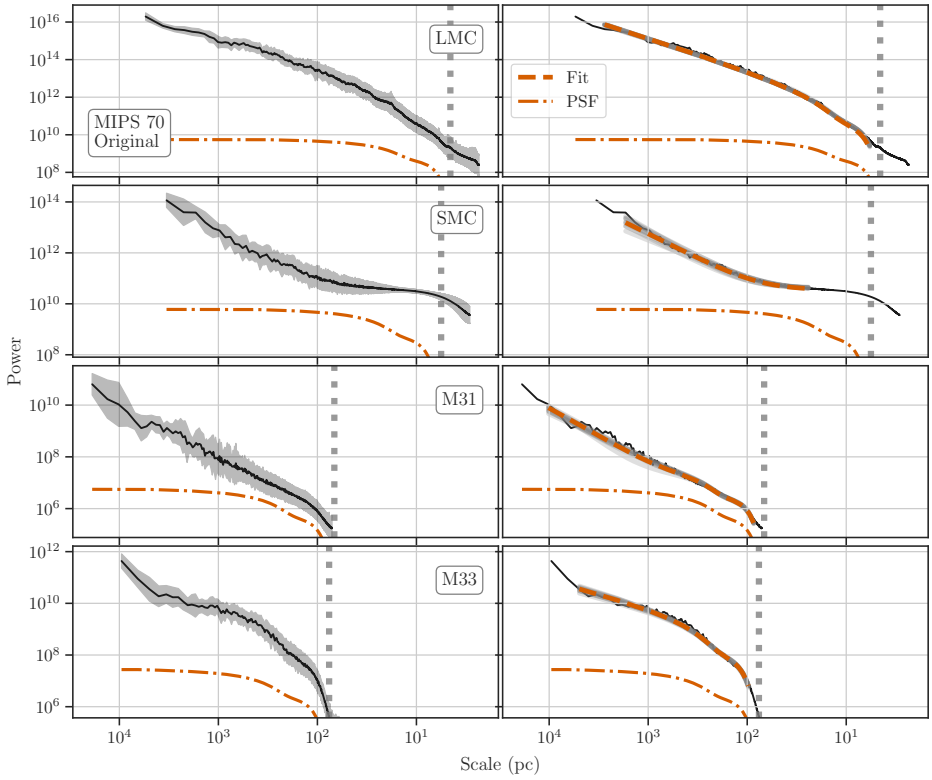

Figure A.2: 1D power spectra of MIPS 70 μm maps with the original PSF (left column) and the fitted models (right column; Table 6.2).

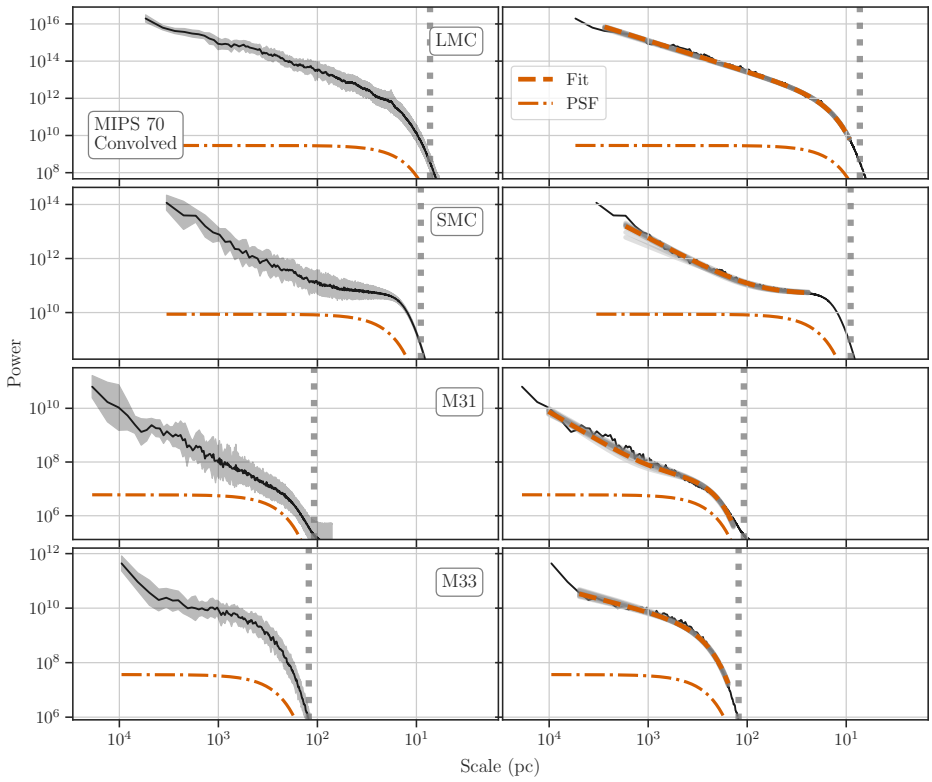

Figure A.3: 1D power spectra of MIPS 70 μm maps convolved to a Gaussian PSF (left column) and the fitted models (right column; Table 6.2).

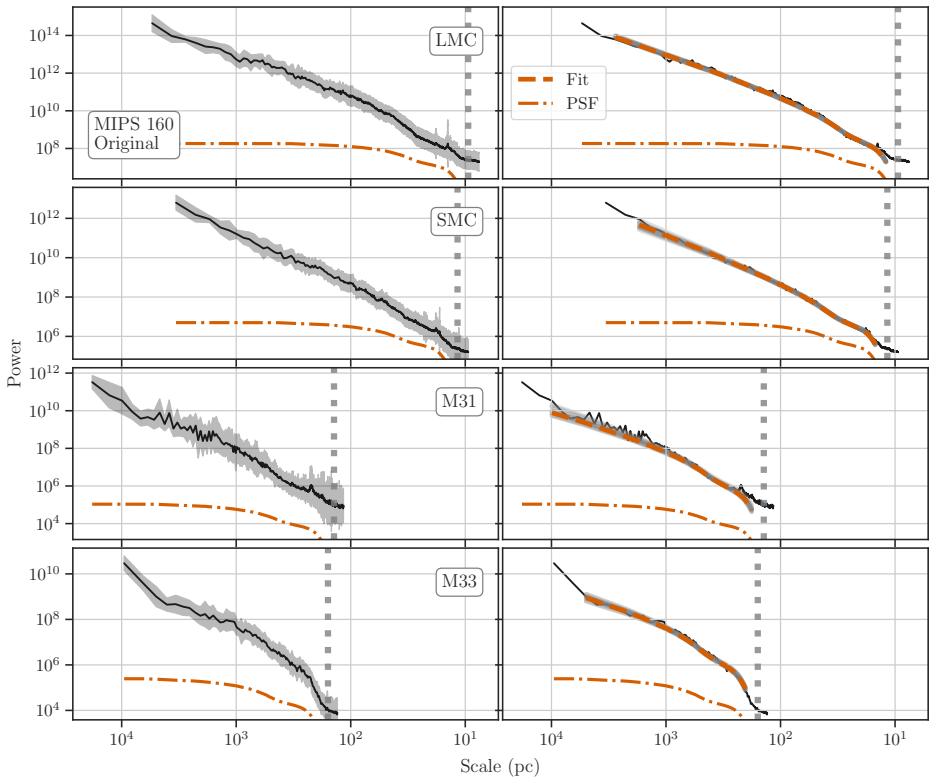

Figure A.4: 1D power spectra of MIPS 160 μm maps with the original PSF (left column) and the fitted models (right column; Table 6.2).

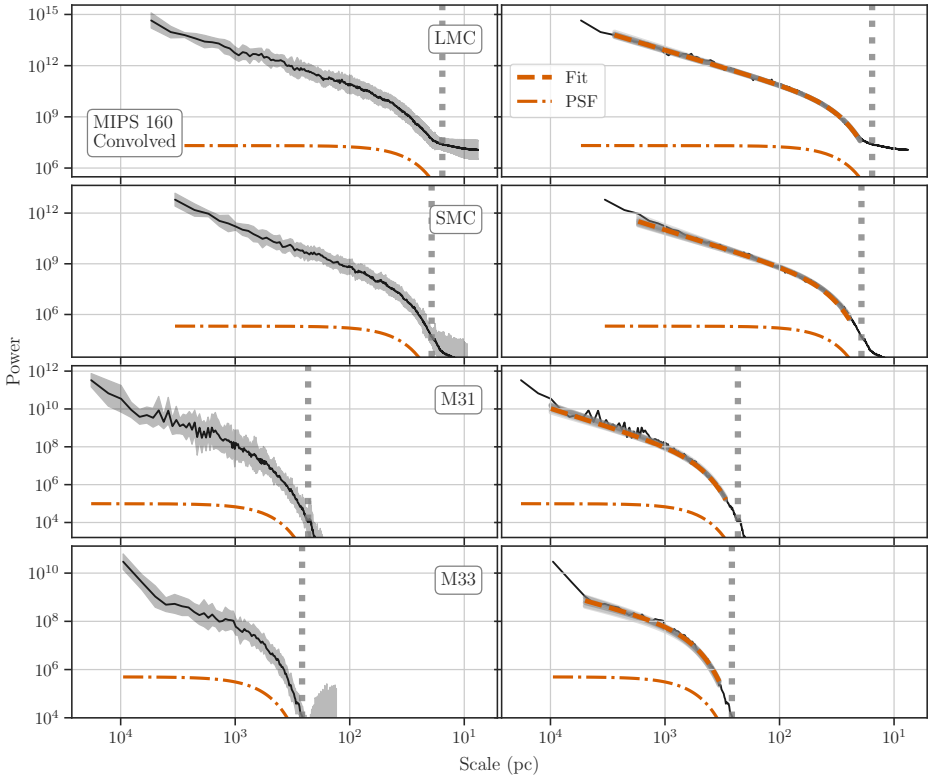

Figure A.5: 1D power spectra of MIPS 160 μm maps convolved to a Gaussian PSF (left column) and the fitted models (right column; Table 6.2).

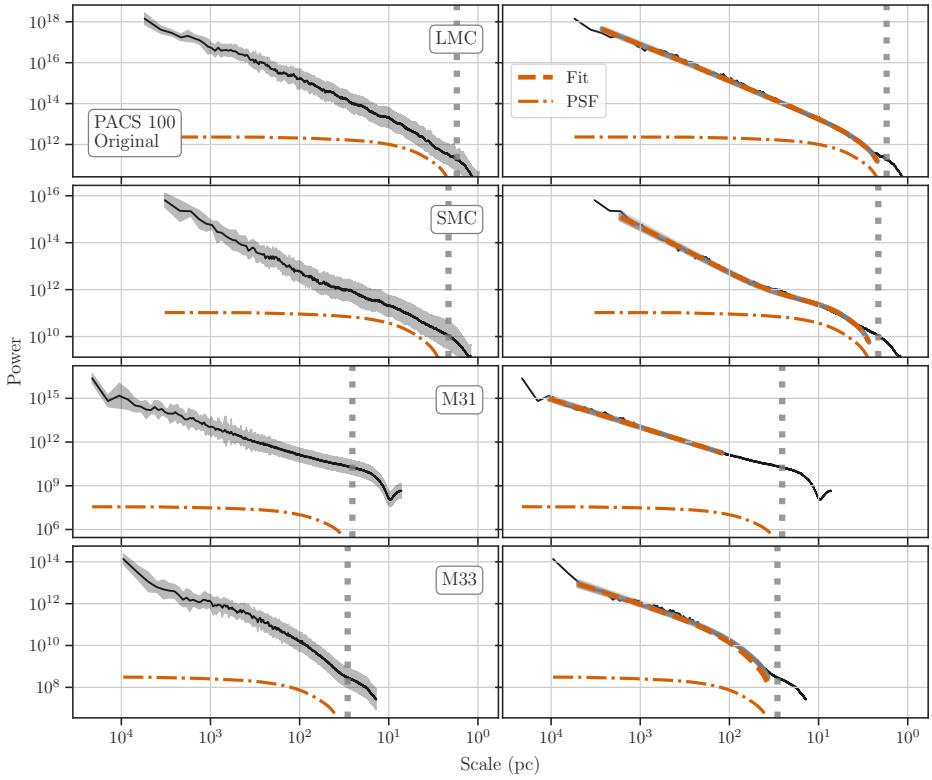

Figure A.6: 1D power spectra of PACS 100 μm maps with the original PSF (left column) and the fitted models (right column; Table 6.2).

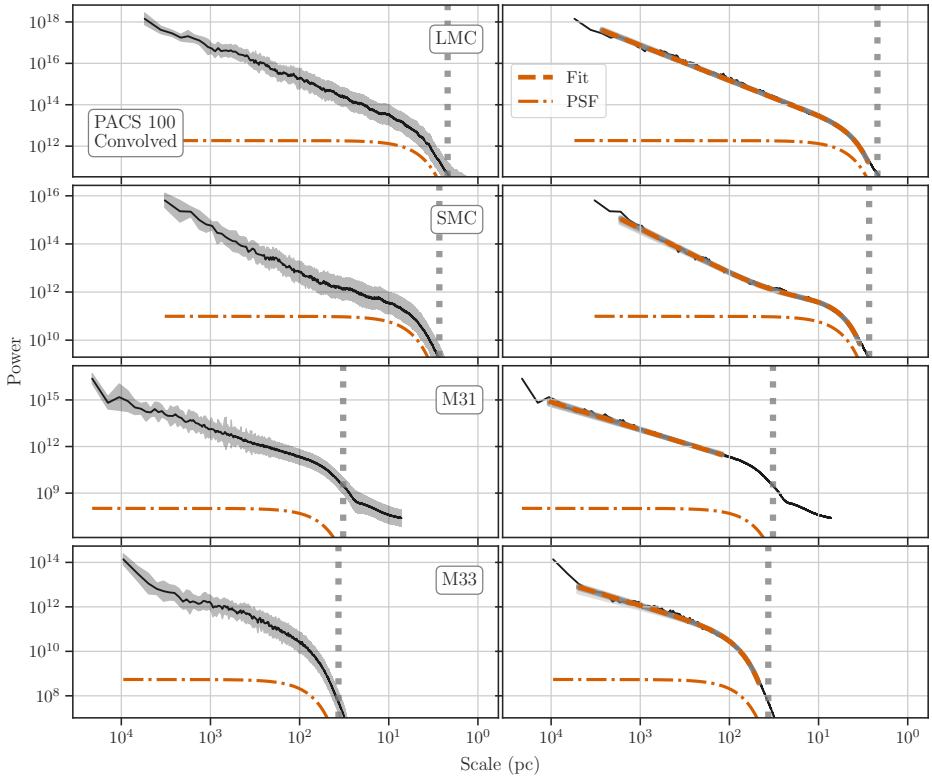

Figure A.7: 1D power spectra of PACS 100 μm maps convolved to a Gaussian PSF (left column) and the fitted models (right column; Table 6.2).

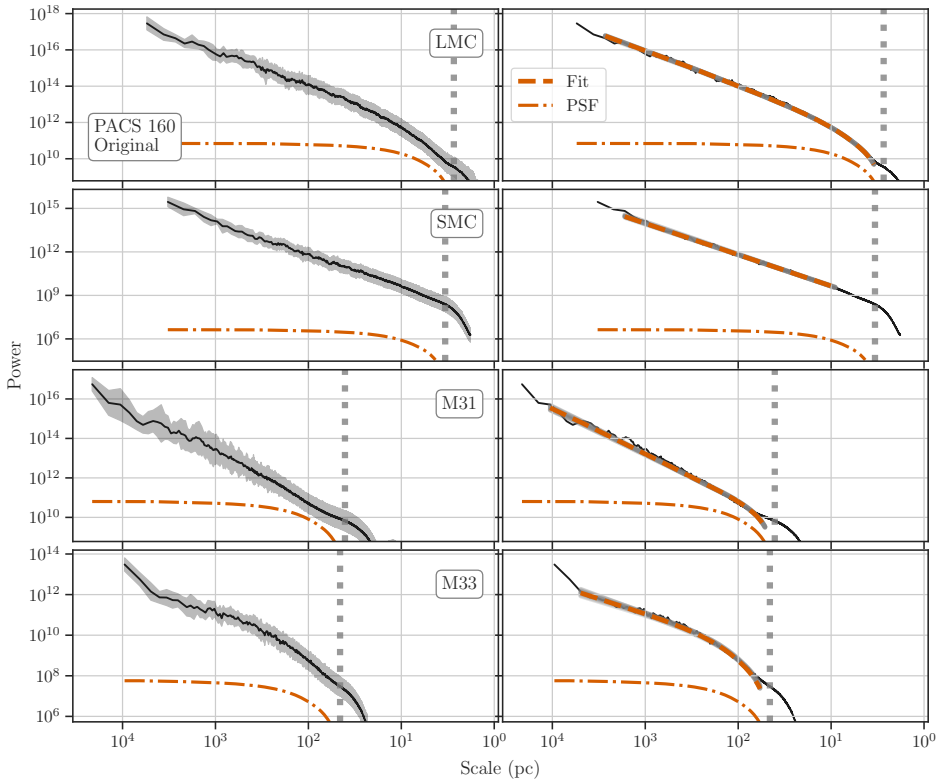

Figure A.8: 1D power spectra of PACS 160 μm maps with the original PSF (left column) and the fitted models (right column; Table 6.2).

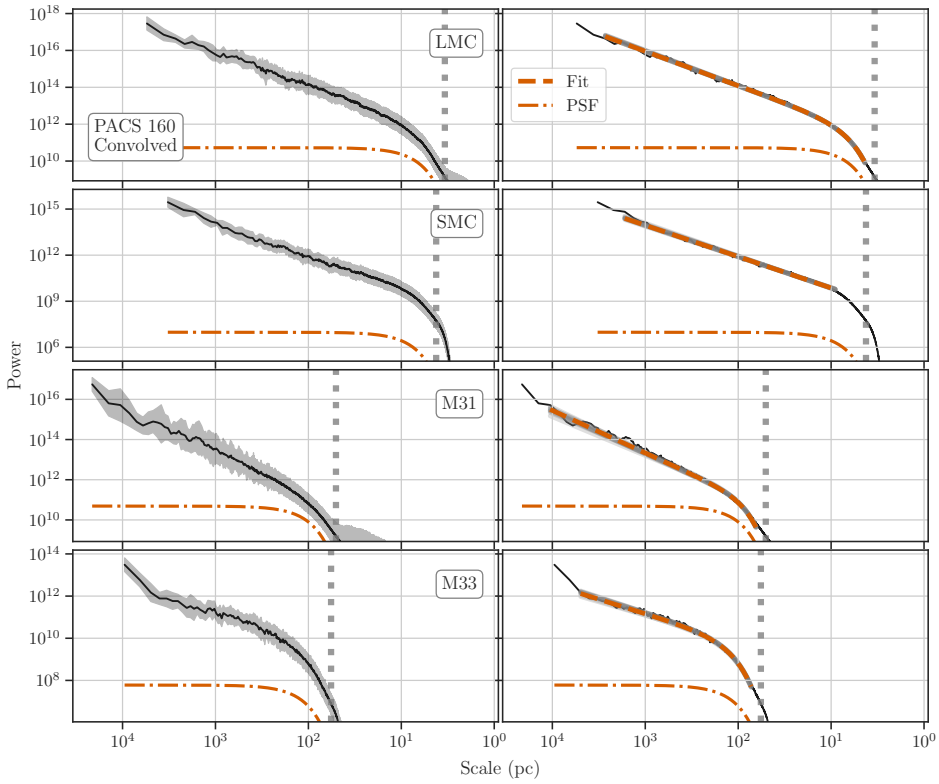

Figure A.9: 1D power spectra of PACS 160 μm maps convolved to a Gaussian PSF (left column) and the fitted models (right column; Table 6.2).

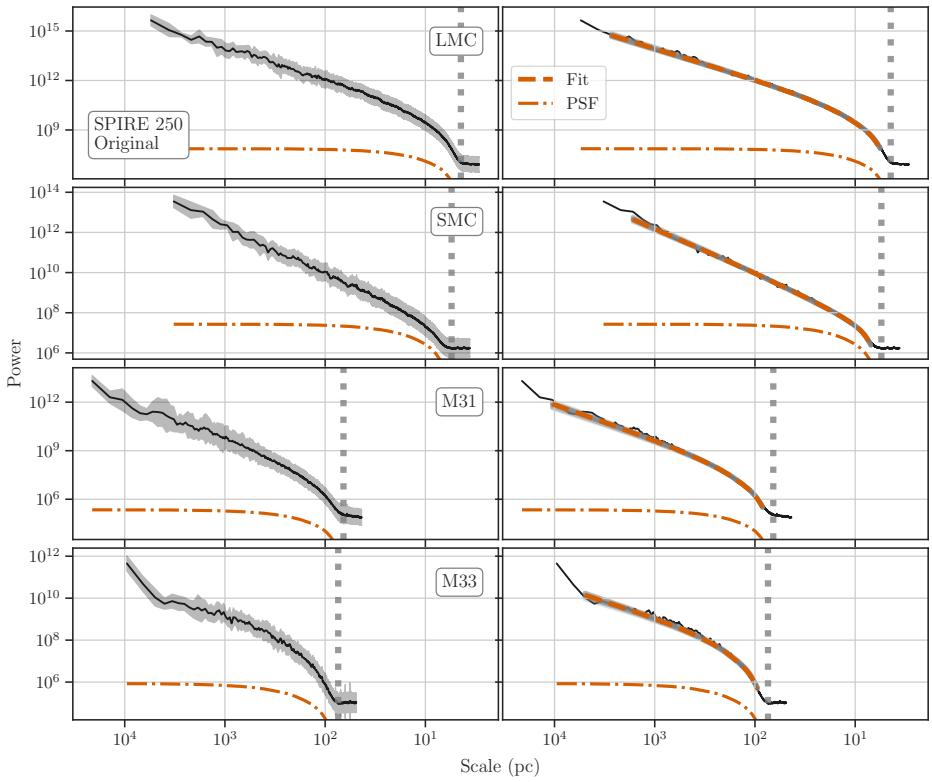

Figure A.10: 1D power spectra of SPIRE 250 μm maps with the original PSF (left column) and the fitted models (right column; Table 6.2).

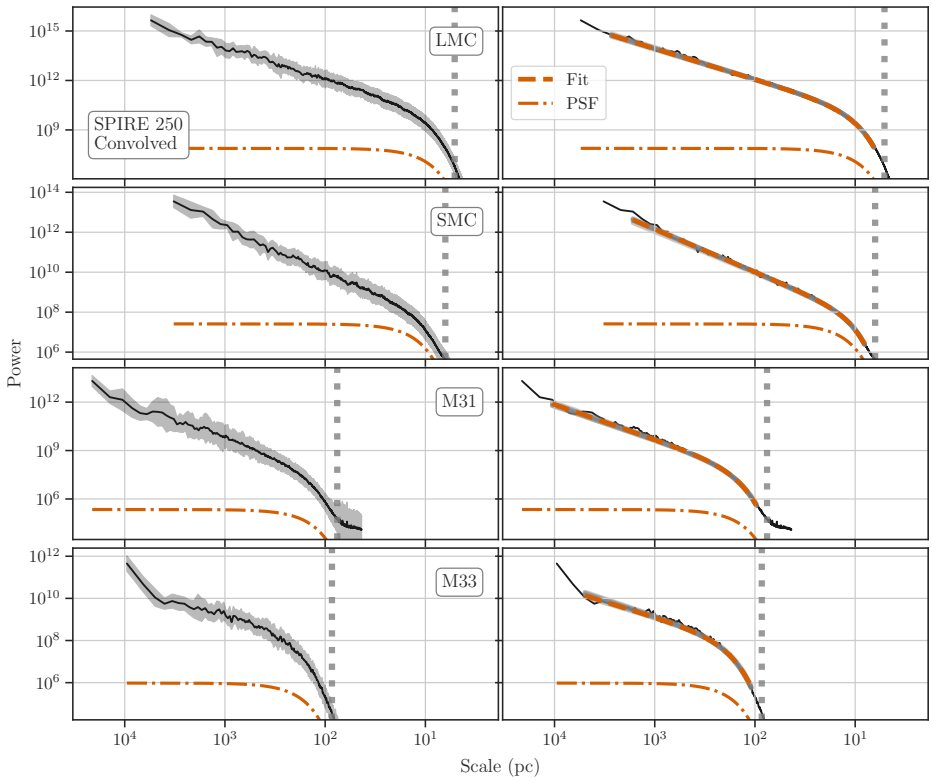

Figure A.11: 1D power spectra of SPIRE 250 μm maps convolved to a Gaussian PSF (left column) and the fitted models (right column; Table 6.2).

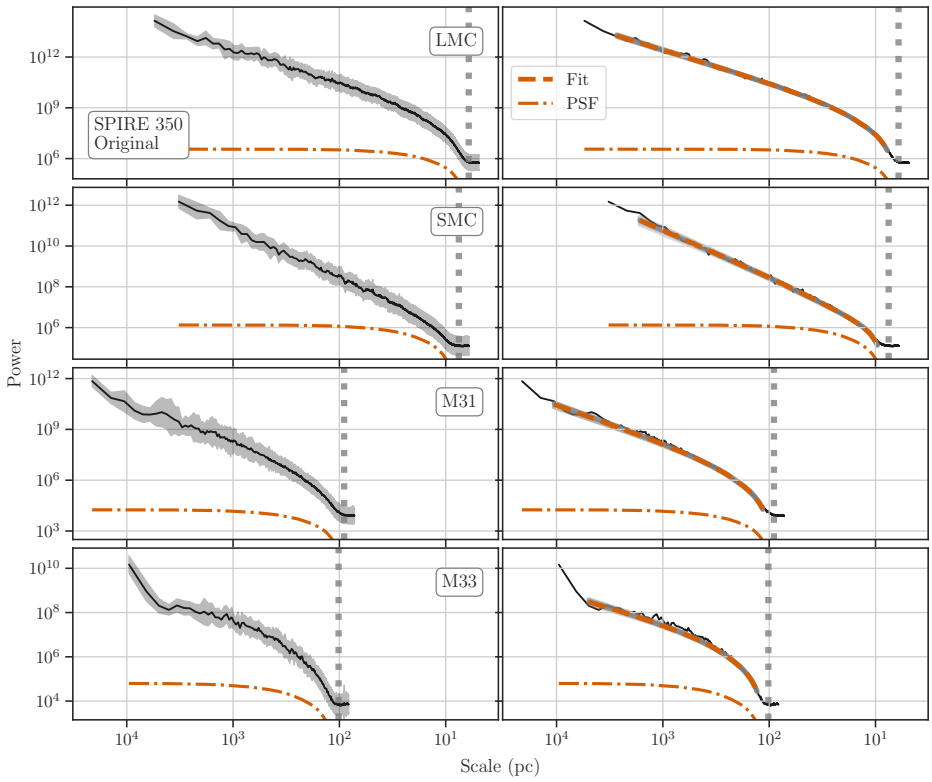

Figure A.12: 1D power spectra of SPIRE 350 μm maps with the original PSF (left column) and the fitted models (right column; Table 6.2).

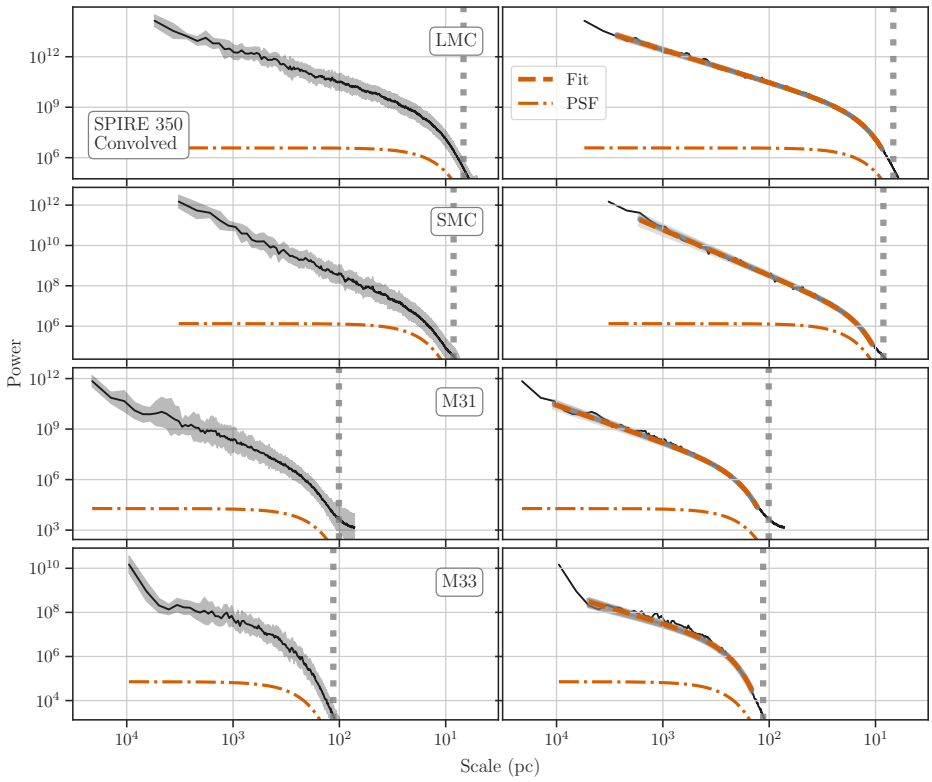

Figure A.13: 1D power spectra of SPIRE 350 μm maps convolved to a Gaussian PSF (left column) and the fitted models (right column; Table 6.2).

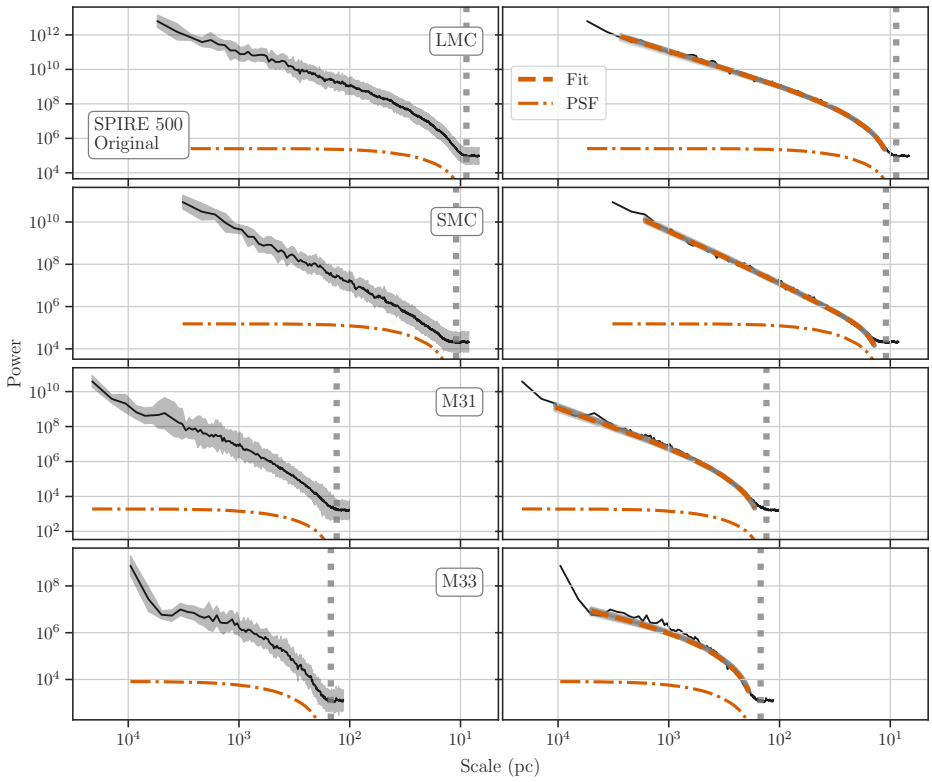

Figure A.14: 1D power spectra of SPIRE 500 μm maps with the original PSF (left column) and the fitted models (right column; Table 6.2).

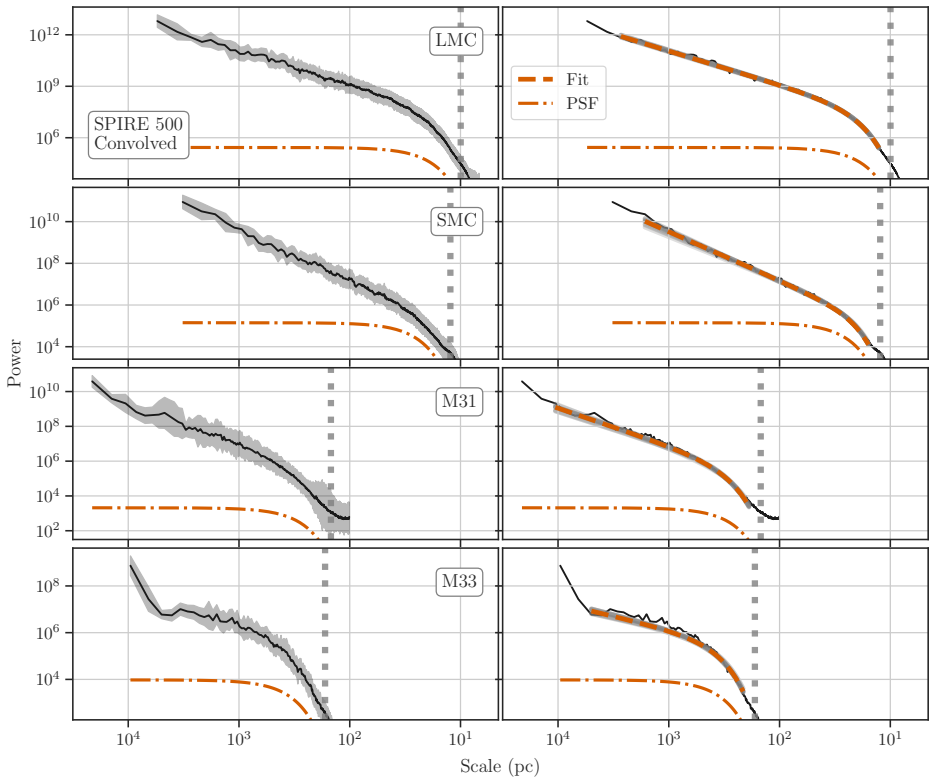

Figure A.15: 1D power spectra of SPIRE 500 μm maps convolved to a Gaussian PSF (left column) and the fitted models (right column; Table 6.2).