

UNVEILING THE DIFFUSE, NEUTRAL INTERSTELLAR MEDIUM: ABSORPTION
SPECTROSCOPY OF GALACTIC HYDROGEN

by

CLAIRE ELIZABETH MURRAY

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

DOCTOR OF PHILOSOPHY
(ASTRONOMY)

at the
UNIVERSITY OF WISCONSIN–MADISON
2017

Date of final oral examination: 4 August 2017

The dissertation is approved by the following members of the Final Oral Committee:

Snežana Stanimirović, Professor, Astronomy
Ellen G. Zweibel, Professor, Astronomy
Eric M. Wilcots, Professor, Astronomy
John S. Gallagher, Professor, Astronomy
W. Miller Goss, Emeritus Scientist, Astronomy
Lindsay Elizabeth Stovall, Assistant Professor, Mathematics

Abstract

The formation of stars and evolution of galaxies depends on the cycle of interstellar matter between supernova-expelled plasma and molecule-rich gas. At the center of this cycle is multiphase neutral hydrogen (HI), whose physical conditions provide key ingredients to theoretical models. However, constraints for HI properties require measurements of gas emission *and* absorption which have been severely limited by previous observational capabilities.

In this thesis, I present the largest survey of Galactic HI absorption ever undertaken with the Karl G. Jansky Very Large Array (VLA). The survey, 21-cm Spectral Line Observations of Neutral Gas with the VLA (21-SPONGE), is a statistical study of HI in all phases using direct absorption measurements. Leveraging novel calibration techniques, I demonstrate the capability of the VLA to detect a significant sample of 21 cm absorption lines from warm, diffuse HI.

To maximize observational sensitivity, I stack the 21-SPONGE spectra and detect a pervasive signature of the warm neutral medium in absorption. The inferred excitation (or spin) temperature is consistent with existing estimates, yet higher than predictions from theoretical models of collisional HI excitation. This suggests that radiative feedback via resonant scattering of Ly α photons, known as the Wouthuysen-Field effect, is influential with important implications for cosmological 21 cm observations.

Next, I compare 21-SPONGE with synthetic HI spectra from 3D numerical simulations using a new, objective decomposition and radiative transfer tool. I quantify the recovery of HI structures and their properties by Gaussian-fitted 21 cm spectral lines for the first time. I find that 21 cm absorption line shapes are sensitive to simulated physics, and demonstrate that my

analysis method is a powerful tool for diagnosing neutral ISM conditions.

Finally, I compare properties inferred from synthetic spectra with “true” simulation results to construct a bias correction function for estimating HI properties. I apply this correction to the mass distribution of HI as a function of temperature from 21-SPONGE, and find a significant fraction of thermally unstable gas. This confirms that non-steady radiative and dynamical processes, such as turbulence and supernovae, have a strong influence on the thermodynamic state of the ISM.

For Mary Alice Duenwald and Bruce Alan Murray

Acknowledgments

Often asked why I chose to study astronomy, I reply that there are no more exciting mysteries to investigate than the origins of Earth, the Milky Way galaxy and the Universe. It has been an extraordinary privilege to spend time and energy on these pursuits, and I am proud to present this dissertation as a small contribution. However, none of it would have been possible without the support of friends, family and colleagues around the world, and no length of acknowledgments can contain my gratitude. Nevertheless, I will do my best to express it here.

First and foremost, I would like to thank my advisor, Snežana Stanimirović. On day one, you trusted me to dive into an enormous pile of data full of possibilities, and I am as excited, humbled and grateful for this opportunity now as I was then. This experience allowed me to build the technical and creative skills that will serve as the foundation for the rest of my career. Thank you for your unwavering dedication, for encouraging me to be the best I could be, and for sending me all over the world to observe and learn. Thank you especially for introducing me to the wonderful world of the interstellar medium and its myriad mysteries. I look forward to exploring them with you as part of many future “sponge baby” projects.

To my unofficial advisor Miller Goss, thank you for your enthusiastic support. I am honored to be counted among your students. Thank you for patiently explaining the intricacies of data reduction from afar, and for generously sharing the rich history of observations which formed the foundation of our survey. You called me a “black-belt” radio astronomer when I was in my first month of graduate school, and I’ve worked to live up to that title ever since.

To my collaborators along the way, thank you for your guidance. I have come a long way

from the nervous student I used to be thanks in large part to your confidence in me, and I am proud to follow in your footsteps. Carl Heiles, I remain baffled by your endless patience in teaching me how to use IDL many years ago. Thank you for the many wonderful conversations we've had around the world – I remember particularly fondly discussing dark molecular gas with you over Thai food on Thanksgiving in Sydney. John Dickey and Naomi McClure-Griffiths, thank you for welcoming me to Australia and teaching me the ropes at the ATCA. I am thrilled to be part of the GASKAP team, and I look forward to tackling the puzzles of this project with you.

To the past and present members of my research group in Madison, thank you for teaching me everything I know. Min, the example you set in every aspect of research, from work ethic to scientific curiosity, has been an inspiration that I will follow every day. Bob, thank you for opening my eyes to new tools and ideas, and for revolutionizing the 21-SPONGE project. Working with GaussPy has been a rewarding adventure, and I am grateful for your trust in continuing to develop this amazing tool. Elijah, thank you for inspiring me to make better plots and for your sincerely heroic efforts in helping me to learn new programming skills. You are a true font of joy, and I look forward to following your future adventures in life (and in home brewing!). Brian, 21-SPONGE would not be complete without you. Thank you for taking on the challenge of AIPS data reduction with such aplomb, and for your invaluable input throughout. To my students along the way, Jesse, Nick, Al, Jake and Skylyn, thank you for the opportunity to grow as a mentor; it has been a pleasure to work with you and watch you succeed.

For inspiring me to start down this path in the first place, I am forever indebted to my generous and caring advisors at Carleton College, Cindy Blaha and Joel Weisberg. Working with you at Goodsell Observatory is why I fell in love with the stars (without doubt, the best job on campus is the Astro 110 “lab ass”). Thank you Cindy for giving me my first experience in research and for welcoming me to the Olin family so warmly. Thank you Joel for fatefully replying to my list of summer research applications with “what, no radio?”, and supporting me wholeheartedly and without reservation throughout my career.

To my UW astronomy family, thank you for defying all norms and making graduate school so astonishingly fun. I couldn't ask for a better classmate than Ben; thank you for battling so valiantly through waves of physics and astronomy with me over the past six years. In repayment, someday I'll pass along my expertise in supermoon-lit sand drawing. Nick, I am so grateful to have been your officemate. You taught me many things (Spotted Cow is boring! Bikes are cool!), and I'd sit next to a pilla' any day to be around you and your infectious enthusiasm. Natalie, you are an idol of astronomy and it has been an honor to look up to you and call you my friend. Thank you for your tireless guidance and support — I am counting on a lifetime of cribbage and Manhattan dates. Jenna, thank you for being the co-manager of my dreams. I'm so happy that you haven't gotten rid of me quite yet! Arthur, you are *the* elite professional. Thank you for always stopping by my office to talk, and for being down to get ice cream no matter the weather. Chris, you are a wonderful friend (when you're not laughing at the Mets). I hope to watch baseball and exchange books with you indefinitely. John, you never fail to remind me what boundless enthusiasm looks like; thank you for the support and inspiration. Corey, I won't soon forget our sabtoeur coup, and I can't thank you enough for always cheering me on. Emily, I count our many nights of games and grog as essential for survival. I can't wait to see where your triumphs take you! Anna, it has been a joy to share an office with you. Thank you not only for watering all the plants in the room, but for lending a caring ear during this last, stressful year.

To my friends in Madison, thank you for reminding me that life outside of work is not only important but delightful. Katrina, I don't know where I would be without you. Thank you for being MMB and inspiring me a hundred times over to live life to its fullest, complete with the occasional flabongo. Anna, you are the most fascinating and wonderful friend and I still can't believe my luck to have met you. Thank you for making me smile so often. Emily, I have lived in awe of your infinite talents since Love House; thank you for sharing them with me so generously. Josh, thank you for being the undisputed best manager ever. Playing softball and soccer with you is one of my absolute favorite parts of life in Madison. And on that note, thank you to the many

generations of Extragalacticos, Infrared Sox, Infield Flies, Baltimore Factorioles and Magnetic Monopoles. You guys are the best, too.

To my family, thank you for setting the bar so high and believing in me to reach it. Mom, I can't wait to continue watching the Mets and traveling the world with you. Thank you for supporting me so constantly and reminding me to take notes. Dad, thank you for teaching me how to cook and how to never quit. I am so lucky to have you to look up to. Nick, you inspire me to break out of my shell. Thank you for making me laugh so effortlessly and for understanding me like no one else can.

Finally, neither this thesis nor anything else would be possible without David, my person. Thank you for keeping bandaids in your wallet for when my shoes hurt my feet. Thank you for sharing my love of baseball. Thank you for (begrudgingly) learning to love a fluffy and ridiculous cat. Thank you for reading every draft. Thank you for believing in me when I refused to. Thank you for accepting my incompetence and teaching me how to fix it. Thank you for showing me what it means to follow your dreams. I owe it all to you.

Contents

Abstract	i
Dedication	iii
Acknowledgments	iv
Contents	viii
List of Tables	xiv
List of Figures	xv
1 Introduction	1
1.1 Theory of the atomic ISM	3
1.2 Observations of the atomic ISM	8
1.2.1 Measuring Physical Properties of HI	10
1.3 The 21-SPONGE Survey	13
1.4 Current Progress and Open Questions	13
1.4.1 What is the temperature of the WNM?	13
1.4.2 How well do simulations reproduce observed HI properties?	16
1.4.3 What is the mass fraction of HI as a function of temperature?	18
1.5 Thesis Outline	19
References	21
2 The 21-SPONGE Survey: Techniques and Initial Results	24
Abstract	25

2.1	Introduction	26
2.1.1	Previous Work	27
2.1.2	21-SPONGE	30
2.2	Data Processing	32
2.2.1	Observations	32
2.2.2	Data reduction	34
2.2.3	Optical depth comparison	36
2.2.4	Matching emission profiles	37
2.2.5	Stray radiation	38
2.3	Analysis	40
2.3.1	Gaussian Fitting	40
2.3.2	Absorption Noise Spectra	44
2.3.3	Component Selection: Absorption	44
2.3.4	Component Selection: Emission	46
2.4	Selected Examples	47
2.4.1	3C138	48
2.4.2	3C286	49
2.4.3	4C16.09	49
2.4.4	PKS0531+19	50
2.5	Properties of Cold and Warm Neutral Gas	50
2.5.1	Temperature	50
2.5.2	Column Density	52
2.6	Comparison with previous surveys	54
2.6.1	Line of sight properties	58
2.7	Comparison with synthetic observations	60

2.8 Comparison of methods for estimating spin temperature from HI absorption and emission spectra	62
2.8.1 One phase per channel T_s	62
2.8.2 Line of sight harmonic mean T_s	64
2.8.3 The Slope Method	64
2.8.4 Applying methods to 21-SPONGE data	65
2.9 Summary and Conclusions	67
References	69
3 The excitation temperature of the Warm Neutral Medium	73
Abstract	74
3.1 Introduction	75
3.2 Observations and Data Processing	77
3.2.1 Observations	77
3.2.2 Derivation of T_s for individual components	78
3.3 Stacking analysis of HI absorption and emission spectra	80
3.4 Discussion	84
3.5 Conclusions	86
References	86
4 Recovering Interstellar Gas Properties with HI Spectral Lines	89
Abstract	90
4.1 Introduction	91
4.2 Observations	95
4.3 Numerical simulations	96
4.4 Gaussian Decomposition with AGD	99
4.5 Assessing the power of the Gaussian-fitting method with synthetic H I spectra	100

4.5.1	Defining gas structures in position and velocity	100
4.5.2	Matching gas structures with H I absorption lines	103
4.5.2.1	Matching H I absorption lines with H I emission lines	104
4.5.3	Quantifying Match Completeness	105
4.5.4	Observed vs. “True” Gas Properties	107
4.5.4.1	Spin Temperature	109
4.5.4.2	Column Density	111
4.6	Comparing Real and Synthetic H I Spectra	113
4.6.1	Number of components along LOS	114
4.6.2	Properties of H I absorption components	116
4.6.2.1	Comparison with previous studies	117
4.6.2.2	Influence of local box simulations	117
4.6.2.3	Minimum CNM temperature?	119
4.6.2.4	Role of the WF effect	120
4.6.3	Properties of H I emission components	121
4.6.3.1	Boundaries in brightness temperature	122
4.6.3.2	Broad emission components	123
4.6.4	Observed Spin Temperature	124
4.6.4.1	Per-channel T_s	129
4.6.5	Observed Column Density	130
4.7	Summary and Conclusions	131
	References	135
5	The 21-SPONGE Survey II: The mass distribution of Galactic HI as a function of temperature	139
	Abstract	140
5.1	Introduction	141

5.2 Data	143
5.2.1 Observing and Data Reduction	143
5.2.2 Matching H I Emission Profiles	146
5.2.3 Absorption Noise Spectra	147
5.3 Integrated Properties	147
5.3.1 Line of Sight Properties	147
5.4 Gaussian Fitting	148
5.4.1 Synthetic H I Fits	157
5.5 Results	159
5.5.1 Fitted Parameters	159
5.5.2 Mass distribution of H I as a function of T_s	161
5.5.3 Correcting for observational bias	161
5.6 Discussion	163
5.6.1 Significant fraction of thermally unstable HI	166
5.6.2 Broad CNM distribution	168
5.6.3 Lack of absorption by WNM with $T_s > 2000$ K	168
5.7 Summary	169
References	171
6 Summary and Future Work	173
6.1 Future Work	177
References	180
A 21-SPONGE Calibration Techniques	181
A.1 Calibration Issues	181
A.1.1 Contamination by HI at $v \sim 0$ km s ⁻¹	181
A.1.1.1 Frequency-Switched Calibration Setup	182

A.1.2 Example: Observing Galactic H I Absorption towards 4C12.50	183
A.1.2.1 Bandpass Calibration of 3C286 via Frequency-Switching	183
A.1.3 Spacing of Frequency-Switched Subbands	186
A.1.4 Increased System Temperature	187
A.1.5 Resolved HI Emission	187
A.2 59 kHz Ripple	189
A.3 Time Averaging	193
References	195
B 21-SPONGE Continuum Images and Fitted Spectral Line Properties	197

List of Tables

2.1	Survey Comparison	28
2.2	VLA Observation Information	32
2.3	Line of sight column density parameters	56
3.1	Source Information	79
4.1	AGD Summary	97
5.1	VLA Observation Information	144
5.2	Mass Fractions as a function of T_s	163
A.1	Bandpass Solution Parameters	192
B.1	Fitted Parameters	202

List of Figures

1.1	Cross section of HI cloud (McKee & Ostriker 1977)	4
1.2	Phase diagram (Wolfire et al. 2003)	5
1.3	Effect of turbulence on phase distribution (Audit & Hennebelle 2005)	7
1.4	21 cm emission and absorption profiles (Dickey & Lockman 1990)	9
1.5	Detection of WNM in absorption (Dwarakanath et al. 2002)	15
1.6	Synthetic 21 cm spectral lines (Hennebelle et al. 2007; Kim et al. 2014)	17
2.1	Optical depth comparison	35
2.2	Example improvement following inclusion of additional components	41
2.3	Selection of example 21-SPONGE absorption spectra	42
2.4	Example broad, high-temperature absorption component	43
2.5	Turbulent Mach number histogram	52
2.6	Properties of fitted Gaussian components	53
2.7	Spin temperature comparison with Millennium Survey	54
2.8	CNM fraction comparison with previous surveys	55
2.9	CNM fraction vs. harmonic mean spin temperature	63
2.10	Temperature derivation method comparison	66
3.1	Stacked residual absorption and emission signal	81
3.2	Stacked absorption equivalent width and spin temperature PDFs	82
3.3	Comparison with previous WNM T_s measurements	84

4.1	Distribution of 21-SPONGE sources	93
4.2	Example synthetic HI spectra from Kim et al. (2014)	101
4.3	Number of gas structures recovered by AGD spectral lines	102
4.4	Simulated versus inferred T_s for synthetic LOS	108
4.5	Example LOS where AGD method overestimates T_s	108
4.6	Simulated versus inferred column density for synthetic LOS	110
4.7	Ratio of simulated to inferred T_s and $N(\text{HI})$ versus LOS properties	111
4.8	Example HI pairs from 21-SPONGE with AGD fits and matching results	114
4.9	Effect of viewing angle on AGD number statistics	115
4.10	Comparison of Gaussian parameters fit to HI absorption	118
4.11	Line width versus amplitude for AGD fits to HI absorption	118
4.12	Comparison of Gaussian parameters fit to HI emission	120
4.13	Line width versus amplitude for AGD fits to HI emission	121
4.14	Inferred T_s for all matched HI features	125
4.15	Line width versus amplitude for HI absorption colored by T_s	126
4.16	Per-channel T_s	127
4.17	Inferred $N(\text{HI})$ for all matched HI features	128
4.18	Line width versus amplitude for AGD fits to HI absorption, including LAB	138
5.1	Histogram of RMS uncertainties in 21-SPONGE	147
5.2	21-SPONGE spectral line pairs and fits	149
5.3	Fit parameters from AGD analysis of 21-SPONGE	158
5.4	Gas fraction vs. spin temperature	159
5.5	Corrected gas fractions using simulated H I spectra	162
5.6	Histogram of T_s for 21-SPONGE and HT03	167
6.1	Stacked absorption properties as a function of latitude	178

A.1	21-SPONGE instrument configuration	182
A.2	Example BP solutions from frequency-switched bands	184
A.3	Combined frequency switched BP solution	184
A.4	21-SPONGE instrument configuration with added subbands	185
A.5	Effect of subband spacing on BP solution	185
A.6	Effect of combining multiple subbands on BP solution	187
A.7	H _I emission structure resolved on short baselines	188
A.8	Effect of baseline removal on resolved H _I emission	189
A.9	Effect of resolved H _I emission on H _I absorption structure	189
A.10	Bandpass solution analysis via Fourier Transform	190
A.11	FT bandpass model residuals	191
A.12	Cascaded filter response for 21SPONGE target subband	193
A.13	Example bandpass solutions	195
A.14	Improvement in bandpass noise due to time averaging	196
B.1	21-SPONGE continuum images.	198

Chapter 1

Introduction

The evolution of galaxies in the Universe is determined by where, when and how stars form. Throughout cosmic time, star formation is fueled by the intervening gas and dust known as the interstellar medium (ISM). During their lifetimes, stars enrich the ISM from which they formed with elements such as carbon, nitrogen, and oxygen via nuclear fusion reactions, and expel this enriched material back into the ISM at the end of their lives. This expelled, shocked gas is ionized and hot, with $T \sim 10^5$ to 10^7 K, but subsequent star formation from gravitational collapse typically takes place in cold gas with $T < 40$ K. The conversion between hot and cold gas phases in the ISM, passing through essential stages consisting primarily of neutral atomic hydrogen (HI), is a crucial bottleneck between gas reservoirs and star formation in galaxies.

As the most abundant atomic species in the ISM ($\sim 78\%$), HI lends insights into the influence of radiative and dynamical processes at work within galaxies (e.g., Governato et al. 2010; Ostriker et al. 2010; Hopkins et al. 2013). Theoretical simulations aiming to predict distributions of gas and stars in galaxies throughout cosmic time must develop models for unresolved ISM physics in order to match observed HI properties. Within these models, prescriptions for interstellar heating, cooling, and star formation are still highly uncertain (e.g., Tasker & Bryan 2008; Christensen et al. 2012), and observations are far from complete.

Furthermore, HI provides the main fuel reservoir for the formation of the molecular gas complexes which host star formation. Numerous observational studies, covering spatial scales from individual clouds to distant galaxy disks, show that the star formation rate scales directly with molecular gas content (Bigiel et al. 2008; Leroy et al. 2013). In addition, the star formation efficiency of giant molecular clouds — the birth places of stars — is roughly constant in disparate galaxy environments (Schruba et al. 2011). This suggests that once molecular gas is formed, it is converted to stars with similar efficiency in both dwarf and spiral galaxies. Understanding the conditions required to initiate or inhibit star formation within galaxies therefore requires a detailed understanding of how the ISM transitions between atomic and molecular gas phases.

Despite the importance of HI for building stellar nurseries and reflecting the effects of

radiative and dynamical processes in the ISM, a comprehensive observational model of basic conditions within HI throughout the Milky Way (let alone other galaxies!) is still missing. This thesis is an observational study of the multiphase properties of Galactic HI and the implications for our theoretical understanding of the physics of the ISM. This work is enabled by the largest, highest-sensitivity survey for HI absorption to date at the Karl G. Jansky Very Large Array (VLA) and novel analysis techniques. In the following sections I review the theory behind the atomic ISM, observational progress to date, and the open questions addressed in this thesis.

1.1 Theory of the atomic ISM

The balance between heating and cooling from radiative and dynamical sources in galaxies generates a complex structure of pressure, temperature and density within the atomic ISM. The original theory to explain the properties of HI emerged in the 1960s. Field (1965) showed that thermal equilibrium within a static, uniform atomic medium is unstable under the influence of thermal conduction, magnetic fields, rotation and gravity. This thermal instability allows structures to condense within the medium if energy losses increase with decreasing temperature (Parker 1953), and is a natural consequence of pressure equilibrium in the medium (Zanstra 1955). The result is a “multiphase” structure of HI, with cool, dense clouds in pressure equilibrium with their warmer, diffuse surroundings.

Subsequent models refined this model of the neutral ISM. Considering a homogeneous gas pressure maintained by ionization and heating from low-energy cosmic rays (Pikel’Ner 1968; Field et al. 1969), or X-rays (Silk & Werner 1969), balanced with cooling by fine-structure emission at low temperatures (e.g., [CII], $158\mu\text{m}$) and Ly α emission at around 8000 K, two thermally-stable “phases” emerge at kinetic temperatures of $\sim 10^4$ K and < 300 K in local pressure equilibrium. The steepness of the cooling curve in the two regimes of $T \sim 10^2$ K and $\sim 10^4$ K maintain this structure, even after heating by photoelectric dust emission (Watson 1972; Shull & Woods 1985) and magnetohydrodynamic dissipation (Silk 1975; Ferrière et al. 1988) are included (Wolfire et al. 1995; Ferrière 2001).

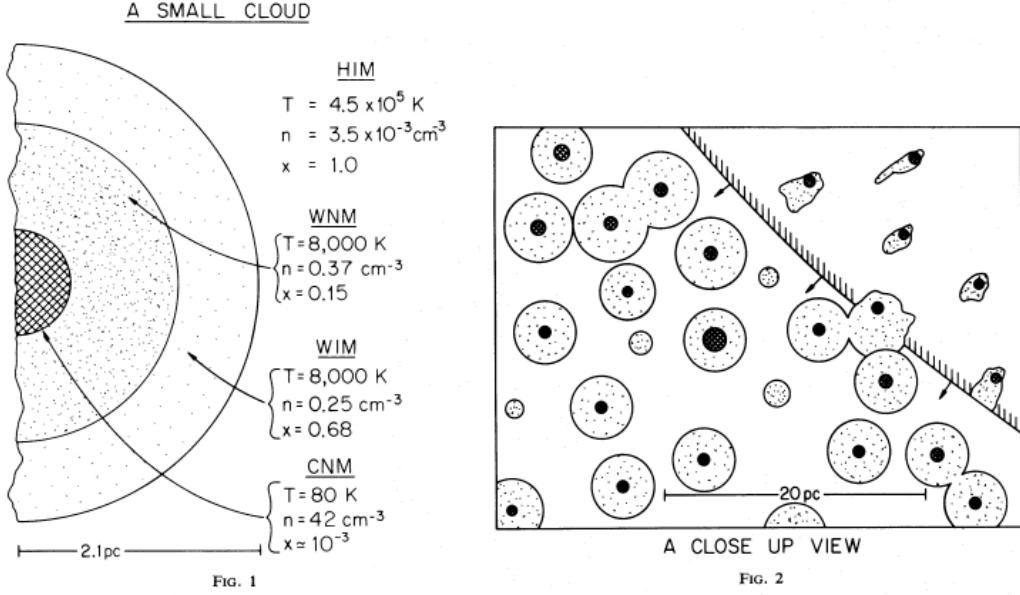


Figure 1.1 Left: Cross section of an interstellar HI cloud from McKee & Ostriker (1977) (their Figure 1). In summary, the cloud is structured as a core of CNM, with WNM envelope partially ionized by soft X-rays, and an outer WIM envelope ionized by stellar UV photons. Outside the cloud, the ambient HIM phase is maintained by supernovae. Right: A representative patch of the Galactic ISM, wherein a supernova blast wave propagates towards a population of HI clouds.

Following Cox & Smith (1974), who suggested that supernovae can maintain low-density but pervasive hot gas, McKee & Ostriker (1977, hereafter MO77) considered the effect of supernova heating and ionization on the ISM. In their model, dense clouds of cold neutral medium (CNM) are embedded in a pervasive, hot, ionized medium (HIM) so that each cloud has an envelope of “warm” gas with kinetic temperature ~ 8000 K, composed of both warm ionized medium (WIM; 70% ionized) and warm neutral medium (WNM; 10% ionized), with all phases in rough pressure equilibrium (Spitzer 1956, MO77). This model of the ISM is summarized in Figure 1.1, which includes a schematic view of an HI cloud (left) and a representative section of the Galactic ISM under the influence of a supernova blast wave (right) from MO77.

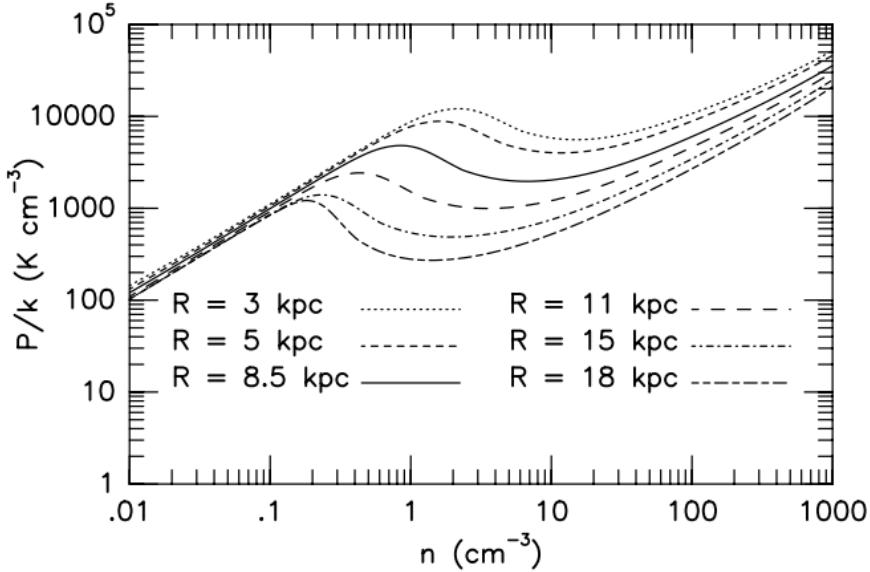


Figure 1.2 From Figure 7 of Wolfire et al. (2003): phase diagrams displaying thermal pressure (P/k) versus volume density (n) for a range of Galactocentric radii (R).

Throughout the Milky Way, the phase structure of the ISM is determined by local interstellar conditions, including the radiation field, gas phase metallicity and local dust-to-gas ratio. To illustrate the conditions required for maintaining a multiphase neutral ISM, Figure 1.2 depicts thermal equilibrium in the atomic ISM as a function of thermal pressure and volume density. The curves are computed by balancing photoelectric heating from dust grains with cooling by radiative line transitions ([CII], [OI], Ly α) and electron recombination (Wolfire et al. 1995) for typical conditions at a range of Galactocentric radii (Wolfire et al. 2003). The ISM is stable to isobaric perturbations when $dP/dn > 0$, producing CNM and WNM regimes at $(n, T_k) = (5\text{--}120 \text{ cm}^{-3}, 40\text{--}200 \text{ K})$, and $(n, T_k) = (0.03\text{--}1.3 \text{ cm}^{-3}, 4100\text{--}8800 \text{ K})$ respectively (Wolfire et al. 2003).

As seen in Figure 1.2, the stable CNM and WNM exist together (i.e., $dP/dn > 0$ at low and high densities) only for a very narrow range of thermal pressures, P , wherein $P_{\min} < P < P_{\max}$ for $P_{\max} \leq 3 P_{\min}$ (Field et al. 1969). The gravitational potential of the disk is sufficient to ensure $P > P_{\min}$, but the average density is too low for all HI to be in the CNM (i.e., $P < P_{\max}$) throughout the Milky Way (Wolfire et al. 2003). Therefore, interstellar HI is predicted to exist in

two thermally stable phases all the way out to radii of ~ 18 kpc (Wolfire et al. 2003). In support, observational studies of thermal pressure in the local ISM from carbon fine structure emission lines estimate thermal pressures of $P/k \sim 2500 - 5600$ K cm $^{-3}$, which fall within the required narrow pressure range for multiphase HI (Jenkins & Tripp 2011).

Intermediate conditions between the CNM and WNM ($dP/dn < 0$) represent a thermally unstable phase (UNM) wherein gas transitions quickly to the thermally-stable ranges of the CNM or WNM. However, “nonsteady” or time-dependent dynamical sources, such as supernovae and turbulence, can populate the unstable regions of $n - T$ parameter space (e.g., Werner et al. 1970; Dalgarno & McCray 1972; Gerola et al. 1974).

Numerical simulations of the ISM have explored the importance of dynamical processes on the phase structure of HI. However, theoretical predictions for the fractions of mass in each phase are strongly dependent on the prescriptions for detailed ISM physics, and are thus highly uncertain. In addition to thermal instability, self-gravity of HI clouds, magnetic fields, turbulence and shear from Galactic rotation are crucial (Vázquez-Semadeni et al. 2000). Turbulent driving can enhance phase transitions to the point where 50% of the HI mass should lie in the UNM (Gazol et al. 2001). However, recent high-resolution studies have shown that turbulence alone is insufficient, and additional sources of external pressure are required (Saury et al. 2014).

Many studies motivate cold cloud condensation via sustained ram pressure from colliding flows of warm gas (e.g., Audit & Hennebelle 2005; Hennebelle et al. 2007). Variations of ~ 10 km s $^{-1}$ in the initial speed of the colliding flows has a huge effect on the resulting fractions of CNM, WNM and UNM (Clark et al. 2012). In addition, shocks driven in turbulent, magnetized gas from supernovae create dense CNM clouds within a medium with a continuum of temperatures, where the unstable fraction is determined by the supernova rate (Koyama & Inutsuka 2002; Mac Low et al. 2005). The fraction of cold atomic gas surrounding molecular clouds also strongly depends on whether supernovae explode within density peaks (high fraction) or randomly (low fraction) (Gatto et al. 2015).

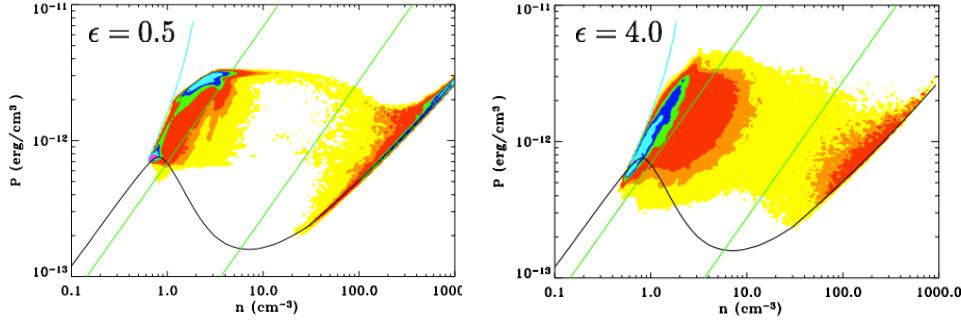


Figure 1.3 From Figures 3 and 7 of Audit & Hennebelle (2005): Gas masses of 1, 5, 10, 50, 100, 200, and 1000 in arbitrary units are shown with yellow, orange, red, green, dark blue, light blue, and pink, respectively. Thermal equilibrium, isothermal, and Hugoniot shock curves are shown in black, green ($T = 5000$ K left, 200 K right) and blue solid lines. Left: Results for turbulence modulation amplitude $\epsilon = 0.5$ (weak turbulent driving); Right: $\epsilon = 4.0$ (strong turbulent driving).

As an example, Figure 1.3 displays pressure-density diagrams from a colliding-flow numerical simulation by Audit & Hennebelle (2005). The two panels compare the gas fractions by mass in the CNM, WNM and unstable regime as a function of turbulent strength, parameterized by the amplitude of modulation (ϵ). With stronger turbulent driving (i.e., larger ϵ), the mass fraction in the UNM increases significantly, from $\sim 10\%$ to 30% (Audit & Hennebelle 2005).

Despite the analytical successes of the steady-state multiphase models of Field et al. (1969) and MO77, uncertainties regarding the thermal stability of HI and the influence of time-dependent radiative and dynamical processes revealed by recent numerical simulations remain prevalent. Ultimately, to constrain analytic and numerical models of the atomic ISM, detailed observational constraints for the temperature and density of HI are required. However, the properties of a significant fraction of the HI mass, including WNM and UNM, remain observationally unconstrained due to insufficient observational capabilities and systematic uncertainties in data analysis. Specifically, the excitation temperature of the WNM, and the fractions of HI mass in

each phase have been indirectly estimated only indirectly. Furthermore, connecting theoretical predictions with observations is difficult, and new methods are required for making objective comparisons and refining our understanding of the thermodynamic state of the ISM.

1.2 Observations of the atomic ISM

A convenient tracer for atomic gas in the ISM is the hyperfine energy splitting caused by magnetic moment interactions between the hydrogen atom’s electron and proton. This transition, located at a wavelength of $\lambda = 21$ cm and a frequency of $\nu = 1420.40575$ MHz, was first detected in emission from astronomical sources by Ewen & Purcell (1951) and Muller & Oort (1951), and since has become an essential tool for investigating the atomic ISM. The Einstein A emission coefficient of the 21 cm transition is $A_{10} = 2.85 \times 10^{-15} \text{ s}^{-1}$, which corresponds to a radiative lifetime of ~ 11 million years. This long lifetime ensures that collisions can maintain local thermodynamic equilibrium in normal galactic conditions. The ubiquity of HI in the ISM further guarantees that 21 cm emission is observable throughout the Milky Way. In the direction of a strong source of radio continuum (i.e., quasar, active galactic nucleus or radio galaxy), absorption by the 21 cm transition can also be measured. The first detection of HI absorption arrived soon after the first detection in emission (Hagen et al. 1955).

A striking result from the first 21 cm observations was that HI absorption and emission look very different (e.g., Clark 1965; Dickey et al. 1978). Figure 1.4 displays a representative set of 21 cm emission and absorption profiles from Dickey & Lockman (1990). All profiles were obtained by observing the same patch of the Milky Way using varying angular resolutions for emission and in the direction of a background radio continuum source for absorption. Figure 1.4 also includes a reference bar converting velocity relative to the local standard of rest (v_{LSR}) to the corresponding distance from the Sun in kiloparsecs. This illustrates the phenomenon of “velocity crowding”, wherein Galactic rotation projects widely-separated clouds into the same velocity channel (Burton 1966).

In contrast with the relatively small differences in HI emission due to varying angular

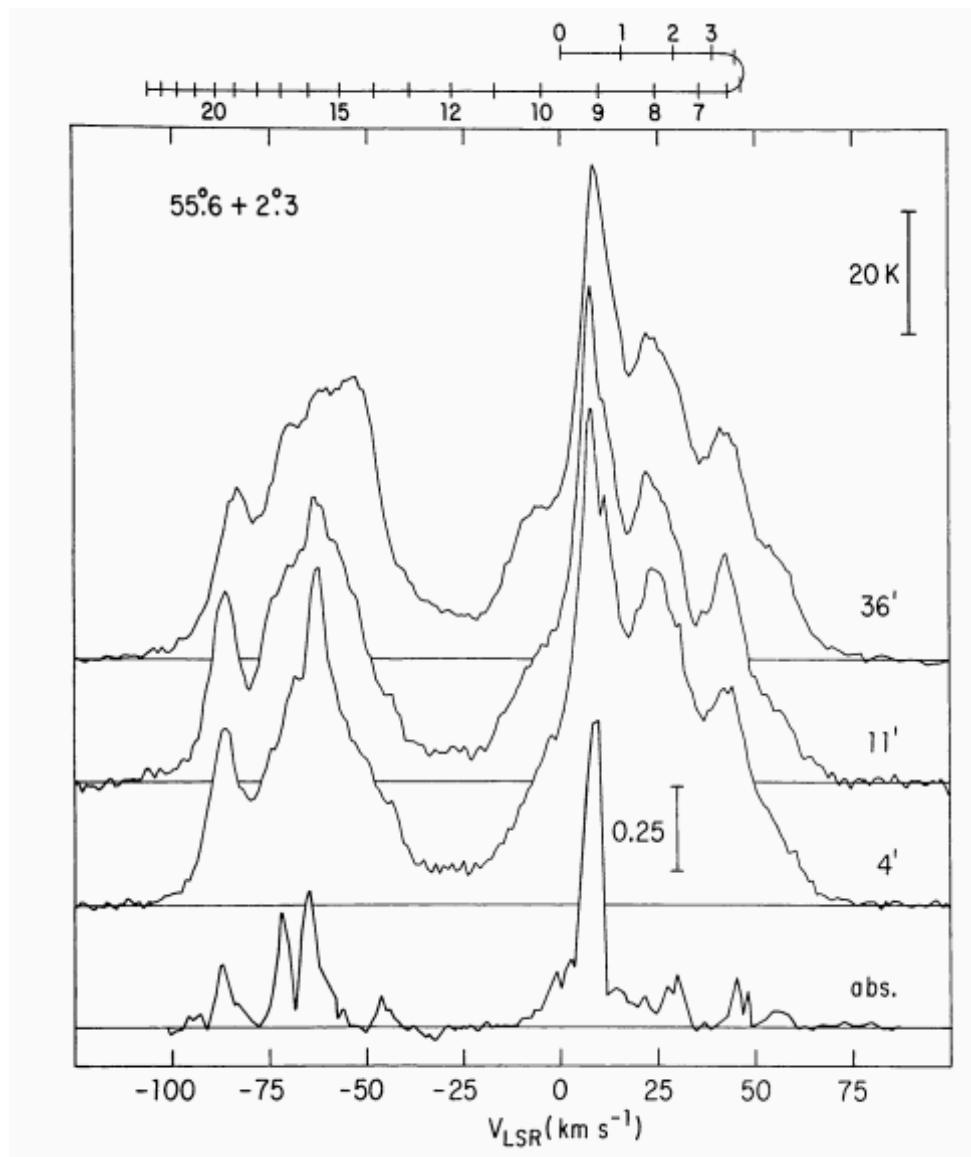


Figure 1.4 From Figure 2 of Dickey & Lockman (1990): Example 21 cm emission ($T_B(v)$; top three) and absorption ($1 - \exp(-\tau(v))$; bottom, inverted) profiles in the same Galactic direction ($55^{\circ}.6, +2^{\circ}.3$). Emission profiles were obtained with varying angular resolutions, $36'$, $11'$, $4'$, and the absorption profile was measured towards a background radio continuum source with $< 0''.1$ angular size. Above the top axis, a reference bar converts velocity to the corresponding distance from the Sun in units of kiloparsecs.

resolution, the velocity structure of HI absorption is significantly different than HI emission. In general, spectral line widths observed in absorption are narrower, and fewer velocity channels

exhibit absorption than emission. Although it is possible that different angular resolutions (i.e., \sim few' in emission versus $< 0''.1$ in absorption) account for these discrepancies, high-resolution observations of 21 cm emission and absorption at matched resolution have established that the differences are due to variations in the excitation temperature of HI and the mass fraction of the CNM along the line of sight (Clark 1965; Dickey & Lockman 1990). In other words, the earliest measurements of 21 cm emission and absorption showed that HI is multiphase, as predicted by the original steady-state models of the ISM.

Fundamental physical properties of HI in its many phases, including the CNM, WNM and UNM, can be measured by comparing properties of 21 cm emission and absorption in detail.

1.2.1 Measuring Physical Properties of HI

The emitting and absorbing properties of HI are difficult to measure simultaneously. The 21 cm line is located in the Rayleigh-Jeans limit of the Planck function ($h\nu \ll kT$), and therefore the intensity of emitted radiation can be parameterized with a brightness temperature (T_B). However, the radio continuum source required to measure HI absorption will contaminate T_B from the same intervening gas, and therefore observers must estimate the T_B in the absence of the radio continuum source. This requires two measurements: the brightness temperature “on” and “off” a radio continuum source. Assuming these measurements sample the same population of gas, the brightness temperature in a given velocity channel (v) on- and off-source is a combination of absorption and emission, given by,

$$T_{B,\text{on}}(v) = (T_c + T_{\text{sky}}) e^{-\tau(v)} + T_s (1 - e^{-\tau(v)}), \quad (1.1)$$

$$T_{B,\text{off}}(v) = T_{\text{sky}} e^{-\tau(v)} + T_s (1 - e^{-\tau(v)}), \quad (1.2)$$

where T_c is the radio continuum source flux density, and T_{sky} is the antenna temperature of the blank sky. These equations together determine the optical depth as a function of velocity ($\tau(v)$) and the excitation (or spin) temperature (T_s) — essential parameters for constraining the thermodynamic state of HI and observationally distinguishing between its varied phases. The

natural line width of the 21 cm transition in typical galactic conditions is extremely narrow, and therefore the observed line width is dominated by Doppler broadening from the gas kinetic temperature, turbulence, or cloud blending along the line of sight. An upper limit to the gas kinetic temperature $T_{k,\max}$, can be computed from the measured spectral line width, Δv ,

$$T_{k,\max} = \frac{m_H}{8k_B \ln(2)} \times \Delta v^2 = 21.866 \Delta v^2, \quad (1.3)$$

for hydrogen mass m_H , and Boltzmann's constant k_B (Draine 2011). Within the high-density CNM, collisions between electrons, ions, and other hydrogen atoms can thermalize the 21 cm transition so that $T_s \approx T_k$. On the other hand, within the low-density WNM, collisions are not sufficient to thermalize the transition, and therefore it is widely expected that $T_s < T_k$ (e.g., Field 1958; Deguchi & Watson 1985; Liszt 2001).

Given T_s and $\tau(v)$, 21 cm line observations can also determine the column density, $N(\text{H}\text{I})$. Beginning with the absorption coefficient for H\text{I}, and assuming a Gaussian velocity distribution, the column density is computed by integrating the volume density (n) along the line of sight (s),

$$N(\text{H}\text{I}) = \int n \, ds = C_0 \int T_s \, \tau(v) \, dv, \quad (1.4)$$

where $C_0 = 1.823 \times 10^{18} \text{ cm}^{-2} (\text{km s}^{-1})^{-1} \text{ K}^{-1}$ and dv is measured in km s^{-1} (e.g., Spitzer 1978; Draine 2011). Substituting Equation 1.2 gives,

$$N(\text{H}\text{I}) = C_0 \int \frac{T_B(v) \tau(v)}{1 - e^{-\tau(v)}} \, dv. \quad (1.5)$$

In the optically-thin limit, assuming that $\tau \ll 1$, this reduces to,

$$N(\text{H}\text{I})_{\text{thin}} = C_0 \int T_B(v) \, dv. \quad (1.6)$$

This case is widely assumed when measurements of τ are unavailable. Generally, the optically-thin assumption is valid in the local ISM, and will only underestimate the true H\text{I} content by $\sim 10\%$, even near to molecular clouds (Lee et al. 2015). However, along sightlines within a few degrees

of the Galactic plane, line blending and high densities of star-forming molecular clouds tend to saturate HI absorption profiles and significantly alter $T_B(v)$ (e.g., Bihr et al. 2015).

The advent of radio interferometers facilitated a revolution in observations of HI absorption (e.g., Clark 1965; Hughes et al. 1971; Radhakrishnan et al. 1972). Absorption lines probe only the gas directly in front of a radio continuum source, whereas emission lines sample all HI within the telescope beam. Therefore, telescopes with the highest possible angular resolution on the scale of the continuum source itself, achievable by interferometers, are preferred for filtering out emission on large angular scales and isolating velocity structures due to absorption alone. Interferometric observations of Galactic HI emission must attempt to restore the “zero-spacing” flux using single dish measurements. However, for absorption line observations this is an advantage, as it further suppresses flux from large-scale emission which might otherwise contaminate the absorption profile.

Ideally, to derive physical properties of HI it is preferable to obtain emission and absorption measurements on the same angular scale with the same telescope. In the interest of measuring absorption, as discussed above, it is best to use an interferometer, which can achieve higher angular resolution than any single dish telescope. However, observing HI emission using an interferometer is prohibitively expensive. The noise in brightness temperature is proportional to $\text{HPBW}^{-2} t^{-1/2}$, where HPBW is the half power at the full width of the beam and t is the integration time (Dickey & Lockman 1990). For example, detecting an emission signal at 21 cm using the same configuration required for measuring absorption at the VLA (i.e., the widest configurations: $\text{HPBW} \sim 1''$) would need a factor of 3600 longer integration time compared with the most compact configuration ($\text{HPBW} \sim 60''$) which is unsuitable for measuring absorption. Another strategy is to extract HI absorption in the direction of continuum sources included in large-area surveys of the sky in HI emission. However, extracting uncontaminated HI absorption spectra from HI emission data is tricky (e.g., Dickey et al. 2003), and typically results in samples of spectra with widely varying sensitivity. A compromise is to measure absorption using an interferometer, and emission using

a single-dish telescope with good angular resolution.

1.3 The 21-SPONGE Survey

With these considerations in mind, we designed a large survey to measure the properties of HI in all phases at the VLA. Recent upgrades to the VLA (formerly EVLA) have enabled observations of 21 cm absorption at higher sensitivity and with drastically better calibration than ever before. Most notably, the wide spectral bandwidths and high frequency resolution achievable with the WIDAR signal correlator has improved the characterization and subtraction of radio continuum dramatically. The survey, 21-cm Spectral Line Observations of Neutral Gas with the VLA (21-SPONGE), targets 59 bright, compact radio continuum sources to measure HI absorption with root mean square uncertainty in HI optical depth of $\sigma_\tau < 10^{-3}$. Matching emission profiles for 21-SPONGE sources were obtained at the Arecibo Observatory, whose $3'.5$ beam at 21 cm provides the closest efficient complement for the interferometric absorption measurements at the VLA.

The 21-SPONGE survey was conducted in 573 hours between February 2011 and June 2015. The result is the largest, highest-sensitivity survey for HI absorption ever undertaken. In the following section, I describe current observational progress in the context of three open questions which motivated 21-SPONGE. These questions will provide the structure for the subsequent chapters of this thesis.

1.4 Current Progress and Open Questions

1.4.1 What is the temperature of the WNM?

Of the HI phases, the CNM is best understood. The densities are high enough that the equilibrium timescale is short, and therefore the physical properties are well-determined (Wolfire et al. 1995). Furthermore, the high optical depth of the CNM makes 21 cm absorption signatures easy to detect, even with low sensitivity (e.g., Radhakrishnan et al. 1972; Lazareff 1975; Dickey et al. 1978; Crovisier et al. 1978; Dickey et al. 1983; Braun & Walterbos 1992; Heiles & Troland

2003a; Mohan et al. 2004). The CNM has been detected throughout the Milky Way, all the way out to a Galactocentric radius of ~ 25 kpc (Dickey et al. 2009). These observational efforts have established the spin temperature of the CNM to be $\sim 10 < T_s < 200$ K.

In contrast, the WNM is characterized by very low optical depth, so detecting it via 21 cm absorption – necessary for constraining τ and T_s – requires extremely high sensitivity. The majority of WNM spin temperature measurements are indirect, providing upper limits on T_k , not T_s (Mebold et al. 1982; Kanekar et al. 2003; Heiles & Troland 2003b; Roy et al. 2013b). Other estimates of WNM T_s have been obtained by integrating strongly-absorbing, complex HI absorption and emission profiles to assign a single, average temperature to a full line of sight (Kanekar et al. 2011; Roy et al. 2013a).

Only three direct measurements of WNM T_s from absorption have been made (Carilli et al. 1998; Dwarakanath et al. 2002). The measured values, $T_s \sim 3600, 4800$, and 6000 K, generally agree with analytical predictions ($1800 \leq T_s \leq 4800$ K; Liszt 2001). However, the small sample size makes it difficult to generalize these results. Detecting the WNM in absorption furthermore requires attention to systematic errors in the baseline structure of HI profiles, as the lines are expected to be shallow and broad, with $\tau < 0.01$ and FWHM $\gtrsim 10$ km s $^{-1}$. Figure 1.5 displays a WNM absorption detection from Dwarakanath et al. (2002), at $v \sim -50$ km s $^{-1}$ using the Westerbork Synthesis Radio Telescope (WSRT). However, the baseline at $v \sim 20$ km s $^{-1}$ has similar but inverted structure as the detected absorption line, illustrating the uncertainty introduced by unphysical bandpass structure in detecting wide, shallow spectral lines, such as those from the WNM in absorption.

Furthermore, the details of 21 cm excitation in the diffuse WNM are poorly constrained. Although the excitation temperature (i.e., T_s) of the WNM is expected to be less than T_k due to insufficient collisional ionization at low densities, the Ly α radiation field from Galactic and extragalactic sources can serve to thermalize the 21 cm transition. This effect, known as the Wouthuysen-Field effect, requires a very large optical depth and a large number of scatterings of

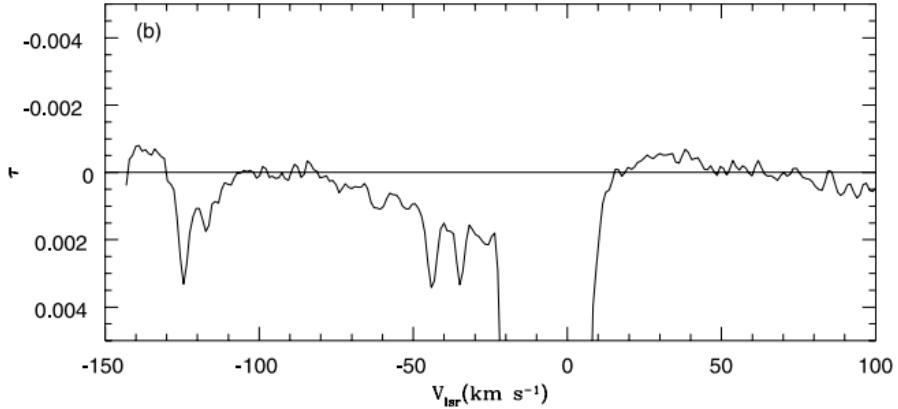


Figure 1.5 From Figure 5 of Dwarakanath et al. (2002): Detection of absorption by the WNM at $v \sim -50 \text{ km s}^{-1}$ with the WSRT.

$\text{Ly}\alpha$ photons to bring the radiation field and the gas into local thermal equilibrium. While the underlying atomic physics is understood (Wouthuysen 1952; Field 1958; Furlanetto et al. 2006), the details of $\text{Ly}\alpha$ radiative transfer are complicated and depend on the topology and the strength of the $\text{Ly}\alpha$ radiation field, which are complex and poorly constrained in the multiphase ISM (Liszt 2001). These properties are furthermore important for interpreting radio signals from early epochs of cosmic structure formation (e.g., the cosmic dark ages and subsequent epoch of reionization), when neutral hydrogen facilitated the formation of the first stars and galaxies (Pritchard & Loeb 2012).

In their seminal paper, MO77 noted, “Probably the greatest weakness of the model presented here concerns the warm neutral medium.” (McKee & Ostriker 1977). In conclusion to the Millennium Survey, Heiles & Troland (2003b) declared the WNM a “key to the Universe” due to its detailed and uncertain dependence on myriad micro- and macro-physical processes at work in the ISM. Until now, a statistical observational study of the temperature and density of the WNM has been unfeasible due to insufficient observational capabilities.

1.4.2 How well do simulations reproduce observed HI properties?

To diagnose the physics underlying observed properties of the CNM, WNM and UNM, we must compare observations with predictions from theory. However, simulations have only recently achieved suitable dynamic range and resolution for capturing both CNM and WNM dynamics (e.g., Saury et al. 2014; Kim & Ostriker 2016). Furthermore, the inherent uncertainty in comparing simulated ISM properties with those inferred from spectral line observations has presented a significant stumbling block to progress.

As simulations have improved in resolution and sensitivity, it has become common practice to construct synthetic “observations” by mimicking real observational conditions and parameters within simulated volumes. Synthetic spectral line observations of atomic gas from numerical simulations provide a particularly convenient opportunity for comparing 21 cm observations and theory as directly as possible. Such comparisons enable us to both quantify the successes of observational diagnostics and to test how complex simulations recover the properties of real systems. For example, the velocity structures of synthetic spectral lines provide important diagnostics of interstellar turbulence (Falgarone et al. 1994).

Figure 1.6 displays example synthetic HI emission and absorption line pairs from the high-resolution two-dimensional hydrodynamic simulations from Hennebelle et al. (2007) and from the three-dimensional simulations presented in Kim et al. (2014). Although the synthetic spectra are qualitatively similar to each other and to real 21 cm observations, detailed comparisons may reveal key differences. For example, Hennebelle et al. (2007) find good agreement in column density and velocity structure of CNM clouds in comparison with Heiles & Troland (2003b), however their HI absorption profiles are generally wider and have a higher ratio of CNM to total HI column density than found in observations. Kim et al. (2014) find good agreement in LOS-integrated quantities such as column density and harmonic mean spin temperature via comparison with observations from Roy et al. (2013a). However, observations have proven that even high galactic latitude lines of sight contain several components of varying spin temperature (e.g., Heiles & Troland 2003a;

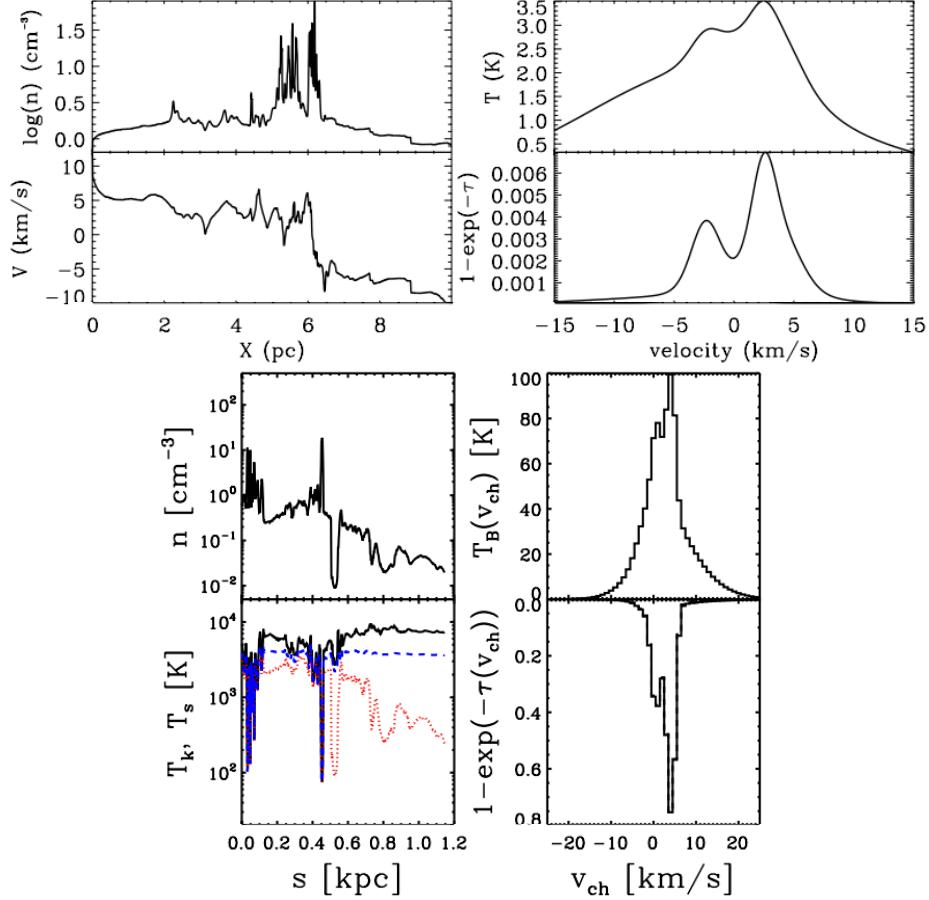


Figure 1.6 Top: From Figure 7 of Hennebelle et al. (2007): A slice in density (n) and velocity (v) with corresponding synthetic 21 cm emission and absorption from a 2-dimensional hydrodynamic high-resolution simulation of the atomic ISM. Bottom: From Figure 3 of Kim et al. (2014): A slice in density (n) and temperature (T) with corresponding synthetic 21 cm spectral lines from a 3-dimensional hydrodynamic simulation of the multiphase ISM in the Milky Way.

Begum et al. 2010; Roy et al. 2013b), implying that LOS-integrated or isothermal diagnostics are missing important complexity of ISM structure.

Decomposing spectral lines into Gaussian components is a widely-used strategy for inferring the properties of multiphase gas along observed lines of sight (e.g., Lazareff 1975; Mebold et al. 1982; Mohan et al. 2004). Simultaneous decomposition of 21 cm emission and absorption profiles determines T_s and $N(\text{HI})$ for individual spectral features (e.g., Heiles & Troland 2003a; Begum et al. 2010; Stanimirović et al. 2014). However, Gaussian decompositions are non-unique, and the extent to which they recover the true properties of HI structures has not been quantified.

Ultimately, detailed decompositions of and comparisons between real and synthetic observations require uniform, objective analysis tools. The sample size of synthetic spectra is potentially limitless, and with the advent of next-generation radio telescopes pouring out terabytes of data per hour, it is especially necessary to develop novel methods for analyzing ever-increasing volumes of real and synthetic data.

1.4.3 What is the mass fraction of HI as a function of temperature?

Beyond the CNM and WNM, 21 cm observations have revealed unexpectedly large fractions of gas in the thermally unstable phase, or UNM (Heiles 2001). Detailed decomposition of 21 cm emission observations revealed three gas populations featuring distinct line-widths, corresponding to the CNM, WNM and thermally unstable temperatures of ~ 3000 K (Verschuur & Magnani 1994). In agreement, Gaussian analysis of the all-sky Leiden Argentine Bonn (LAB) survey (Kalberla et al. 2005) revealed a population of thermally-unstable gas with FWHM $\sim 11.8 \text{ km s}^{-1}$ comprising $\sim 40\%$ of the total column density (Haud & Kalberla 2007). However, these results were based on 21 cm emission alone and therefore only provide upper limits for the HI spin temperature and UNM mass fraction.

Observations of both emission and absorption at 21 cm have found similarly significant fractions of unstable gas. A high-resolution study HI emission and absorption using VLBI found that $\sim 80\%$ of mass in the direction of 3C147 has thermally unstable temperatures (Kalberla et al. 1985). From the Millennium Survey, Heiles & Troland (2003b) concluded that 48% of the WNM over 79 sightlines exists in the thermally unstable regime, with $500 \leq T_s \leq 5000$ K. However, due to limited observational sensitivity in optical depth, the T_s estimates for the UNM and WNM from Heiles & Troland (2003b) are upper limits from line width-derived $T_{k,\max}$ only. Using 35 high-sensitivity HI absorption observations, Roy et al. (2013b) argued that $\leq 28\%$ of the HI mass is in the unstable regime. Unfortunately, high-sensitivity absorption measurements are extremely expensive to undertake, and to date have been limited to a small sample of sources.

Significant mass fractions of HI in the thermally unstable regime of phase space imply that

the steady-state model of MO77, invoking rough pressure equilibrium between phases, is invalid. The influence of dynamical phenomena in the ISM, such as turbulence and supernovae, is crucial, and a high-sensitivity survey of a statistically significant sample of 21 cm absorption lines is sorely needed. Furthermore, observational results must be corrected for biases introduced in HI spectral line analysis.

1.5 Thesis Outline

In this thesis, I address the open questions described above with data from the 21-SPONGE survey. The exceptional optical depth sensitivity of 21-SPONGE enables it to perform the first statistical survey of HI in the full temperature range of $10 < T_s < 10^4$ K. In Chapter 2, I present the observational methods and initial results from the first half of 21-SPONGE (Murray et al. 2015). I discuss the detailed observing and calibration considerations which enabled 21-SPONGE to detect the first statistical sample of broad ($\sim 10 \text{ km s}^{-1}$) absorption lines and thereby extend the mass distribution of HI as a function of T_s detected in absorption beyond the CNM (Murray et al. 2015)

However, despite the excellent sensitivity of 21-SPONGE, we initially detected little evidence for diffuse, warm HI with $T_s > 1000$ K (< 10%; Murray et al. 2015). In Chapter 3, I describe a spectral stacking method for HI absorption lines to maximize sensitivity to absorption from diffuse, warm HI (Murray et al. 2014). Applying this method to 21-SPONGE spectra, I present evidence for a pervasive population of WNM with $T_s = 7200_{-1200}^{+1800}$ K. This temperature has implications not only for HI physics in the local ISM, but also for 21 cm signals from early epochs of cosmic time when HI dominated the baryonic content of the Universe.

To facilitate direct comparisons between 21-SPONGE and numerical simulations, in Chapter 4 I present an objective, efficient method for decomposing HI spectral line pairs into Gaussian functions and estimating the temperature and densities of individual gas structures via radiative transfer (Murray et al. 2017). This method applies the Autonomous Gaussian Decomposition algorithm (AGD; Lindner et al. 2015): a tool for deriving initial guesses for multi-component fits

via derivative spectroscopy and machine learning. I analyze 21-SPONGE and synthetic spectra from Kim et al. (2014) simultaneously and compare properties inferred from synthetic observations with true simulated conditions to quantify for the first time how well spectral lines reproduce real cloud properties in the ISM.

In Chapter 5, I present the final 21-SPONGE data release and apply the AGD-based radiative transfer method developed in Murray et al. (2017) to compute temperatures and column densities of HI structures along all 57 lines of sight. I compute the mass fractions of HI in each phase detected by 21-SPONGE. For the first time, I correct for biases introduced in 21 cm analysis via comparison with an identical decomposition of synthetic data from Kim et al. (2014). The HI mass fractions have important implications for dynamical models of turbulence and star formation in the ISM.

Finally, in Chapter 6, I summarize my results, and discuss the future directions of this work.

References

- Audit, E., & Hennebelle, P. 2005, A&A, 433, 1
- Begum, A., Stanimirović, S., Goss, W. M., et al. 2010, ApJ, 725, 1779
- Bigiel, F., Leroy, A., Walter, F., et al. 2008, AJ, 136, 2846
- Bihr, S., Beuther, H., Ott, J., et al. 2015, A&A, 580, A112
- Braun, R., & Walterbos, R. A. M. 1992, ApJ, 386, 120
- Burton, W. B. 1966, AJ, 71, 848
- Carilli, C. L., Dwarakanath, K. S., & Goss, W. M. 1998, ApJ, 502, L79
- Christensen, C., Quinn, T., Governato, F., et al. 2012, MNRAS, 425, 3058
- Clark, B. G. 1965, ApJ, 142, 1398
- Clark, P. C., Glover, S. C. O., Klessen, R. S., & Bonnell, I. A. 2012, MNRAS, 424, 2599
- Cox, D. P., & Smith, B. W. 1974, ApJ, 189, L105
- Crovisier, J., Kazes, I., & Aubry, D. 1978, A&AS, 32, 205
- Dalgarno, A., & McCray, R. A. 1972, ARA&A, 10, 375
- Deguchi, S., & Watson, W. D. 1985, ApJ, 290, 578
- Dickey, J. M., Kulkarni, S. R., van Gorkom, J. H., & Heiles, C. E. 1983, ApJS, 53, 591
- Dickey, J. M., & Lockman, F. J. 1990, ARA&A, 28, 215
- Dickey, J. M., McClure-Griffiths, N. M., Gaensler, B. M., & Green, A. J. 2003, ApJ, 585, 801
- Dickey, J. M., Strasser, S., Gaensler, B. M., et al. 2009, ApJ, 693, 1250
- Dickey, J. M., Terzian, Y., & Salpeter, E. E. 1978, ApJS, 36, 77
- Draine, B. T. 2011, Physics of the Interstellar and Intergalactic Medium (Princeton University Press)
- Dwarakanath, K. S., Carilli, C. L., & Goss, W. M. 2002, ApJ, 567, 940
- Ewen, H. I., & Purcell, E. M. 1951, Nature, 168, 356
- Falgarone, E., Lis, D. C., Phillips, T. G., et al. 1994, ApJ, 436, 728
- Ferrière, K. M. 2001, Reviews of Modern Physics, 73, 1031
- Ferrière, K. M., Zweibel, E. G., & Shull, J. M. 1988, ApJ, 332, 984
- Field, G. B. 1958, Proceedings of the IRE, 46, 240
- . 1965, ApJ, 142, 531

- Field, G. B., Goldsmith, D. W., & Habing, H. J. 1969, *ApJ*, 155, L149
- Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, *Phys. Rep.*, 433, 181
- Gatto, A., Walch, S., Low, M.-M. M., et al. 2015, *MNRAS*, 449, 1057
- Gazol, A., Vázquez-Semadeni, E., Sánchez-Salcedo, F. J., & Scalo, J. 2001, *ApJ*, 557, L121
- Gerola, H., Kafatos, M., & McCray, R. 1974, *ApJ*, 189, 55
- Governato, F., Brook, C., Mayer, L., et al. 2010, *Nature*, 463, 203
- Hagen, J. P., Lilley, A. E., & McClain, E. F. 1955, *ApJ*, 122, 361
- Haud, U., & Kalberla, P. M. W. 2007, *A&A*, 466, 555
- Heiles, C. 2001, *ApJ*, 551, L105
- Heiles, C., & Troland, T. H. 2003a, *ApJS*, 145, 329
- . 2003b, *ApJ*, 586, 1067
- Hennebelle, P., Audit, E., & Miville-Deschénes, M.-A. 2007, *A&A*, 465, 445
- Hopkins, P. F., Cox, T. J., Hernquist, L., et al. 2013, *MNRAS*, 430, 1901
- Hughes, M. P., Thompson, A. R., & Colvin, R. S. 1971, *ApJS*, 23, 323
- Jenkins, E. B., & Tripp, T. M. 2011, *ApJ*, 734, 65
- Kalberla, P. M. W., Burton, W. B., Hartmann, D., et al. 2005, *A&A*, 440, 775
- Kalberla, P. M. W., Schwarz, U. J., & Goss, W. M. 1985, *A&A*, 144, 27
- Kanekar, N., Braun, R., & Roy, N. 2011, *ApJ*, 737, L33
- Kanekar, N., Subrahmanyan, R., Chengalur, J. N., & Safouris, V. 2003, *MNRAS*, 346, L57
- Kim, C.-G., & Ostriker, E. C. 2016, ArXiv e-prints, arXiv:1612.03918
- Kim, C.-G., Ostriker, E. C., & Kim, W.-T. 2014, *ApJ*, 786, 64
- Koyama, H., & Inutsuka, S.-i. 2002, *ApJ*, 564, L97
- Lazareff, B. 1975, *A&A*, 42, 25
- Lee, M.-Y., Stanimirović, S., Murray, C. E., Heiles, C., & Miller, J. 2015, *ApJ*, 809, 56
- Leroy, A. K., Walter, F., Sandstrom, K., et al. 2013, *AJ*, 146, 19
- Lindner, R. R., Vera-Ciro, C., Murray, C. E., et al. 2015, *AJ*, 149, 138
- Liszt, H. 2001, *A&A*, 371, 698
- Mac Low, M.-M., Balsara, D. S., Kim, J., & de Avillez, M. A. 2005, *ApJ*, 626, 864
- McKee, C. F., & Ostriker, J. P. 1977, *ApJ*, 218, 148
- Mebold, U., Winnberg, A., Kalberla, P. M. W., & Goss, W. M. 1982, *A&A*, 115, 223

- Mohan, R., Dwarakanath, K. S., & Srinivasan, G. 2004, Journal of Astrophysics and Astronomy, 25, 185
- Muller, C. A., & Oort, J. H. 1951, Nature, 168, 357
- Murray, C. E., Stanimirović, S., Kim, C.-G., et al. 2017, ApJ, 837, 55
- Murray, C. E., Lindner, R. R., Stanimirović, S., et al. 2014, ApJ, 781, L41
- Murray, C. E., Stanimirović, S., Goss, W. M., et al. 2015, ApJ, 804, 89
- Ostriker, E. C., McKee, C. F., & Leroy, A. K. 2010, ApJ, 721, 975
- Parker, E. N. 1953, ApJ, 117, 431
- Pikel'Ner, S. B. 1968, ARA&A, 6, 165
- Pritchard, J. R., & Loeb, A. 2012, Reports on Progress in Physics, 75, 086901
- Radhakrishnan, V., Murray, J. D., Lockhart, P., & Whittle, R. P. J. 1972, ApJS, 24, 15
- Roy, N., Kanekar, N., Braun, R., & Chengalur, J. N. 2013a, MNRAS, 436, 2352
- Roy, N., Kanekar, N., & Chengalur, J. N. 2013b, MNRAS, 436, 2366
- Saury, E., Miville-Deschénes, M.-A., Hennebelle, P., Audit, E., & Schmidt, W. 2014, A&A, 567, A16
- Schruba, A., Leroy, A. K., Walter, F., et al. 2011, AJ, 142, 37
- Shull, J. M., & Woods, D. T. 1985, ApJ, 288, 50
- Silk, J. 1975, ApJ, 198, L77
- Silk, J., & Werner, M. W. 1969, ApJ, 158, 185
- Spitzer, L. 1978, Physical processes in the interstellar medium, doi:10.1002/9783527617722
- Spitzer, Jr., L. 1956, ApJ, 124, 20
- Stanimirović, S., Murray, C. E., Lee, M.-Y., Heiles, C., & Miller, J. 2014, ApJ, 793, 132
- Tasker, E. J., & Bryan, G. L. 2008, ApJ, 673, 810
- Vázquez-Semadeni, E., Gazol, A., & Scalo, J. 2000, ApJ, 540, 271
- Verschuur, G. L., & Magnani, L. 1994, AJ, 107, 287
- Watson, W. D. 1972, ApJ, 176, 103
- Werner, M. W., Silk, J., & Rees, M. J. 1970, ApJ, 161, 965
- 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
- Wouthuysen, S. A. 1952, AJ, 57, 31
- Zanstra, H. 1955, Vistas in Astronomy, 1, 256

Chapter 2

The 21-SPONGE HI Absorption Survey I: Techniques and Initial Results

*A version of this chapter has previously appeared
in the Astrophysical Journal*

C. E. Murray, S. Stanimirović, W. M. Goss, et al. 2015, ApJ, 804, 89

Abstract

We present methods and results from “21-cm Spectral Line Observations of Neutral Gas with the EVLA” (21-SPONGE), a large survey for Galactic neutral hydrogen (HI) absorption with the Karl G. Jansky Very Large Array (VLA). With the upgraded capabilities of the VLA, we reach median root-mean-square (RMS) noise in optical depth of $\sigma_\tau = 9 \times 10^{-4}$ per 0.42 km s^{-1} channel for the 31 sources presented here. Upon completion, 21-SPONGE will be the largest HI absorption survey with this high sensitivity. We discuss the observations and data reduction strategies, as well as line fitting techniques. We prove that the VLA bandpass is stable enough to detect broad, shallow lines associated with warm HI, and show that bandpass observations can be combined in time to reduce spectral noise. In combination with matching HI emission profiles from the Arecibo Observatory ($\sim 3.5'$ angular resolution), we estimate excitation (or spin) temperatures (T_s) and column densities for Gaussian components fitted to sightlines along which we detect HI absorption (30/31). We measure temperatures up to $T_s \sim 1500 \text{ K}$ for individual lines, showing that we can probe the thermally unstable interstellar medium (ISM) directly. However, we detect fewer of these thermally unstable components than expected from previous observational studies. We probe a wide range in column density between $\sim 10^{16}$ and $> 10^{21} \text{ cm}^{-2}$ for individual HI clouds. In addition, we reproduce the trend between cold gas fraction and average T_s found by synthetic observations of a hydrodynamic ISM simulation by Kim et al. (2014). Finally, we investigate methods for estimating HI T_s and discuss their biases.

2.1 Introduction

Star formation drives the evolution of galaxies, and the material from which stars form constantly cycles in and out of the interstellar medium (ISM). Within this process, neutral hydrogen (HI) represents a crucial transitional stage between supernovae-expelled plasma and dense star-forming gas. Consequently, the presence and influence of radiative and dynamic physical processes in the ISM are imprinted in the mass distribution of HI as a function of temperature and density. Observational constraints over the full range of temperature and density are therefore necessary for realistic heating, cooling and feedback prescriptions within accurate models of star formation and galaxy evolution on all galactic scales (e.g. Bryan et al. 2007, Governato et al. 2010, Ostriker et al. 2010, Hill et al. 2012, Christensen et al. 2012).

Analytical models of heating and cooling in the ISM predict the existence of two neutral HI phases— the cold neutral medium (CNM) and the warm neutral medium (WNM)— each in thermal equilibrium (e.g., Pikel’ner 1968, Field et al. 1969, McKee & Ostriker 1977). McKee & Ostriker (1977) and Wolfire et al. (2003) predict that the two phases should be mostly in thermal equilibrium, with the CNM dominating and the WNM comprising only a few percent of the total HI column density. Over the range of pressures for which the two stable phases can coexist, $\sim 1000 < P/k < 8000 \text{ cm}^{-3} \text{ K}$, we theoretically expect a volume density (n) and kinetic temperature (T_k) of $(n, T_k) = (5\text{--}120 \text{ cm}^{-3}, 40\text{--}200 \text{ K})$ for the CNM, and $(n, T_k) = (0.03\text{--}1.3 \text{ cm}^{-3}, 4100\text{--}8800 \text{ K})$ for the WNM (Wolfire et al. 2003). Within the high-density CNM, collisions between electrons, ions and other H atoms can thermalize the 21-cm transition so that $T_s = T_k$. On the other hand, within the low-density WNM, collisions are not sufficient to thermalize the transition, and therefore it is widely expected that $T_s < T_k$ (e.g. Field et al. 1958, Deguchi & Watson 1985, Liszt 2001).

In addition, the presence of HI in the intermediate, thermally unstable regime between the CNM and WNM has a range of theoretical implications. Dynamical processes are required to

maintain this unstable material, as thermal pressure alone is not sufficient to force gas across extreme density boundaries (Cox et al. 2005). Shocks driven by supernovae into turbulent, magnetized gas create dense CNM clouds within a multiphase medium, where the fraction of thermally unstable gas is determined by the supernova rate (Koyama & Inutsuka 2002, MacLow et al. 2005). Audit & Hennebelle (2005) showed that turbulent flow collisions produce unstable fractions of 10–30%, depending on the amount of turbulence injected. Modeling collisions of wind-blown superbubbles, Ntormousi et al. (2011) found most gas in cold clumps, with only 8–10% thermally-unstable WNM by mass. In another colliding flow model, Clark et al. (2012) showed that the initial flow speed, including a time-dependent chemical network, has a huge effect on the resulting fractions of cold and warm gas. Clearly, observational constraints of the WNM properties and the full temperature distribution over 10– 10^4 K are important to constrain the role of self-gravity, turbulence, and large-scale gas flows in models of the ISM, as well as to guide initial conditions for numerical simulations.

2.1.1 Previous Work

Observationally, the hyperfine 21-cm transition of HI provides an excellent tracer of the CNM and WNM. Given measurements of *both* HI emission and HI absorption, it is possible to directly measure the excitation temperature (or spin temperature, T_s) and column density ($N(\text{HI})$) of CNM and WNM structures along the line of sight. Due to the high optical depth of the CNM, 21-cm absorption signatures are easy to detect, even with low sensitivity (e.g. Lazareff 1975, Dickey et al. 1977, Crovisier et al. 1978, Payne et al. 1978, Dickey et al. 1983, Braun & Walterbos 1992, Heiles & Troland 2003, Kanekar et al. 2003, Mohan et al. 2004, Roy et al. 2006, Begum et al. 2010). Extensive studies have been undertaken in the Milky Way and the CNM is easily detected all the way out to a Galactocentric radius of ~ 25 kpc (e.g., Dickey et al. 2009).

However, the optical depth of the WNM is so low ($\tau \leq 10^{-3}$) that extremely high sensitivity is required to detect it in absorption. Few studies have targeted individual detections of the WNM in absorption to directly measure T_s (e.g. Carilli et al. 1998, Dwarakanath et al. 2002, Chapter 3),

Table 2.1. Survey Comparison

Survey	Galactic latitude ($^{\circ}$) (1)	Angular res. (arcmin) (2)	Velocity res. (km s^{-1}) (3)	Sensitivity (σ_{τ}) ^a (4)	Number of spectra (5)	Telescopes (absorption, emission) (6)
Absorption targeted:						
21-SPONGE	$ b > 3.7$	4	0.5	0.0006	58	VLA, Arecibo
Roy et al. 2013	$ b > 1.6$	36	1.0	0.0005	35	WSRT/GMRT/ATCA, LAB
Heiles & Troland 2003	all $ b $	4	0.16	0.002	78	Arecibo
Stanimirović et al. 2014	$-32 < b < -8$	4	0.16	0.002	27	Arecibo
Mohan et al. 2004	$ b > 15$	36	3.3	0.005	102	GMRT, LDS
Emission mapped:						
VGPs ^b	$ b < 1.3$	1	1.56	0.025-0.125	113	VLA, GBT
CGPs ^c	$-3.6 < b < 15$	1	1.32	0.023-0.115	364	DRAO, DRAO 26m
SGPs ^d	$ b < 1.5$	2	1.0	0.02-0.1	96	ATCA, Parkes
GASKAP ^e	$ b < 10$	0.5	1.0	0.02	>1000	ASKAP

Note. — ^a: Median RMS noise in HI optical depth per 1 km s^{-1} channel. ^b: Stil et al. (2006); ^c: Taylor et al. (2003); ^d: McClure-Griffiths et al. (2005); ^e:Dickey et al. (2013)

while many others estimate WNM T_s from upper limits (i.e. T_k) or as line of sight averages in the presence of strongly absorbing CNM gas (Mebold et al. 1982, Heiles & Troland 2003, Kanekar et al. 2003, Kanekar et al 2011, Roy et al. 2013a, Roy et al. 2013b). Therefore, the mass distribution of HI in WNM-like conditions, i.e., at high T_s and low n , is statistically unconstrained.

Previous observational studies of CNM and WNM properties, although ubiquitous, typically do not reach sufficient sensitivity to detect the WNM or thermally unstable gas directly in HI absorption. The Millennium Arecibo 21 Centimeter Absorption-Line Survey (Heiles & Troland 2003; HT03), comprised of 79 HI absorption and emission spectral pairs spread over the full Arecibo Observatory field of view, provided important constraints on the T_s distribution of the CNM. A key result from HT03 is that that 48% of the non-CNM spectral features had thermally unstable temperatures, with $500 \leq T_s \leq 5000$ K. However, the high- T_s measurements in the HT03 study are indirect (i.e., not based on direct detections of HI absorption). Furthermore, single dish telescopes have the disadvantage that emission fluctuations within the beam can contaminate absorption measurements, while interferometers resolve out this large scale structure. In addition, the HT03 results could be biased by low sensitivity. Increased integration time on non-detection sightlines from HT03 reveal absorption lines with $\tau < 10^{-3}$, as shown by Stanimirović & Heiles (2005). Previously-undetected, weak absorption lines with CNM-like T_s will account for emission along the line of sight that was originally attributed to thermally unstable material, thereby reducing the estimated fraction of thermally unstable gas. Significant changes to the observed fraction of thermally unstable gas in the ISM will in turn have strong implications for the influence of any radiative and dynamic processes responsible for creating and maintaining the instability.

In a study of HI absorption at the Karl G. Jansky Very Large Array (VLA), Begum et al. (2010) demonstrated that the VLA bandpass is stable enough to detect shallow (peak $\tau \sim 10^{-3}$), wide (full width at half maximum (FWHM) $\sim 7 - 8$ km s $^{-1}$) absorption lines. In combination with HI emission from the Arecibo Observatory, they identified individual absorption lines with spin temperatures in the thermally unstable regime, $T_s = 400 - 900$ K, in observations

towards 12 bright background sources. However, only < 30% of their detected absorption lines had thermally unstable spin temperatures, which is much lower than the 48% unstable fraction reported by HT03. This emphasized the need for a larger interferometric study of HI absorption at high sensitivity to further constrain the fractions of gas in all HI phases.

Recently, Roy et al. (2013a, b) conducted an HI absorption survey of 35 continuum sources at the Westerbork Synthesis Radio Telescope (WSRT), Giant Metrewave Radio Telescope (GMRT) and Australia Telescope Compact Array (ATCA). Using the Leiden Argentine Bonn (LAB; Kalberla et al. 2005) survey for HI emission, they found at least 28% of the absorption-detected HI (by number) in the thermally unstable regime (Roy et al. 2013b). However, their T_s values are upper limits provided by linewidth-derived kinetic temperatures, and therefore their thermally unstable fraction is also an upper limit.

2.1.2 21-SPONGE

To measure the physical properties of all neutral HI in the ISM, we are conducting a large statistical survey, “21-cm Spectral Line Observations of Neutral Gas with the EVLA” (21-SPONGE), to obtain high-sensitivity Milky Way HI absorption spectra using the VLA. The recently upgraded capabilities of the VLA allow us to achieve median RMS noise levels in optical depth of $\sigma_\tau = 9 \times 10^{-4}$ per 0.42 km s^{-1} channel, which are among the most sensitive observations of HI absorption to date. Currently 31 sources are complete after over 200 hours of observing time. We have a very high detection rate so far, and we detect HI absorption in the direction of 30/31 sightlines.

In Table 2.1, we compile properties of other single dish and interferometric surveys for HI absorption for comparison, including: (1) estimates of the area covered in Galactic latitude ($^\circ$), (2) the angular resolution in HI emission (arcmin), (3) the velocity resolution (km s^{-1}), (4) the HI optical depth sensitivity (σ_τ) per 1 km s^{-1} channels, (5) the number of HI absorption spectra, and (6) the telescopes used for HI absorption and emission. Those studies which targeted particular sources for measuring HI absorption at a desired sensitivity are classified as “Absorption targeted”,

and those studies which mapped large spatial areas in HI emission and then extracted absorption from continuum sources within the map are classified as “Emission mapped”. We note that the 4 surveys in the latter category were all conducted at low- $|b|$ and therefore their spectra are more complex than those at high-latitude in the targeted absorption line surveys.

21-SPONGE is one of the most sensitive and extensive absorption line surveys ever undertaken. We emphasize that our analysis will improve on previous work at similar sensitivity, as we estimate physical properties of interstellar HI using emission data obtained with spatial resolution closer to the $\sim 1''$ interferometric absorption resolution from the VLA than previous studies. For example, Roy et al. (2013) use emission from the $\sim 36'$ -wide LAB beam, and we use emission data from the Arecibo Observatory, whose beam area ($\sim 3.5' \times 3.5'$) is smaller than the LAB beam by two orders of magnitude. In addition, 21-SPONGE will occupy a unique space in terms of sensitivity and size for a long time, given that future HI absorption surveys with the Australian Square Kilometer Array Pathfinder (ASKAP) and the Square Kilometer Array (SKA) will not likely seek such high sensitivity in optical depth (e.g. McClure-Griffiths et al. 2015).

In this paper, we present the current progress of 21-SPONGE. In Section 2, we describe the observations and data reduction; in Section 3 we discuss our methods of spectral line analysis via Gaussian fitting; in Section 4 we show example decompositions for select sources; in Section 5 we calculate temperatures and column densities for all individual component fits; in Section 6 we compare our results with previous observational surveys; in Section 7 we compare our results with synthetic observations of a hydrodynamic Galaxy simulation by Kim et al. (2014); in Section 8 we discuss the pros, cons and biases of various methods for calculating T_s from HI emission and absorption; and in Section 9, we summarize the results.

2.2 Data Processing

2.2.1 Observations

VLA observations for 21-SPONGE began in February 2011, and targeted 58 sources from the NRAO/VLA Sky Survey (NVSS) catalog with flux densities at 1.4 GHz, $S_{1.4\text{GHz}} > 3\text{ Jy}$

Table 2.2. VLA Observation Information

Source (name)	RA(J2000) (hh:mm:ss)	Dec (J2000) (dd:mm:ss)	l ($^{\circ}$)	b ($^{\circ}$)	$S_{1.4\text{GHz}}$ (Jy) ^a	σ_{τ} (10^{-3}) ^b	Synthesized Beam (arcsec 2)	Time (hrs) ^c
4C32.44	13:26:16	31:54:10	67.240	81.049	4.6	1.5	2.8×1.3	6.6
4C25.43	13:30:37	25:09:11	22.464	80.991	6.9	0.9	2.7×1.2	2.9
3C286	13:31:08	30:30:33	56.526	80.676	14.9	0.7	4.1×3.0	2.8
4C12.50	13:47:33	12:17:24	347.220	70.173	5	1.3	4.2×1.5	2.8
3C273	12:29:06	02:03:05	289.945	64.358	54.9	0.6	7.8×4.2	1
3C298	14:19:08	06:28:35	352.159	60.667	6	0.8	2.3×1.4	2.8
4C04.51	15:21:14	04:30:19	7.290	47.748	4	2.9	12.2×4.7	5.6
3C237	10:08:00	07:30:16	232.117	46.627	6.5	0.9	6.5×4.4	5
3C225A	09:42:15	13:45:51	219.866	44.025	4.4	1.5	3.9×1.6	6
3C225B	09:42:15	13:45:49	220.010	44.007	4.4	3.0	3.9×1.6	6
3C345	16:42:59	39:48:37	63.455	40.948	7	1.3	3.4×1.4	2.5
3C327.1	16:04:45	01:17:51	12.181	37.006	4.1	1.4	6.1×4.8	4.7
3C147	05:42:36	49:51:07	161.686	10.298	22.9	0.4	4.4×3.8	1.1
4C33.48	19:24:17	33:29:29	66.388	8.373	3.8	2.8	4.8×1.8	5.7
3C154	06:13:50	26:04:36	185.594	4.006	5.2	0.9	14×13	3.5
3C410	20:20:06	29:42:12	69.210	-3.768	10	1.9	9.9×5.5	2.5
B2050+36	20:52:52	36:35:35	78.858	-5.124	5	2.3	4.3×2.0	5
3C409	20:14:27	23:34:52	63.397	-6.120	14	3.3	1.6×1.4	3.3
PKS0531+19	05:34:44	19:27:21	186.761	-7.109	7	0.5	1.3×1.1	1.9
3C111	04:18:21	38:01:35	161.675	-8.821	4.3	0.7	23×23	5.1
3C133	05:02:58	25:16:24	177.725	-9.914	5.7	1.8	1.3×1.1	5.1
3C138	05:21:09	16:38:22	187.403	-11.346	9	0.9	15×5.2	2.2
3C123	04:37:04	29:40:13	170.581	-11.662	47	0.7	19×5.5	2.2
3C433	21:23:44	25:04:10	74.475	-17.693	12	2.4	9.7×5.7	3
3C120	04:33:11	05:21:15	190.373	-27.397	3.4	1.1	4.7×4.3	7
3C48	01:37:41	33:09:35	133.961	-28.720	15.7	0.7	1.3×1.2	2.8
4C16.09	03:18:57	16:28:32	166.633	-33.598	8	0.8	1.3×1.2	1.5
3C454.3	22:53:57	16:08:53	86.108	-38.182	11	0.9	7.5×3.9	2.5
J2232	22:32:36	11:43:50	77.436	-38.582	7	0.7	5.2×4.3	2.1
3C78	03:08:26	04:06:39	174.857	-44.514	5.7	4.1	4.6×3.7	5.2
3C459	23:16:35	04:05:17	83.038	-51.285	4	0.8	5.7×4.5	5

Note. — ^a: Condon et al. (1998). ^b: RMS noise in τ per 0.42 km s^{-1} channel, measured off-line. ^c: Total on-source time, not including calibration overheads.

(Condon et al. 1998). We chose most sources to lie at high galactic latitude ($|b| > 10^\circ$) to avoid complicated HI profiles associated with the Galactic plane, and also to have angular sizes less than $1'$ to avoid resolving substantial continuum flux. The range of source parameters allows us to conduct observations in all array configurations and configuration-moves except for D-array (most compact configuration). Currently, 53 sources have received full integration time and 31 of these 53 have been fully reduced and analyzed. Table 2.2 contains information about the completed observations.

All VLA observations use simultaneously three separate, standard L-band configurations, each with one dual-polarization IF of bandwidth 500 kHz and 1.95 kHz per channel spacing. One IF is centered at 1.420408 GHz (HI line rest frequency, “standard”), one at 1.421908 GHz (1.5 MHz, or about 300 km s^{-1} , higher than the HI rest frequency, called “high”) and one at 1.418908 GHz (1.5 MHz lower, called “low”). We use the high and low configurations for observing our bandpass calibrators to avoid contamination by local HI in the direction of all calibrators at the HI rest frequency. Because we normalize our solution with respect to the continuum level, the absolute phase change associated with the frequency switching method is not an issue. We use the “standard” configuration for observing target sources and measuring HI absorption, allowing for a velocity coverage of 107.5 km s^{-1} with 0.42 km s^{-1} channels, corresponding to a velocity resolution of 0.5 km s^{-1} (Rohlfs & Wilson 2004). Relative flux calibration is performed via self-calibration on the target source and phase and amplitude calibration are performed by observing a nearby VLA calibrator source.

For bandpass calibration, we observe the strong calibrators 3C286, 3C48 or 3C147 in both high and low instrument configurations following each on-source scan. For each full-length observing session (between 3 and 5 hours) we allocate up to 80% overhead for the high and low bandpass observations. However, if the source was stronger than $\sim 7 \text{ Jy}$, it was observed as its own bandpass calibrator to conserve slewing time over the course of the observation, which is particularly beneficial for filler-project length observing files (maximum 1 hour total).

2.2.2 Data reduction

All data sets were reduced using the Astronomical Image Processing System¹ (AIPS). We decided to use AIPS instead of CASA² to take advantage of the suite of bandpass analysis tasks unique to AIPS when we began the survey. During observations conducted in configuration moves involving D-array, all baselines shorter than 300 m were excluded to avoid contribution from partially resolved HI emission. After initial flagging of all observations, the task BPLOT was used to examine initial bandpass solutions for the high and low bandpass calibration observations in detail for each antenna. These solutions were then combined to produce a final bandpass solution. For all sources, we consistently reach noise levels in normalized bandpass amplitude profiles of $< 10^{-3}$ per 0.42 km s^{-1} channels.

In addition to a linear slope in the bandpass solution, we found a stable, periodic ripple which was revealed only because of the extreme sensitivity of our observations. The ripple is caused by the finite impulse response (FIR) filters applied to the data prior to correlation, and is stable in amplitude, phase and time. Appendix A contains information on the properties and stability of the bandpass structure.

Furthermore, we demonstrated that the VLA bandpass is stable enough to combine bandpass observations conducted at different times or on different days. This is especially helpful for filler observations (max total observation time ≤ 1 hour). We are able to combine bandpass observations between separate days, thereby increasing the total bandpass observation time. The noise level in the bandpass solution decreases with added time by index of -0.44 ± 0.02 , which is nearly the theoretically expected index of -0.5 (i.e. by $1/\sqrt{t_2/t_1}$ for combined integration time t_2 and initial integration time t_1). Appendix A contains further details on the noise improvement due to time averaging in the cases of additional datasets.

Following amplitude and phase calibration on the target dataset, we apply the bandpass

¹<http://www.aips.nrao.edu>

²<http://casa.nrao.edu/>

solution from the combined bandpass calibration observation. We then perform self-calibration on the target dataset by isolating the target source continuum, constructing an image, and using it to calibrate the continuum dataset. After applying the self-calibration solution, we repeat the process until the signal-to-noise level in the resulting continuum image no longer improves significantly. For most sources, this occurs after only two iterations.

Returning to the original target dataset, the continuum contribution to each source is determined by fitting a linear model to line-free channels. We subtract this from the source visibilities in all channels using the AIPS task UVLSF. We then use the task CVEL to correct source velocities for Earth's rotation and motion towards the Local Standard of Rest (LSR) within the Solar System.³

³We initially applied CVEL before continuum subtraction (UVLSF). However, we observed considerable Gibbs ringing in the edge channels of our reduced spectra. This forced us to Hanning smooth all datasets, reducing our velocity resolution by a factor of 2. After investigating the origin of the Gibbs ringing, we found that the ringing phenomenon arose in the order of CVEL and continuum subtraction, and reversing their order of application

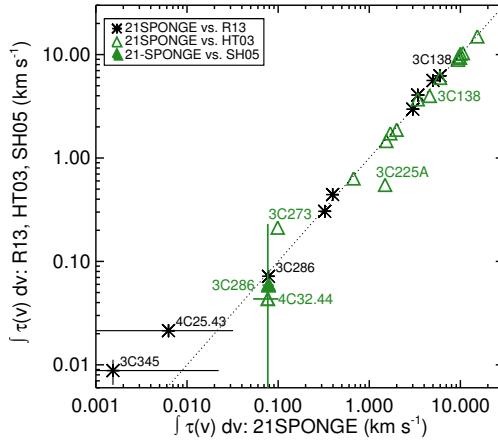


Figure 2.1 Comparing integrated optical depth values between 21-SPONGE and Roy et al. 2013 (R13; black stars), 21-SPONGE and HT03 (green diamonds) and 21-SPONGE and Stanimirović & Heiles (2005) (SH05; filled green diamond). The dotted line shows $y=x$ to highlight the agreement between the majority of points. Sources which strongly deviate from $y=x$ are indicated by their source names. Errors on the measurements are included, but are smaller than the symbol sizes unless visible.

The final calibrated data cubes are constructed using the AIPS task IMAGR, and are cleaned to three times the background noise in a test image of one channel. For each cube, the pixel sizes are calculated to be 4 times smaller than the beam size to properly sample the beam. Finally, we extract the absorption spectrum from the central pixel of each image, as in most cases, the sources are unresolved by the observations. However, of the 31 completed sources, 5 were resolved (3C327.1, 3C154, 3C111, 3C133, 3C123). To extract the spectra from these 5 data cubes, we average all spectra over the extent of the source.

The achieved sensitivities for the final data products are listed in Table 2.2, and the quoted values for σ_τ are per 0.42 km s^{-1} channel. The median RMS noise values are $\sigma_\tau = 9 \times 10^{-4}$ (mean 1×10^{-3}) per 0.42 km s^{-1} channel, and $\sigma_\tau = 6 \times 10^{-4}$ (mean 8×10^{-4}) per 1 km s^{-1} channel. Several sources have high noise due to the fact that they were resolved by the VLA, and we plan to include additional data in these directions to reduce the noise for the final data release.

2.2.3 Optical depth comparison

To verify the accuracy of our absorption spectra, we compare integrated optical depths with two previously published surveys: the Arecibo Millennium Survey (HT03) composed of 79 absorption spectra with matching expected emission spectra from the Arecibo Observatory, and 35 high-sensitivity absorption spectra from the GMRT, WSRT and ATCA by Roy et al. (2013). We compare the sources overlapping between 21-SPONGE and HT03 (green diamonds) and between 21-SPONGE and Roy et al. (2013) (black stars) in Figure 2.1.

The three catalogs generally agree well in terms of integrated optical depth. In Figure 2.1, we indicate the source names of those which differ most between the catalogs. In the cases of 4C25.43 and 3C345, and Roy et al. (2013) have better sensitivity ($\sigma_\tau = 5 \times 10^{-4}$ and 5×10^{-4} whereas we achieve $\sigma_\tau = 13 \times 10^{-4}$ and 9×10^{-4} respectively), so our $\int \tau(v)dv$ values have larger error bars. For source 3C273, the HT03 baseline is not perfectly flat, with the consequence that their $\int \tau(v)dv$ value is higher than ours. HT03 did not measure any HI absorption above

eliminated the ringing effect.

the 1σ limit towards 3C286, so we compare our optical depth spectrum to the higher-sensitivity result from Stanimirović & Heiles (2005; SH05) (filled green diamond), and observe that our integrated values are consistent. In addition, we agree very well with the Roy et al. (2013) value for 3C286. Towards source 4C32.44, HT03 have lower sensitivity and so their value has larger error bars. Finally, the HT03 profile towards 3C225A, a resolved double-lobed radio galaxy, contains “positive absorption” due to an artifact caused by the HI emission fluctuations and large uncertainties in estimating HI emission profiles.

Remaining residual differences that exist could possibly be caused by fluctuations in absorption profiles due to transient clouds, as has been well-studied in the cases of 3C147 (Lazio et al. 2009) and 3C138 (Brogan et al. 2005, Roy et al. 2012). These studies find differences in optical depth between observational epochs (spaced by several years) of up to ~ 0.5 .

Overall, the excellent agreement between 21-SPONGE absorption and HT03, Roy et al. (2013) and Stanimirović & Heiles (2005) absorption demonstrates that our observing and reduction strategies are sound. It is also encouraging to see that in the majority of cases, 21-SPONGE observations agree well with HT03 and Stanimirović & Heiles (2005). This proves that Arecibo - a single dish telescope - can reliably produce accurate HI absorption profiles in the directions of strong sources ($S_{1.4\text{GHz}} > 3\text{ Jy}$).

2.2.4 Matching emission profiles

We obtained matching HI emission profiles along most sightlines from the 305-m Arecibo Observatory as part of project a2770. For the observations, we used the L-wide receiver to simultaneously observe HI at 1420 MHz and three OH lines (1665, 1667 and 1720 MHz) with two linear polarizations. We achieve a velocity resolution of 0.16 km s^{-1} over 2048 channels between -164 and 164 km s^{-1} for the HI observations. The $\sim 3.5'$ angular resolution of the Arecibo telescope at these frequencies complements the VLA observations by minimizing the effects of mismatched beam sizes on interpreting the absorption and emission spectra. It would require prohibitively long integration times to acquire emission profiles at similar sensitivity from the VLA in D array.

Following the methods of HT03, we constructed an “expected” HI emission profile ($T_{\text{exp}}(v)$) toward each source by interpolation using a pattern of 16 observed off-source positions. The expected profiles represent the profile one would observe if the continuum source were turned off. The profile construction via least squares fit is described fully by HT03, and was also implemented by Stanimirović & Heiles (2005) and Stanimirović et al. (2014). The error in $T_{\text{exp}}(v)$ as a function of velocity, $\sigma_{T_{\text{exp}}}(v)$, is also computed in this process based on the difference in system temperature and spatial offset between on-source and off-source pointings (HT03, Section 2.7). As discussed by Stanimirović et al. (2014), we used a simpler method than HT03 and excluded fine tuning for gain variations under the assumption that correctly accounting for them requires accurate knowledge of the spatially-varying gain and beam shape. In addition, we used a second-order Taylor expansion to construct $T_{\text{exp}}(v)$, which, as found by Stanimirović et al. (2014), is a noisier but more accurate approach than using a first-order Taylor expansion. We then divided the expected profiles by a total beam efficiency factor of 0.85 (P. Perillat private communication) to convert the expected profiles to brightness temperature.

To compare our results, we scaled the expected emission profiles from HT03 to our new Arecibo expected emission profiles by their velocity integrals and found that applying a median beam efficiency factor of 0.81 to the HT03 profiles produced the best agreement with our beam efficiency-scaled profiles from a2770. For each source, we then selected the scaled emission profile with the best sensitivity. For 11/31 sources we use the HT03 expected emission profiles, and for 16/31 sources we use a2770 data. The remaining 4 sources do not have Arecibo emission profiles because they are on the edge of the Arecibo field of view, and so we use emission spectra from the Leiden Argentine Bonn (LAB) survey for these sources (Kalberla et al. 2005).

2.2.5 Stray radiation

Due to the complex beam pattern of the Arecibo telescope, our emission profiles are likely contaminated at some level by radiation from higher order side-lobes, known as stray radiation. To estimate the level of this contamination, we use additional data from the Galactic Arecibo

L-band Feed Array Survey in HI (GALFA-HI; Stanimirović et al. 2006, Peek et al. 2011). GALFA-HI is an all-sky survey for HI emission in the Galaxy using the seven beam-array ALFA receiver. Peek et al. (2011) smoothed GALFA-HI data cubes to the $\sim 36'$ angular resolution of the stray radiation-corrected LAB survey (Kalberla et al. 2005) and found that the differences generally fall within typical 1σ errors (see Figure 12 in Peek et al. 2011). Stray radiation signatures appear as broad, wing-like features at high velocities, which are easy to isolate visually from the narrower, peak-like differences due to mismatched beam shapes between the LAB survey and the smoothed GALFA-HI data cubes.

To compare our results for all sources covered by GALFA-HI to date, we constructed expected GALFA-HI emission profiles around each source by averaging 5×5 pixel 2 regions around each source. The central pixels of this average are likely affected by absorption, so we removed the central 3×3 pixel 2 region which corresponds to roughly one beam area. We then subtracted our $T_{\text{exp}}(v)$ profiles from the expected GALFA-HI profiles and found that the differences fall mostly within 3σ errors in $T_{\text{exp}}(v)$, which is very encouraging. The best method of removing the stray radiation contamination in lieu of modeling the complex, zenith-angle-dependent beam pattern of the Arecibo telescope would be to bootstrap the GALFA-HI data to the LAB survey and simply remove the differences from our profiles (e.g. as discussed by Peek et al. 2011). However, applying this method necessarily injects noise from the LAB survey into the higher-sensitivity GALFA-HI profiles. In addition, the method assumes that the contamination to GALFA-HI (observed with ALFA) is the same as to our $T_{\text{exp}}(v)$ profiles (observed with L-wide), which may not necessarily be the case. In the future we plan to investigate alternative methods for modeling the stray radiation contribution to our emission profiles without degrading our sensitivity unnecessarily.

2.3 Analysis

2.3.1 Gaussian Fitting

To estimate physical properties of individual interstellar clouds along the line of sight, we decompose the HI spectra into Gaussian functions (e.g. HT03). Although the CNM dominates the absorption spectra, by achieving extremely high sensitivities we aim to detect shallow, wide signatures of warm gas directly in absorption.

Based on the two-phase ISM model, discussed in detail by Dickey et al. (2003) and HT03, both CNM and WNM contribute to the expected profile, $T_{\text{exp}}(v)$. We denote the CNM and WNM components detected in absorption by the subscript “ABS” and the WNM components detected in emission by the subscript “EM”. Thus $T_{\text{exp}}(v)$ is given by,

$$T_{\text{exp}}(v) = T_{B,\text{ABS}}(v) + T_{B,\text{EM}}(v), \quad (2.1)$$

where $T_{B,\text{ABS}}(v)$ is the brightness temperature due to absorption-detected gas, and $T_{B,\text{EM}}(v)$ is the brightness temperature due to non-absorbing gas detected only in emission.

For each absorption spectrum, we identify a number (N) of Gaussian components using a least-squares fit (described in Section 3.3), so that the absorption spectrum $\tau(v)$ is given by,

$$\tau(v) = \sum_{n=0}^{N-1} \tau_{0,n} e^{-[(v-v_{0,n})/\delta v_n]^2}, \quad (2.2)$$

where $(\tau_0, v_{0,n}, \delta v_n)$ are the (peak optical depth, central velocity, $1/e$ width) of the n^{th} component in the fit.

Following the fit to the absorption spectrum, and assuming each fit component is independent and isothermal with spin temperature $T_{s,n}$, the brightness temperature contribution from the N Gaussian functions detected in absorption to the expected emission profile is given by,

$$T_{B,\text{ABS}}(v) = \sum_{n=0}^{N-1} T_{s,n} (1 - e^{-\tau_n(v)}) e^{-\sum_{m=0}^{M-1} \tau_m(v)}. \quad (2.3)$$

Here, the subscript m and the associated optical depth $\tau_m(v)$ run through the M number of CNM clouds lying in front of cloud n .

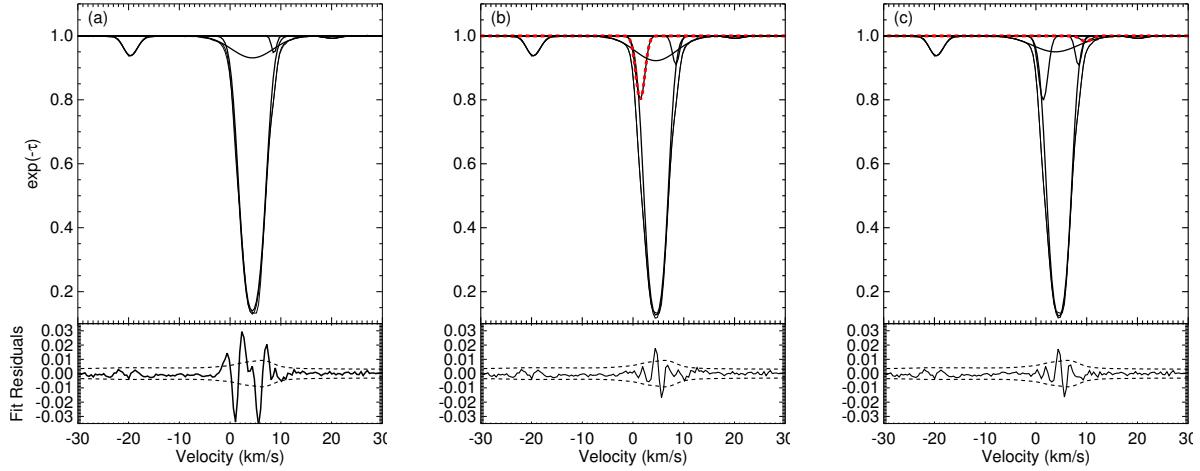


Figure 2.2 Three Gaussian fits to the optical depth profile in the direction of 3C123 demonstrating the fit improvement with additional components. Top panels: absorption spectrum ($\exp(-\tau)$) with Gaussian components overlaid; Bottom panels: fit residuals with $\pm\sigma_\tau(v)$ overlaid (see Section 2.3.2). (a): base fit, 5 components, (b): additional component added (thick red dashed), 6 total, improving the fit with 99% confidence by the F-test, so the component is retained, (c): additional component added (thick red dashed), 7 total, improving the fit with 61% confidence by the F-test, so the additional component is rejected.

We model the remaining contribution to the expected profile by WNM not detected in absorption using a set of K independent Gaussian functions. For each k^{th} component, we assume that a fraction F_k lies in front of all N absorption components, and thus a fraction $(1 - F_k)$ is absorbed by the intervening gas. The emission brightness temperature from these K non-absorbing Gaussian components is given by,

$$T_{B,\text{EM}}(v) = \sum_{k=0}^{K-1} [F_k + (1 - F_k)e^{-\tau(v)}] \times T_{0,k} e^{-[(v - v_{0,k})/\delta v_k]^2}, \quad (2.4)$$

where $(T_{0,k}, v_{0,k}, \delta v_k)$ are the (peak in units of brightness temperature, central velocity, $1/e$ width) of the k^{th} emission component. For a given set of absorption components, the expected profile is fit for the spin temperatures of the N absorption components and the Gaussian parameters of the K emission components. This fit is repeated for all permutations of absorption components along the line of sight, for every value of F_k for each emission component. Following the method of

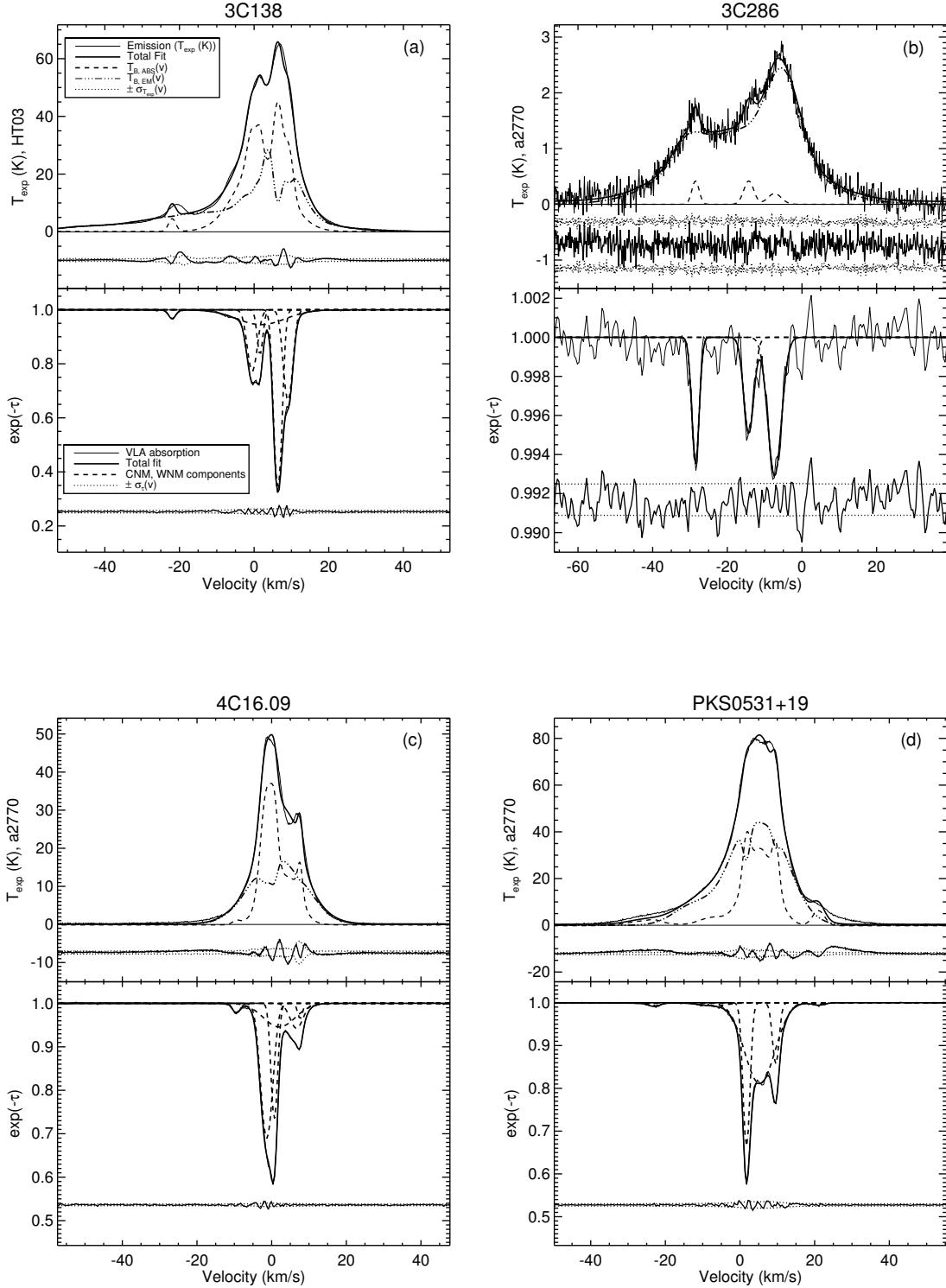


Figure 2.3 Gaussian fitting decomposition for 4/31 completed sources. The sources are displayed in two-panel plots, where the top panel displays the emission profile (from HT03, a2770, or the LAB survey) and the bottom panel displays the VLA absorption profile. The fits are overlaid in several forms, including total CNM and WNM contribution and individual components, according to the inset legend in panel (a). Residuals in the fits are offset at the bottom of each panel with

HT03, we allow F_k to be 0, 0.5, or 1. The values of F_k affect the derivation of spin temperatures, but any finer tuning of the values produces results that are difficult to distinguish statistically. Overall, there are $N! \times 3^K$ permutations for each line of sight. The final fit selected is the one with the smallest residuals, and the final spin temperatures are calculated by a weighted average over all trials, where the weights are equal to the inverse of the variance of the residuals to the $T_{\exp}(v)$ fit (HT03). The error in the final spin temperature is estimated from the variation in T_s with F_k , as described in Section 3.5 of HT03.

The expected emission profile (Equation 2.1) has been baseline-corrected, meaning that the contribution from diffuse radio continuum emission, including the cosmic microwave background (CMB) and Galactic synchrotron emission, has been removed. This value, T_{sky} , is equal to 2.725 K from the CMB plus an estimate of the Galactic synchrotron background at the source position. Given that the spectral index of the synchrotron background is about 2.7, we divide the 408 MHz brightness temperature taken from the Haslam et al. (1982) Galactic survey by a factor of $(1420/408)^{2.7}$ to estimate the synchrotron contribution at 1420 MHz. The 21-SPONGE sources

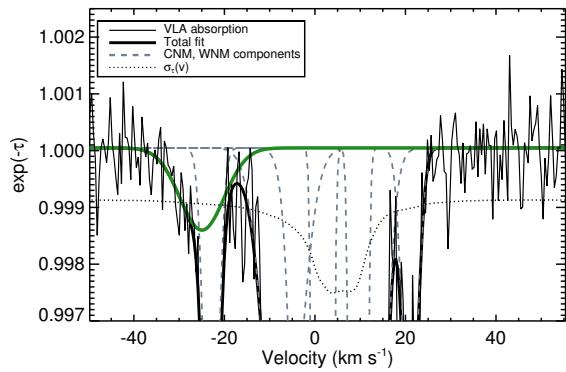


Figure 2.4 Zoom-in of a broad, warm component (green) in the complex profile towards PKS0531+19. Velocity: $v_{0,\text{LSR}} = -24.96 \pm 4.25 \text{ km s}^{-1}$, FWHM: $\Delta v = 11.5 \pm 9.3 \text{ km s}^{-1}$, peak optical depth: $\tau_{\text{peak}} = 0.15 \pm 0.17$, maximum kinetic temperature: $T_{k,\text{max}} = 2900 \pm 5000 \text{ K}$, spin temperature: $T_s = 1451 \pm 263 \text{ K}$. The dotted line indicates $\sigma_\tau(v)$.

are mostly at high latitude, and therefore the values of T_{sky} are generally small, ranging between 2.76 K and 2.85 K. To account for its presence, we add T_{sky} back to $T_{\text{exp}}(v)$ before performing the fit.

2.3.2 Absorption Noise Spectra

The noise in HI absorption spectra is not constant as a function of velocity. HI emission at Galactic velocities will increase the system temperature of the VLA antennas, producing different noise properties across the absorption spectra. To estimate the “noise spectra”, $\sigma_\tau(v)$, of the VLA absorption profiles, we follow the methods discussed by Roy et al. (2013a). The spectral noise is composed of (1) on-source noise, $\sigma_{\text{on}}(v)$ (varies with velocity) and (2) off-line noise from the frequency-switched bandpass solution, σ_{BP} (constant with velocity). We first estimate the RMS noise in off-line channels, σ_τ (see Table 2), which includes contributions from both $\sigma_{\text{on}}(v)$ and σ_{BP} . We then scale σ_τ by $((T_{\text{B}}(v) + T_{\text{sys}})/T_{\text{sys}})$ to calculate $\sigma_{\text{on}}(v)$. The antenna used to produce the LAB survey is of similar size to an individual VLA antenna, so we use HI emission from the LAB survey to estimate the brightness temperature as a function of frequency for each source ($T_{\text{B}}(v)$) and we assume that the system temperature of the VLA in the absence of HI emission is $T_{\text{sys}} \sim 25$ K. Finally, we calculate the noise spectrum by: $\sigma_\tau(v)^2 = \sigma_{\text{on}}(v)^2 + \sigma_{\text{BP}}^2$ (also see discussion by Roy et al. 2013a). Examples of these noise spectra are displayed as dotted lines above and below the residuals of the Gaussian fits in Figures 2.2, 2.3, and 2.4. We use the noise spectra to visually inspect the fit residuals, not to quantify the goodness-of-fit.

2.3.3 Component Selection: Absorption

The process of Gaussian fitting has many inherent uncertainties, including the determination of number of fit components. The noise level in the fit residuals can be used as a goodness-of-fit indicator, but it cannot easily distinguish between different numbers of components, because adding more components will always reduce the residual noise level without necessarily bringing the fit any closer to an accurate depiction of reality.

To determine the best-fit number of Gaussian components to the absorption spectra, we begin by using a reduced chi-squared (χ_{red}^2) test. We calculate χ_{red}^2 using,

$$\chi_{\text{red}}^2 = \frac{1}{\nu} \sum \frac{(O(v) - M(v))^2}{\sigma_\tau^2}, \quad (2.5)$$

where $O(v)$ are the observed data, $M(v)$ are the fitted model data, σ_τ is off-line RMS noise in the absorption spectrum (from Table 2) and ν is the number of degrees of freedom (DOF: equal to number of measurements (256 spectral channels) minus the number of fitted parameters, which includes the height, width and center of each Gaussian function and a constant continuum offset).

For each profile, we would like to include the minimum number of components that brings χ_{red}^2 near to 1.0. However, in many cases, the difference in χ_{red}^2 between fits of different numbers of components is too small to confidently determine the best fit. As an example, we present the case of 3C123 in Figure 2.2. The central component of the absorption spectrum is very strong, and is poorly fit by a single Gaussian component. Including an additional broad component at the base of the central component does improve the fit, and including additional narrow lines improves the fit further.

To compare the χ_{red}^2 values of different possible Gaussian fits, such as those shown for 3C123 in Figure 2.2, we employ an F-test (e.g., Westmoquette et al. 2007, Dawson et al. 2011) using the IDL procedure mpftest.pro⁴. The F-test allows us to quantify the improvement in χ_{red}^2 as a function of added DOF to the Gaussian model. First, the procedure computes the F-value, which is defined as a ratio of χ_{red}^2 and ν values for two models:

$$F = \frac{\chi_{\text{red},1}^2/\nu_1}{\chi_{\text{red},2}^2/\nu_2}, \quad (2.6)$$

where the subscript “1” denotes the base model which involves the fewest number of Gaussian components, and “2” denotes a model including an additional component (i.e., three additional DOF). The probability density function $p_f(f; \nu_1, \nu_2)$ for all F-values (f) described by DOF ν_1 and

⁴<http://purl.com/net/mpfit>

ν_2 is given by⁵,

$$p_f(f; \nu_1, \nu_2) = \frac{\Gamma[(\nu_1 + \nu_2/2)]}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2} \times \frac{f^{1/2(\nu_1-2)}}{(1 + f\nu_1/\nu_2)^{1/2(\nu_1+\nu_2)}}. \quad (2.7)$$

From this we calculate the probability that a value f drawn from $p_f(f; \nu_1, \nu_2)$ equals or exceeds the given value of F (Equation 2.6), which is given by,

$$P_F(F; \nu_1, \nu_2) = \int_F^\infty df p_f(f; \nu_1, \nu_2), \quad (2.8)$$

where P_F (or confidence level $C_F = 1 - P_F$) is a measure of the probability that the added component merits inclusion in the model.

For the majority of sources, the inclusion of additional Gaussian components results in $C_F > 0.99$ before C_F drops far below $C_F \sim 0.9$ and indicates low confidence in the most recent component-addition. In three cases (sources 3C111, 3C133 and 3C459), the fit parameters varied strongly with the initial guesses, and we observed that the fits remained stable between the final component-additions if we applied a cutoff value of $C_F = 0.97$. Therefore, to apply a uniform standard to all sources, we use a strict threshold of $C_F = 0.97$ for including components, noting that in all but three cases, the effective cutoff is $C_F > 0.99$.

As an example, in the case of 3C123, the confidence level of including one additional component (Figure 2.2b) to the base fit (Figure 2.2a) is $C_F = 0.99$, indicating that the additional component exists with high confidence. The confidence level of including a second additional component (Figure 2.2c) in addition to the new fit (Figure 2.2b) is $C_F = 0.61$, and therefore we reject this additional component.

2.3.4 Component Selection: Emission

The selection of Gaussian components for the expected emission profiles is more difficult. Taking the absorption line fits as fixed in central velocity, peak optical depth and line width (FWHM), we simultaneously solve for the Gaussian parameters of additional emission components

⁵NIST/SEMATECH e-Handbook of Statistical Methods, <http://www.itl.nist.gov/div898/handbook/>, 01/2015.

and the spin temperatures of the absorption components given emission information along the line of sight (Equations 2.3 and 2.4). Considering the effects of stray radiation on the emission profiles and the complexity of the fits, we select a minimum number of components that brings the fit residuals below the $3\sigma_{T_{\text{exp}}}(v)$ level. We then apply a similar F-test as in the absorption component selection and add additional components until the confidence level of the fit improvement drops below the cutoff ($C_F = 0.97$).

Given the final selection of Gaussian components, we then solve the radiative transfer equations again for all possible orderings of the N absorption components along the line of sight. We repeat this line of sight iteration for all permutations of F_k , the fraction of WNM unabsorbed by intervening CNM (see Section 3.1) among the K emission components. If the total number of absorption components was large (defined here as greater than or equal to 6), we only varied the order of the most blended components at the center of the spectrum. The effect of different CNM cloud order on T_s is very small compared with the effect of different values of F_k per emission component, therefore we held the more isolated components in these complex spectra fixed in their LOS order with little consequence to the derived T_s values for the overall fit.

2.4 Selected Examples

To illustrate the fitting method further, we select four example sources and display their spectra and Gaussian decompositions in Figure 2.3. The top panel of each plot contains the emission profile, $T_{\text{exp}}(v)$, from HT03, a2770, or LAB (thin solid), the total fit to $T_{\text{exp}}(v)$ (thick solid), the total absorption contribution to the emission profile, $T_{\text{B, ABS}}(v)$ (Equation 1, dashed), and the total emission-only contribution, $T_{\text{B, EM}}(v)$ (Equation 1, triple dot-dash). The bottom panel of each plot contains the VLA absorption profile (thin solid), total fit to the absorption profile (thick solid), and individual Gaussian components (dashed). The residuals of each fit are plotted at the bottom of each panel, with noise envelopes $\sigma_{T_{\text{exp}}}(v)$ (Section 2.4) and $\pm\sigma_{\tau}(v)$ (Section 3.2) to illustrate the goodness of fit (dotted).

We compare in detail our Gaussian decomposition with overlapping HI absorption line

surveys, including HT03, Mohan et al. (2004) and Roy et al. (2013b). All three surveys fit Gaussian functions to their absorption lines and followed different strategies. For example, Roy et al. (2013b) fit increasing numbers of Gaussian functions until χ^2_{red} for the fit was as close to 1 as possible. This resulted in a minimum of 1, maximum of 20 and of median of 7 Gaussian components over their 30 spectra. This approach requires detailed knowledge of the spectral noise properties, as well as the assumption that Gaussian functions are the proper representation of all components, which can fail if line broadening is not due to thermal and/or turbulence motions. Our method also makes these assumptions, but we fit fewer components per spectrum (minimum of 1, maximum of 11 and median of 5) to not “over-fit” at the expense of inferior χ^2_{red} values for the fits. In another approach, HT03 stress that human judgment is necessary for producing reliable Gaussian fits. For example, two fitted components with similar central velocities and FWHMs could in reality be a single component with a non-Gaussian intrinsic lineshape. No Gaussian decomposition is a unique solution, and it is impossible to prove which solution is most physically relevant, but the differences between our respective Gaussian fitting strategies will affect the comparison of our results. We emphasize that our F-test method for component selection imposes a uniform standard to a non-unique fitting problem, and aims to make our decisions reproducible.

2.4.1 3C138

3C138 is included in the Millennium Arecibo survey (HT03), and they identified many similar, strong components, as we do, in their Gaussian decomposition. As seen in Figure 2.3a, in addition to narrow cold components at velocities 9.1, 6.4, 1.6 and -0.5 km/s, we identify a shallow broad underlying component, centered at 1.9 ± 0.3 km/s with a FWHM of 15.1 ± 0.6 km s $^{-1}$. This corresponds to a maximum kinetic temperature of $T_{k,\text{max}} = 5000$ K, and with a fit to the emission profile (from HT03), corresponds to a spin temperature of 456 ± 40 K. The maximum spin temperature derived by HT03 for an absorption component in the direction of this source is $T_s = 380 \pm 23$ K, corresponding to a similar component at 1.8 km s $^{-1}$.

This source is also included in the Roy et al. (2013b) WSRT/GMRT/ATCA survey. Whereas we fit a total of 6 components before the fit ceases to improve significantly by the F-test, Roy et al. (2013b) fit 13 components. Given the complex nature of the spectrum, and the large difference in total numbers of components, there are few closely matching components between the two decompositions.

2.4.2 3C286

With a sensitivity of $\sigma_\tau = 9 \times 10^{-3}$, HT03 were unable to detect HI absorption in the direction of 3C286. Stanimirović & Heiles (2005) demonstrate that with additional integration times at the Arecibo Observatory, weak absorption lines below the HT03 sensitivity threshold are easily detected, which can have important implications for HT03's measured fractions of HI in the CNM, WNM and thermally unstable phases. For 3C286, they find 3 velocity components at -28.8 , -14.3 and -7.4 km s^{-1} , with $T_{k,\max} = 106$, 115 and 315 K and $T_s = 89 \pm 7$, 37 ± 4 and $30 \pm 20 \text{ K}$. We find similar velocity components, at -28.5 , -14.3 and -7.3 km s^{-1} , with $T_{k,\max} = 116$, 220 and 405 K and $T_s = 63 \pm 13$, 84 ± 12 and $27 \pm 9 \text{ K}$. Roy et al. (2013b) find the same three components, in addition to a broad line centered at -8.4 km s^{-1} with peak optical depth $\tau = 0.001$ and $T_{k,\max} = 5140 \text{ K}$. Their achieved sensitivity of $\sigma_\tau = 3 \times 10^{-4}$ allows them to detect this weak component whereas our sensitivity of $\sigma_\tau = 7 \times 10^{-4}$ puts it below our detection limit. We note that they do not discuss the effect of bandpass instability on their detection of broad, weak lines.

2.4.3 4C16.09

Although 4C16.09 is not included in the Millennium survey (HT03), it is included in the Mohan et al. (2004) GMRT survey of HI absorption. They achieve a mean survey sensitivity in optical depth of $\sigma_\tau \sim 3 \times 10^{-3}$ and the spectra have velocity resolution of 2.1 km s^{-1} over roughly 350 km s^{-1} . Our absorption profile agrees well in shape with their results. They detect two narrow absorption components, at velocities 7.5 and 0.2 km s^{-1} with corresponding spin temperatures

$T_s = 30\text{ K}$ and 45 K respectively (Mohan et al. 2004). We detect components at similar velocities, 7.6 and 0.7 km s^{-1} , in addition to four additional absorption components (see Figure 2.3c). We identify two additional, broad emission components which are prominent in the expected emission profile and produce a maximum spin temperature of $T_s = 240 \pm 10\text{ K}$. In comparison, Roy et al. (2013b) have the same sensitivity of $\sigma_\tau = 8 \times 10^{-4}$ and find 8 Gaussian components.

2.4.4 PKS0531+19

The profile in the direction of PKS0531+19 is highly complex. Detected by HT03 in the Millennium Survey, they identify four narrow absorption components at velocities $9.6, 5.5, 1.8$ and -23.1 km s^{-1} (HT03). We detect features at similar velocities, in addition to two broad, shallow components centered at 21 and -25 km s^{-1} corresponding to spin temperatures of $T_s = 1129 \pm 74\text{ K}$ and $1451 \pm 263\text{ K}$. In Figure 2.4 we zoom in on the second of these components, shown by the thick green line. This source is a good example of the benefit of high sensitivity. The RMS noise in the absorption profile is $\sigma_\tau = 5 \times 10^{-4}$, so that the broad component with peak optical depth equal to $\tau = 0.0015$ is a 3σ detection. HT03 achieve a sensitivity of $\sigma_\tau \sim 0.0015$ in their absorption profile towards this source, and therefore did not fit a Gaussian function to this line given that it fell below their 1σ level.

2.5 Properties of Cold and Warm Neutral Gas

Following complete decomposition of all absorption and emission profiles with Gaussian functions, we can use estimates of temperature and column density to infer properties of the CNM and WNM.

2.5.1 Temperature

The FWHM (Δv) of each Gaussian function determines its maximum kinetic temperature, $T_{k,\max}$, by

$$T_{k,\max} = m_H / (8k_B \ln(2)) \times \Delta v^2 = 21.866 \times \Delta v^2, \quad (2.9)$$

for hydrogen mass m_H , Boltzmann's constant k_B , and FWHM Δv in km s^{-1} (e.g., HT03). This quantity contains contributions from both thermal broadening and turbulent gas motions so that,

$$T_{k,\max} = T_k + \frac{m_H v_{\text{turb}}^2}{2k_B}, \quad (2.10)$$

where T_k is the true kinetic temperature of the gas (e.g. Liszt 2001). Therefore, $T_{k,\max}$ is an upper limit to the kinetic temperature in the presence of turbulent broadening. As shown by Heiles & Troland (2003b), nonthermal motions in the CNM are characterized by a turbulent Mach number (M_t) of $M_t \sim 3$, implying that CNM gas is supersonic. To estimate M_t , we solve Equation (2.10) for v_{turb}^2 , and then assume that this line of sight turbulent velocity is $\sqrt{3}$ times the three-dimensional turbulent velocity ($v_{\text{turb,3D}}$). The turbulent Mach number M_t is given by $v_{\text{turb,3D}}^2/C_s^2$, where C_s is the isothermal sound speed. We use C_s because thermal equilibrium is established quickly in the CNM, which dominates the absorption profiles (Heiles & Troland 2003b). Following Heiles & Troland (2003b), we adopt $C_s^2 = kT_s/1.4m_H$ for mean atomic weight $1.4m_H$. Therefore, the turbulent Mach number is given by,

$$M_t^2 = \frac{v_{\text{turb,3D}}^2}{C_s^2} = 4.2 \left(\frac{T_{k,\max}}{T_s} - 1 \right), \quad (2.11)$$

which is Equation (17) from Heiles & Troland (2003b). In Figure 2.5, we display a histogram of M_t for all 21-SPONGE absorption-detected Gaussian components, which has a peak at $M_t \sim 3$ (median $M_t = 2.86$) and agrees with the results from HT03. The distribution is strongly peaked, and 40% of all components have Mach numbers between 2 and 4.

The spin temperature is determined by the ambient radiation field, collisions between hydrogen atoms, electrons and protons, and Ly α scattering. For the CNM, high densities allow the collisional contribution to dominate 21-cm excitation so that T_s is expected to be equal to T_k . At low densities, i.e. in more diffuse WNM gas, collisions are not sufficient to thermalize the 21-cm transition and therefore T_s is expected to be less than T_k (e.g. Liszt 2001).

In the case of 12 sources, one or more of the Gaussian functions fit to the absorption spectrum did not have a corresponding spectral feature at the same velocity in emission (likely

due to the cold component occupying a very small solid angle, Heiles & Troland 2003a). Therefore, the converged T_s values for these components are unreasonably low, < 1 K. Given the uncertain presence of stray radiation in the Arecibo emission profiles, we take a conservative approach for estimating T_s for these components. We adopt the median M_t value measured over all fitted Gaussian functions, $M_t = 2.86$ and solve Equation (2.11) for T_s and its error given $T_{k,\max}$ and its error.

2.5.2 Column Density

We estimate the HI column density of the absorbing HI using the expression:

$$N(\text{HI, ABS}) = C_0 \cdot T_s \int \tau(v) \, dv \quad (2.12)$$

where $C_0 = 1.823 \times 10^{18} \text{ cm}^{-2} \text{ K}^{-1} (\text{km s}^{-1})^{-1}$, T_s is the derived spin temperature and $\tau(v)$ is the optical depth as a function of velocity in km s^{-1} . For each n^{th} absorption component, we use parameters from the Gaussian fit to calculate the column density, so that,

$$N(\text{HI, ABS})_n = 1.064 \times C_0 \times T_{s,n} \times \tau_{0,\text{peak},n} \times \Delta v_n, \quad (2.13)$$

where $(T_{s,n}, \tau_{0,\text{peak},n}, \Delta v_n)$ are the (spin temperature, peak optical depth and FWHM in km s^{-1}) of the n^{th} component and 1.064 converts the product to the area under a Gaussian function with

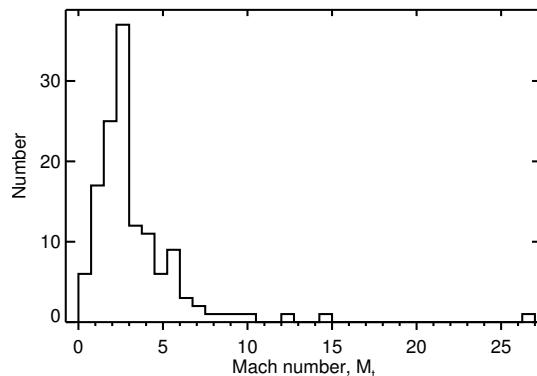


Figure 2.5 Histogram of turbulent Mach number M_t for all absorption-detected Gaussian components, calculated from Equation (2.11) (Heiles & Troland 2003b Equation 17).

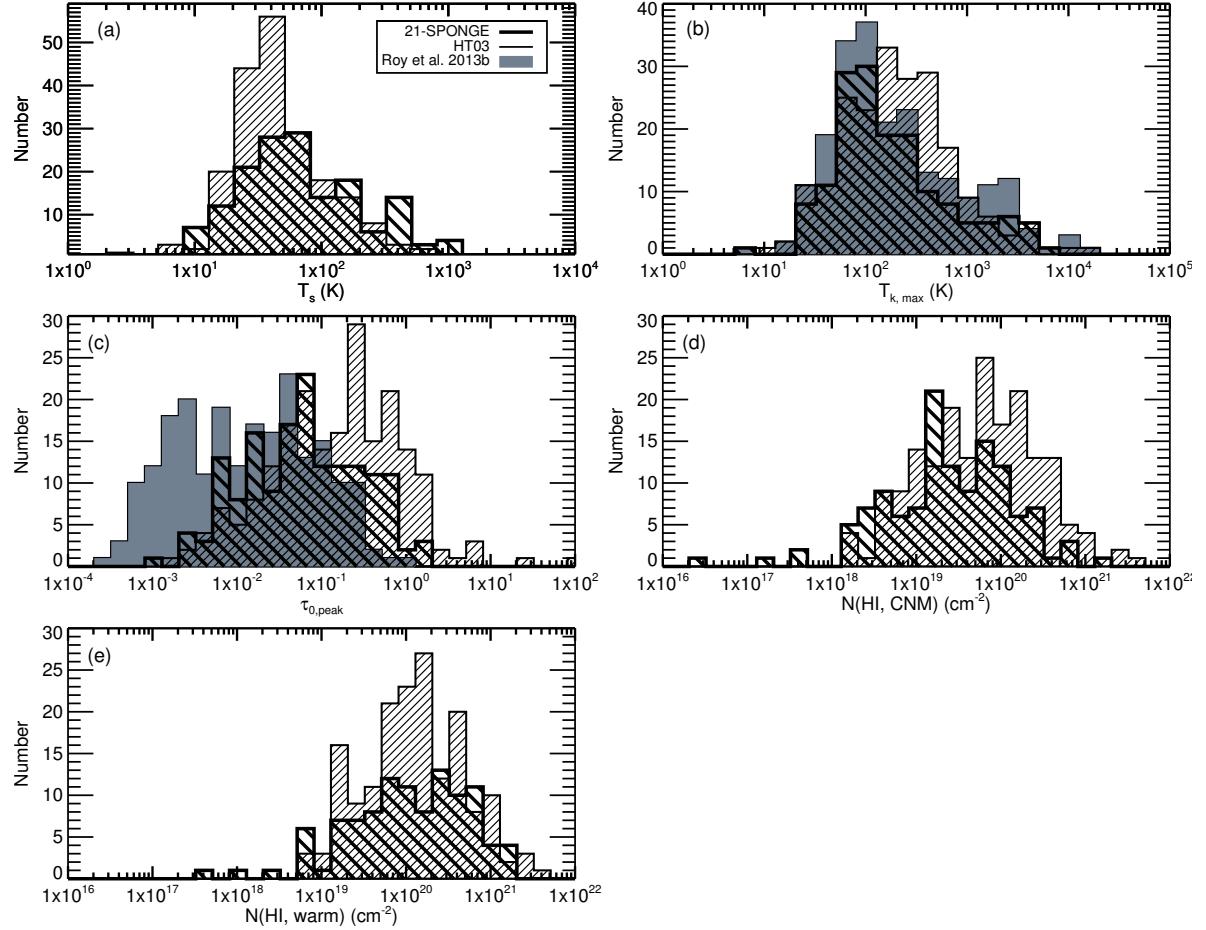


Figure 2.6 Histograms of various properties from the Gaussian fits to the current 31 sources in 21-SPONGE (thick black), including: (a) T_s (K), (b) $T_{k,\max}$ (K), (c) peak optical depth in absorption, $\tau_{0,\text{peak}}$, (d) $N(\text{HI, CNM})$: column density of all CNM components, defined as those with $T_s < 200$ K, (d) $N(\text{HI, warm})$: column density of all “warm” components, defined as those with $T_s > 200$ K and those detected only in emission. The results for the Millennium Survey (HT03, thin black) and the WSRT/GMRT/ATCA survey by Roy et al. (2013b) (shaded gray) are included for comparison. Results from Roy et al. (2013b) are not included in panels a, d and e because they do not compute these parameters for individual Gaussian components.

the given height and width. For the emission components not detected in absorption, we estimate the column density as,

$$N(HI, EM) = C_0 \int T_B(v) dv, \quad (2.14)$$

where T_B is the brightness temperature. Using the Gaussian fit to the emission profile, we estimate the column density associated with each k^{th} non-absorbing component as,

$$N(HI, EM)_k = 1.064 \times C_0 \times T_{0,k} \times \Delta v_k, \quad (2.15)$$

where $T_{0,k}$ and Δv_k are the peak and FWHM in km s^{-1} of the k^{th} component.

2.6 Comparison with previous surveys

In Figure 2.6, for all absorption-detected Gaussian functions we display histograms of: (a) spin temperature (T_s), (b) maximum kinetic temperature ($T_{k,\max}$), (c) peak optical depth ($\tau_{0,\text{peak}}$), (d) column density for all CNM clouds, defined as those with $T_s \leq 200 \text{ K}$, ($N(HI, \text{CNM})$), and finally (e) column density for all warm clouds, defined as those emission-only components and absorption-detected components with $T_s \geq 200 \text{ K}$ ($N(HI, \text{warm})$). We overlay the same quantities

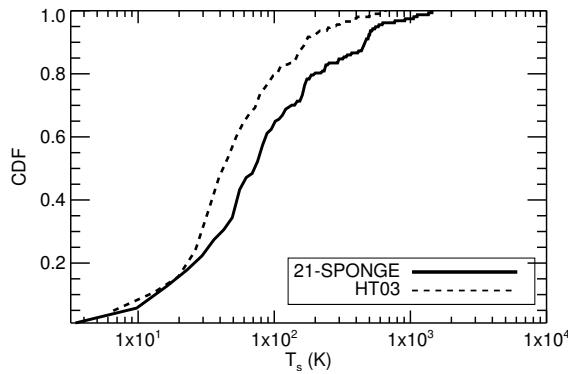


Figure 2.7 Cumulative distribution function of the T_s values shown in Figure 2.6a for 21-SPONGE and HT03.

for the 79 sources from the Millennium survey (HT03) in the panels of Figure 2.6 for comparison.

The superior sensitivity of 21-SPONGE allows us to detect absorption lines with smaller τ_{peak} and/or higher T_s than in HT03, which explains the poor agreement between the 21-SPONGE and HT03 distributions in Figures 2.6a and 2.6c. We are therefore more sensitive to low-N(HI, CNM), as shown in Figure 2.6d and in Equation (2.13).

In Figure 2.6a, we detect a strong tail of high- T_s components in comparison with HT03. This results in our median $T_s = 77$ K, which is higher than their median value of $T_s = 48$ K. The maximum T_s we detect is $T_s = 1451 \pm 263$ K, compared with the maximum value measured by HT03 of $T_s = 656$ K, demonstrating that it is possible to directly measure T_s from warm gas with improved sensitivity in HI absorption.

However, although we have the sensitivity to detect HI at higher spin temperatures than 1500 K directly in absorption, we only detect 32/157 Gaussian components above the expected range for the CNM ($T_s = 40 - 200$ K from Wolfire et al. 2003), and only 4/157 above $T_s = 1000$ K. This is a surprising result from 21-SPONGE so far, given that HT03 estimate (based on indirect

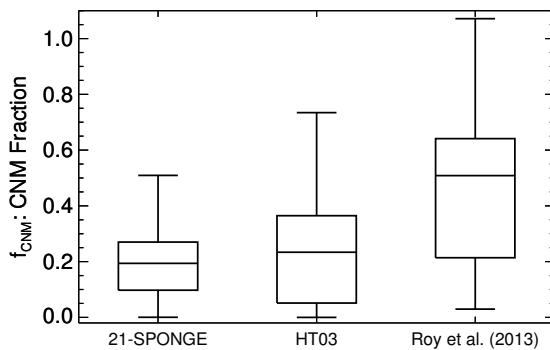


Figure 2.8 “Box and whisker” plot comparing the CNM fractions (f_{CNM}) for 21-SPONGE with HT03 and Roy et al. (2013). Each box spans the 25th through 75th percentiles of each distribution, the median value is indicated by the horizontal line within each box, and the bars extend to the minimum and maximum values.

Table 2.3. Line of sight column density parameters

Source	$\Sigma N(HI, CNM)_{20}$ (10^{20} cm^{-2})	$\Sigma N(HI, \text{warm})_{20}$ (10^{20} cm^{-2})	$N(HI, \text{total})_{20}$ (10^{20} cm^{-2})	f_{CNM}
4C32.44	0.17 ± 0.09	1.19 ± 0.03	1.36 ± 0.09	0.12 ± 0.06
3C286	0.06 ± 0.02	1.28 ± 0.09	1.34 ± 0.09	0.05 ± 0.02
4C12.50	0.36 ± 0.11	2.01 ± 0.03	2.37 ± 0.11	0.15 ± 0.05
3C273	0.05 ± 0.19	2.32 ± 0.03	2.37 ± 0.19	0.02 ± 0.08
3C298	0.06 ± 0.37	2.33 ± 0.09	2.39 ± 0.38	0.03 ± 0.15
4C04.51	0.24 ± 0.06	4.31 ± 0.11	4.55 ± 0.13	0.05 ± 0.01
3C237	0.17 ± 0.20	2.43 ± 0.11	2.60 ± 0.23	0.07 ± 0.08
3C225A	1.19 ± 0.30	3.05 ± 0.06	4.24 ± 0.30	0.28 ± 0.07
3C225B	0.98 ± 0.18	3.57 ± 0.11	4.55 ± 0.21	0.22 ± 0.04
3C345 ^a	0.00 ± 0.00	1.01 ± 0.13	1.01 ± 0.13	0.00 ± 0.00
3C327.1	1.76 ± 1.17	6.96 ± 0.11	8.72 ± 1.18	0.20 ± 0.14
3C147	7.44 ± 1.79	13.15 ± 0.78	20.59 ± 1.95	0.36 ± 0.09
4C33.48	1.54 ± 3.00	14.22 ± 0.48	15.76 ± 3.04	0.10 ± 0.19
3C154	9.75 ± 14.94	36.19 ± 0.49	45.94 ± 14.94	0.21 ± 0.33
3C410	30.39 ± 6.64	29.29 ± 0.81	59.69 ± 6.69	0.51 ± 0.13
B2050	3.00 ± 2.02	30.86 ± 1.33	33.86 ± 2.42	0.09 ± 0.06
3C409	5.33 ± 3.49	26.36 ± 0.55	31.69 ± 3.53	0.17 ± 0.11
PKS0531	8.39 ± 9.61	20.45 ± 1.48	28.84 ± 9.72	0.29 ± 0.35
3C111	3.99 ± 3.78	32.44 ± 2.29	36.43 ± 4.42	0.11 ± 0.10
3C133	9.76 ± 4.02	26.22 ± 0.46	35.98 ± 4.05	0.27 ± 0.12
3C138	6.01 ± 2.83	20.42 ± 0.58	26.43 ± 2.89	0.23 ± 0.11
3C123	5.15 ± 4.41	19.41 ± 1.90	24.55 ± 4.80	0.21 ± 0.18
3C433	1.15 ± 2.09	9.23 ± 0.13	10.38 ± 2.10	0.11 ± 0.20
3C120	10.86 ± 2.24	11.36 ± 0.26	22.22 ± 2.26	0.49 ± 0.11
3C48	0.70 ± 0.36	4.25 ± 0.19	4.95 ± 0.41	0.14 ± 0.07
4C16.09	3.27 ± 1.10	8.83 ± 0.24	12.10 ± 1.13	0.27 ± 0.09
3C454.3	1.99 ± 15.74	5.92 ± 0.63	7.90 ± 15.75	0.25 ± 2.05
J2232	1.21 ± 0.37	4.29 ± 0.17	5.50 ± 0.41	0.22 ± 0.07
3C78	3.81 ± 1.91	9.93 ± 0.27	13.74 ± 1.93	0.28 ± 0.14
3C459	1.17 ± 0.49	5.12 ± 0.04	6.29 ± 0.49	0.19 ± 0.08

Note. — ^a: $N(HI, CNM) = 3 \pm 1 \times 10^{16} \text{ cm}^{-2}$.

measurements from HI emission only) that 48% of components should be in the thermally unstable range of $500 \leq T_s \leq 5000$ K. Our current estimate for the thermally unstable fraction in this same range— $\sim 20\%$ by number—is more consistent with estimates from colliding flow models (e.g., Audit & Hennebelle 2005).

In Figure 2.7, we present cumulative distribution functions of the data in Figure 2.6a. A K-S test indicates that the two distributions are highly unlikely to be drawn from the same parent population (K-S = 0.23, $p = 0.00$). The 21-SPONGE distribution is much shallower at high temperature, indicating again that we are detecting more gas at higher T_s than was done by HT03 using similar analysis methods.

In Figures 2.6b and 2.6c we also include Gaussian decomposition results from the survey presented in Roy et al. (2013b) (grey shaded histograms; see Table 2.1 for survey information). Our two surveys have similar sensitivity in absorption. Following Gaussian decomposition of their absorption spectra, Roy et al. (2013b) do not estimate T_s , $N(\text{HI, CNM})$ or $N(\text{HI, warm})$ for their Gaussian components. As discussed in Section 2.4, Roy et al. (2013b) fit many more components per spectrum than we do, and therefore they detect many more low- τ lines than either we or HT03 do, as seen by the strong low- τ tail in the gray-shaded histogram in Figure 2.6c.

We agree well with both Roy et al. (2013b) and HT03 in $T_{k,\max}$, shown in Figure 2.6b. From Equation (2.9), this implies that we are sensitive to similar absorption line widths. We find a median value of $T_{k,\max} = 160$ K, compared with HT03’s median value of $T_{k,\max} = 220$ K and Roy et al. (2013b)’s value of $T_{k,\max} = 160$ K. We find that 5.7% of components have $T_{k,\max} \leq 40$ K, which is below the minimum theoretically expected kinetic temperature for the CNM (e.g. Wolfire et al. 2003), compared with 4.7% by HT03 and 7.0% by Roy et al. (2013). At the extreme end of the low- $T_{k,\max}$ values, we find one weak component with $T_{k,\max} = 9 \pm 4$ K in the direction of 3C345. These low-temperature components can be explained by ineffective photoelectric heating from dust and polycyclic aromatic hydrocarbons (PAHs) along those sightlines, which allows for the existence of very cold clouds with $T_{k,\max} \sim 10$ K (Wolfire et al. 1995, Heiles & Troland 2003b,

Roy et al. 2013b).

In Figures 2.6d and 2.6e, it is clear that we generally agree with HT03 in measurements of $N(\text{H}\text{I}, \text{CNM})$ and $N(\text{H}\text{I}, \text{warm})$. However, our superior sensitivity in optical depth allows us to detect lines with much lower $N(\text{H}\text{I}, \text{CNM})$. For example, we are able to detect an individual absorption line with $N(\text{H}\text{I}, \text{CNM}) = 3 \pm 1 \times 10^{16} \text{ cm}^{-2}$, whereas HT03 measure a minimum column density of $2 \times 10^{18} \text{ cm}^{-2}$ for individual CNM clouds. An important result from the final 21-SPONGE data release will be constraints on the low- $N(\text{H}\text{I})$ end of the CNM column density distribution. McKee & Ostriker (1977) predicted a lower-limit to CNM column density of $N(\text{H}\text{I}, \text{CNM}) \sim 10^{17} - 10^{18} \text{ cm}^{-2}$, derived from balancing radiative losses and conductive heating at the interface between the CNM and WNM at various possible temperatures. As discussed by Stanimirović & Heiles (2005), observational constraints on this lower limit (and its prevalence) can distinguish between different formation and evolution histories of these structures, whether by evaporation into the surrounding medium, or transient, dynamic formation at shock interfaces, or condensation from turbulent flows.

2.6.1 Line of sight properties

In Table 2.3 we compute line of sight column density-related quantities for the current 21-SPONGE sources, where the subscript “20” indicates units of 10^{20} cm^{-2} . These quantities include: (a) the total CNM column density ($\Sigma N(\text{H}\text{I}, \text{CNM})_{20}$), equal to the sum of the column densities of the absorption-detected Gaussian components with $T_s < 200 \text{ K}$; (b) the total WNM column density ($\Sigma N(\text{H}\text{I}, \text{warm})_{20}$), equal to the sum of the column densities of the absorption-detected Gaussian components with $T_s \geq 200 \text{ K}$ plus the column densities of all emission-only Gaussian components (which is a combination of both thermally-stable WNM and thermally-unstable gas); (c) the total H_I column density per line of sight ($N(\text{H}\text{I}, \text{total})_{20} = \Sigma N(\text{H}\text{I}, \text{CNM})_{20} + \Sigma N(\text{H}\text{I}, \text{warm})_{20}$); (d) the CNM fraction, f_{CNM} , defined as:

$$f_{\text{CNM}} = \frac{\Sigma N(\text{H}\text{I}, \text{CNM})}{N(\text{H}\text{I}, \text{total})}. \quad (2.16)$$

To compare our results, we obtained or calculated the same quantities for the HT03 and Roy et al. (2013b) samples. For HT03, we used their published fit information to calculate all quantities in the same manner as for 21-SPONGE, described above. To calculate $\Sigma N(HI, CNM)_{20}$, Roy et al. (2013b) began by assuming that all absorbing gas is from the CNM with $T_s = 200$ K and solved Equation (2.12) to produce upper limits to $\Sigma N(HI, CNM)_{20}$ for each line of sight. We estimated $\Sigma N(HI, \text{warm})_{20}$ from their results by subtracting $\Sigma N(HI, CNM)_{20}$ from $N(HI, \text{total})_{20}$, which they calculated using the “isothermal” method (Spitzer 1978, Dickey & Benson 1982, Chengalur et al. 2013).

We generally agree well with the results from HT03 and Roy et al. (2013b). All three studies measure the same range in $\Sigma N(HI, CNM)_{20}$, between ~ 0.01 and ~ 10 . We find a median $\Sigma N(HI, CNM)_{20} = 1.76$, compared with 1.45 by HT03 and 1.76 by Roy et al. (2013). From Figure 2.7, we detect many more components in absorption with $T_s > 200$ K, which have high $N(HI, ABS)_n$ according to Equation (2.13) and which will increase the total $\Sigma N(HI, \text{warm})_{20}$ along the line of sight. We measure median $\Sigma N(HI, \text{warm})_{20} = 8.83$, which is much higher than HT03’s median value of 3.97 and Roy et al. (2013)’s median value of 1.41. We measure a median total column density per LOS of $N(HI, \text{total})_{20} = 10.4$ compared with 5.54 by HT03 and 4.33 by Roy et al. (2013).

Figure 2.8 displays a “box and whisker” plots of f_{CNM} for 21-SPONGE, HT03 and Roy et al. (2013a). By inspection, our results are consistent with the f_{CNM} distribution found by HT03 and Roy et al. (2013a). HT03 do not detect HI absorption in the direction of 15/79 sources, resulting in many values of $f_{CNM} = 0$. Whereas we only have one non-detection, we find 9/31 sources with $f_{CNM} \leq 0.1$. We measure median $f_{CNM} = 0.20$ compared with 0.23 by HT03. We appear to find lower f_{CNM} than Roy et al. (2013b). Out of our 31 sources, we find a maximum $f_{CNM} \sim 0.51$, whereas Roy et al. (2013b) find much higher values, all the way up to $f_{CNM} \sim 1$ with median 0.51. As discussed above, their $\Sigma N(HI, CNM)_{20}$ are strong upper limits, relying on the assumption that all absorbing gas has $T_s = 200$ K, which neglects gas along the line of sight

at higher and lower temperatures. Therefore, their f_{CNM} values, calculated by dividing the upper limits to $\Sigma N(\text{H}\text{I}, \text{CNM})_{20}$ by the total isothermal column density (Chengalur et al. 2013), are all strong upper limits.

2.7 Comparison with synthetic observations

Recently, Kim et al. (2014) used a hydrodynamical simulation of a Milky Way-like Galactic disk to construct synthetic H I absorption and emission profiles for thousands of sight lines probing the turbulent, multiphase ISM. They calculate “observed” T_s , T_k , $N(\text{H}\text{I})$ and f_{CNM} for all lines of sight and found that the majority of sight lines with average LOS spin temperature below 200 K have a CNM fraction between 0.4 and 0.7 (98% have $f_{\text{CNM}} < 0.7$; C.-G. Kim et al. 2014 private communication). Whereas we measure several CNM fractions within this range, by inspection of Figure 2.8 and Table 2.3, we have more measurements below $f_{\text{CNM}} = 0.3$ than predicted by the Kim et al. (2014) results. However, in agreement with HT03, Kim et al. (2014) also found many WNM-dominated sightlines with $f_{\text{CNM}} = 0$.

Interestingly, Kim et al. (2014) found that synthetically-observed line-of-sight (harmonic mean) T_s values agree within a factor of 1.5 with “true” harmonic mean T_s values computed from the simulation over the full range of integrated optical depth ($10^{-3} < \int \tau dv < 10^2 \text{ km s}^{-1}$). They argue that this implies that there is little overlap between CNM clouds along a line of sight within their 1 km s^{-1} velocity channels. However, by inspection of our absorption profiles and fits (e.g. Figure 4), it is clear that cloud complexes can overlap significantly along every real observed line of sight.

In Figure 2.9 we plot CNM fraction versus the optical depth-weighted LOS spin temperature, $T_{s,\text{obs}}$. This weighted spin temperature, $T_{s,\text{obs}}$, is calculated from Equation (15) of Kim et al. (2014), which was derived from the “isothermal” column density estimate (Spitzer 1978, Dickey & Benson 1982, Chengalur et al. 2013, Kim et al. 2014), and is given by,

$$T_{s,\text{obs}} = \frac{\int \tau(v) \frac{T_B(v)}{1-e^{-\tau(v)}} dv}{\int \tau(v) dv}. \quad (2.17)$$

We display the error in the weighted T_s values when they exceed the symbol size. The squares represent same quantities with a cutoff for the CNM fraction at $T_s = 350$ K instead of $T_s = 200$ K. This increases the value of f_{CNM} in only a few cases, because we only detect 6% of all components in the $200 < T_s < 350$ K regime by number. This shows that our choice of temperature cutoff does not strongly affect the observed trends in f_{CNM} .

Kim et al. (2014) plotted the same quantities (their Figures 7a and 8a) and found that f_{CNM} , defined by a T_k cutoff at 184 K, varies with $T_{s,\text{obs}}$ roughly as $f_{\text{CNM}} \approx T_c/T_{s,\text{obs}}$ for $50 < T_c < 100$ K. These $T_c/T_{s,\text{obs}}$ curves are shown in Figure 2.9 in dashed purple for $T_c = 20, 50$ and 100 K. Although we see a similar tight trend between CNM fraction and $T_{s,\text{obs}}$, we find points below the $50/T_{s,\text{obs}}$ curve, indicating again that we are measuring lower f_{CNM} and motivating us to extend the dashed curves down to $20/T_s$. Kim et al. (2014) possibly produced higher f_{CNM} because they did not include HI-H₂ chemistry, and therefore could have overestimated f_{CNM} if some cold, dense HI should have transitioned to H₂ form. This was also noticed by Stanimirović et al. (2014) in an HI emission and absorption line study of the Perseus molecular cloud at the Arecibo Observatory.

Overall, we generally agree with the trend between CNM fraction and optical depth-weighted LOS T_s ($T_{s,\text{obs}}$) produced by the synthetic observations of the Kim et al. (2014) simulation, although our observations increase the scatter in the trend. The agreement is very encouraging considering that our CNM fractions rely on temperature calculation by Gaussian decomposition, whereas theirs are taken directly from their simulated lines of sight. Kim et al. (2014) derived an expression describing this trend based on several assumptions: $T_s = T_k$ in the CNM, the “isothermal” column density equals the true column density (truest for low integrated τ sightlines), and the synthetically observed harmonic mean T_s equals the true harmonic mean T_s (where the agreement is best for $\tau < 0.1$, above which there is scatter of up to 1.5). Finally, in computing f_{CNM} , they assume that the CNM fraction is not “extremely small” in order to express the f_{CNM} as a ratio of temperatures instead of column densities (see their Section 3.2, Equation (16), Kim et al. 2014). This last assumption may be the most important to increasing the trend

scatter, because we (and HT03) find many sightlines with both small f_{CNM} and small $T_{s,\text{obs}}$.

The differences between our results are likely due in part to the fact that Kim et al. (2014) do not include magnetic fields or chemistry to map the transitions between HI and H₂, but the similarities found so far suggest that only small adjustments to the simulations may be necessary to match the observed f_{CNM} . We also emphasize that the spatial resolution of a simulation is key for resolving the CNM and appropriately calculating quantities like the CNM fraction (e.g. Audit & Hennebelle 2005). In the future we will apply our Gaussian fitting methods to large volumes of synthetic data from Kim et al. (2014) and other higher spatial-resolution numerical simulations using Autonomous Gaussian Decomposition (Lindner et al. 2015). This will be essential to assessing the effect of our analysis techniques on these results and also understanding the biases of the numerical models.

2.8 Comparison of methods for estimating spin temperature from HI absorption and emission spectra

Several independent methods are commonly used in the literature to measure T_s from HI absorption and emission observations. The first is the method we use in this work, whereby we estimate a single T_s value for each Gaussian component fitted to both absorption and emission spectra (i.e., solving Equation 1). This method relies on initial guesses for number of components and their shapes, but it is the only method based on a physical model. As is generally true for the 21 cm line, a Gaussian model is a good representation of the spectral line because the damping wings are insignificant. Although this assumption fails in some cases, shown in this study by the components with $T_{k,\text{max}} \gg T_s$ (6% have $T_{k,\text{max}} > (T_s + 3\sigma)$), the method is convenient and produces reasonable fits to both simple and complex profiles.

2.8.1 One phase per channel T_s

Assuming that the gas in each velocity channel is at a single temperature (the “one-phase” assumption), dividing the brightness temperature spectrum ($T_B(v)$) by the absorption profile

$(1 - e^{-\tau(v)})$ produces a spin temperature spectrum, $T_{s,O}(v)$, via the following equation,

$$T_{s,O}(v) = \frac{T_B(v)}{1 - e^{-\tau(v)}}. \quad (2.18)$$

From this it is possible to measure a single T_s value for every velocity channel. For example, Roy et al. (2013a) use this method to compute $T_{s,O}$ for each velocity channel above 3σ significance in absorption at 1 km s^{-1} resolution.

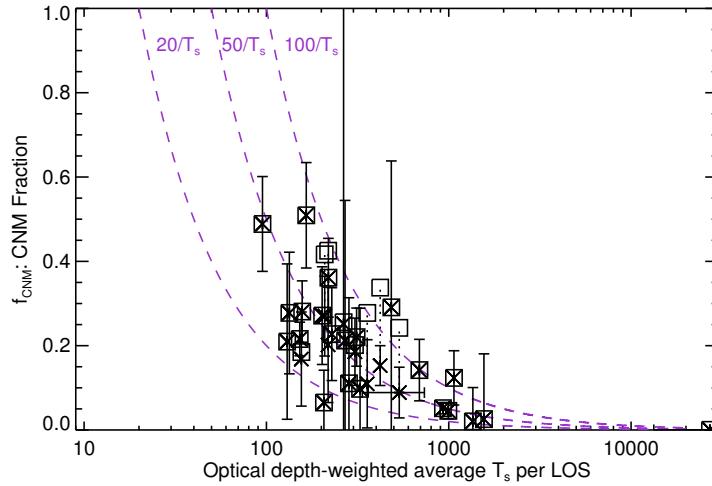


Figure 2.9 CNM fraction (f_{CNM}) versus optical depth weighted average T_s per LOS (Equation 2.17). Purple dashed lines indicate the observed trends in CNM fraction with LOS spin temperature from Kim et al. (2014) (their Figures 7a and 8a). Crosses indicate f_{CNM} calculated from Equation (2.16) with a cutoff of $T_s = 200 \text{ K}$ to define the CNM, squares indicate f_{CNM} using a cutoff of $T_s = 350 \text{ K}$. Errors in f_{CNM} are shown (solid vertical lines) when they exceed the symbol size. Dotted lines connect crosses and squares when the difference between the two values exceed the symbol size.

2.8.2 Line of sight harmonic mean T_s

Assuming that there is one value of T_s along a full line of sight, the “LOS” harmonic-mean temperature, $T_{s,L}$ from the absorption and emission profiles is given by,

$$T_{s,L} = \frac{\int T_B(v)dv}{\int(1 - e^{-\tau(v)})dv}. \quad (2.19)$$

For example, Kanekar et al. (2011) computed $T_{s,L}$ for the same data from Roy et al. (2013) to distinguish between CNM and WNM-dominated lines of sight.

2.8.3 The Slope Method

It is also possible to measure T_s for individual gas components by estimating the slope or “ridge line” of each linear feature in $T_B(v)$ vs. $1 - e^{-\tau(v)}$ space (e.g. Mebold et al. 1997, Dickey et al. 2000). The advantage of this approach compared with Gaussian decomposition is that the solution can be determined directly from the data without fitting iterations or initial guesses. However, this method also requires the assumption that all gas in a given velocity range exists at a single temperature, as is also assumed in the one phase per channel and LOS methods. To improve the slope method parameter estimation, Dickey et al. (2003) developed a two-phase nonlinear least squares fitting technique to determine the T_s slopes, which we will call the “NLLSQ slope method”.

Dickey et al. (2003) carried out a detailed comparison of T_s estimation methods, including Gaussian decomposition and the NLLSQ slope method, and found that in cases of isolated, unblended CNM components, these two methods agree reasonably. In another comparison study, HT03 computed T_s for a sample of their sources using a by-eye version of the slope method and similarly concluded that the two methods agree in the case of unblended CNM components, even if there are blended, weaker WNM components present. Dickey et al. (2003) re-processed all 79 HT03 emission and absorption pairs using the NLLSQ slope method and found a median spin temperature per components of $T_s = 31$ K, which is lower than the median value per fitted Gaussian functions by HT03 of $T_s = 48$ K. This suggests that there may be systematic differences

between the two methods.

In the comparisons made by Dickey et al. (2003) and HT03, the T_s values from the NLSSQ slope method tend to be lower than the Gaussian-derived T_s . This is due to the fact that spectral features in HI emission and absorption profiles do not always have the same shape. If channels in emission corresponding to narrow absorption have similar brightness temperature as channels with no corresponding absorption, the slope of the opacity feature will be small. HT03 note that this is because component emission is determined by the product of both optical depth and T_s , so that broad, warm components dominate emission and narrow, cold components dominate only in absorption. We will conduct a detailed comparison of the Gaussian and NLSSQ T_s derivation methods upon completion of the 21-SPONGE data collection. In addition, we plan to use synthetic spectra from numerical simulations, where the input T_s values are known, to assess and quantify their associated biases.

2.8.4 Applying methods to 21-SPONGE data

From a preliminary analysis of 5 21-SPONGE sources (3C286, 3C225A, 3C225B, 3C237, and 3C298), we find that the NLSSQ slope method temperatures are mostly in the range $25 < T_s < 100$ K, which agrees well with the peak of our T_s distribution (Figure 2.6a). In Figure 2.10 we compare the Gaussian fit, one-phase per channel, and LOS harmonic mean methods by applying them all to the 31 21-SPONGE emission/absorption spectral pairs. For the one-phase per channel method, for each source we used Equation (2.18) and included the T_s values for all channels with absorption signals above the 3σ level. In Figure 2.10a, we plot histograms of T_s for (1) the Gaussian decomposition method (thick green, “G”) (2) the one-phase per channel method (thin black, “O”) (3) the LOS harmonic mean method (shaded gray, “L”). In Figure 2.10b we display cumulative distribution functions for the three T_s methods.

Based on the CDFs of the three T_s distributions in Figure 2.10b, the one-phase per channel and LOS-integrated methods agree well ($K-S = 0.18, p = 0.35$). Both methods rely on the assumption that there is a single-phase medium in each measurement interval (either velocity

channel or full line of sight), which over-simplifies the reality that the ISM is a complex, multi-phase mixture of gas at many temperatures. The one-phase assumption overestimates T_s , because within each velocity channel (and line of sight) narrow components at lower temperatures can cumulatively produce a stronger absorption or emission signal, thereby biasing the average T_s measurement to high values. As previously discussed by Dickey et al. (2000) and Dickey et al. (2003), the NLLSQ slope and Gaussian methods provide systematically lower values of T_s than the single-phase approximations. This is supported by strong disagreements in Figure 2.10 ($p=0.00$ between both G and O, and G and L).

The median values of the Gaussian, one-phase per channel and LOS methods are $T_{s,G} = 77\text{ K}$, $T_{s,O} = 340\text{ K}$ and $T_{s,L} = 290\text{ K}$. Wolfire et al. (2003) model the multi-phase ISM with

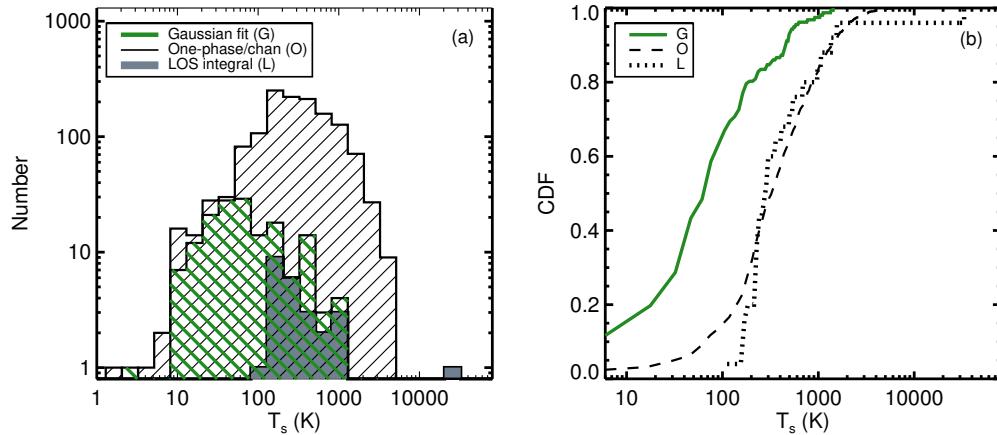


Figure 2.10 Comparison of temperature derivation methods for the 21-SPONGE survey. (a): Histograms of T_s for: Gaussian decomposition of HI absorption and emission lines (thick green), “one-phase” T_s per 0.42 km/s channel (thin black, Equation 2.18) and integrated T_s per LOS (shaded gray, Equation 2.19). (b): Cumulative distribution functions comparing the three temperature distributions from panel (a).

detailed analytical prescriptions for heating and cooling, and show that the CNM should have a spin temperature of $40 \leq T_s \leq 200$ K. The Gaussian fit method best reproduces this predicted range, although the T_s values in Figure 2.10 also contain components associated with thermally unstable gas in addition to the dominant CNM. Furthermore, Figure 2.10 shows that the Gaussian fit method is the most successful in reproducing the analytical prediction for the lower limit to CNM T_s . The LOS method in particular misses a large fraction of the CNM T_s distribution, as it predicts all gas should have $T_s > 200$ K. As discussed by Dickey et al. (2009), the T_s single-phase (i.e. O and L) T_s methods provide estimates of the mean temperature or the mixture of CNM and WNM along the line of sight, rather than physical temperatures of individual structures.

2.9 Summary and Conclusions

In this paper we present methods, data and preliminary results from the first half of 21-SPONGE, a large HI absorption survey at the Karl G. Jansky VLA. The absorption data are complemented by emission data from the Arecibo Observatory, whose $\sim 3.5'$ beam at 21-cm provides the best available single-dish complement to the $\sim 1''$ VLA beam so that we come closer by orders of magnitude than previous studies to sampling the same gas populations in absorption and emission. By obtaining extremely high-sensitivity VLA HI absorption lines, we detect signatures of both cold and warm ISM directly in absorption and, in combination with emission, directly determine the temperatures and column densities of individual components along the line of sight. We describe the observation and data reduction strategies and discuss details of line fitting and parameter estimation in this paper. Important initial results from the first half of the survey include:

1. We achieve median RMS noise in optical depth of $\sigma_\tau \sim 9 \times 10^{-4}$ per 0.42 km s^{-1} channel over the 31 sightlines probed by 21-SPONGE so far. This allows us to directly probe absorption signals from both cold and warm HI. We are able to achieve this sensitivity in part by maximizing the signal-to-noise in our bandpass calibration solutions, which

we accomplish by characterizing periodic structures within the solutions and combining calibration observations over time.

2. Following a careful Gaussian decomposition of all 31 HI absorption detections, we estimate T_s and N(HI) for 157 Gaussian components. We find a maximum spin temperature per individual component of $T_s = 1451 \pm 263$ K, which is more than a factor of 2 higher than the maximum Gaussian component spin temperature measured by the Millennium Arecibo HI Survey (Heiles & Troland 2003; HT03). We are sensitive to lower N(HI, CNM) due to our better sensitivity in optical depth, and detect individual clouds with N(HI, CNM) as low as $3 \pm 1 \times 10^{16}$ cm $^{-2}$.
3. We compute total CNM column density, total WNM column density, total HI column density and CNM fraction along each of the 31 emission/absorption pair sightlines. We detect similar total CNM column densities, defined as the total column density of all Gaussian components per sightline with $T_s \leq 200$ K, as HT03 and Roy et al. (2013). We find larger values of total “warm gas” column density, defined as the total column density of all absorption-detected Gaussian components with $T_s > 200$ K and the components only detected in emission, due to the fact that we are sensitive to warmer gas in absorption than HT03. We find consistent CNM fractions (f_{CNM}) with those found by HT03 and Roy et al. (2013), although the Roy et al. (2013) values are all upper limits.
4. We observe the same trend between CNM fraction and harmonic mean LOS spin temperature found by the synthetic-observation analysis of a hydrodynamic Galaxy simulation by Kim et al. (2014). Although we find lower CNM fractions (median 0.20) than the simulation predicts (98% LOS with $0.4 < f_{\text{CNM}} < 0.7$), the agreement in the trend is encouraging given that our calculations rely on Gaussian decomposition to estimate CNM fraction and theirs are taken directly from the simulation, suggesting that the Gaussian decomposition method is successful.

5. In a comparison of T_s estimation methods, we find that one-phase or isothermal assumptions overestimate T_s values relative to individual Gaussian fits. The T_s distribution produced by multi-phase fits via Gaussian decomposition better reproduces theoretically expected lower limits for individual CNM clouds based on detailed heating and cooling analytical models.

This work was supported by the NSF Early Career Development (CAREER) Award AST-1056780. C. M. acknowledges support by the National Science Foundation Graduate Research Fellowship and the Wisconsin Space Grant Institution. S.S. thanks the Research Corporation for Science Advancement for their support. We thank Nirupam Roy and Crystal Brogan for sharing absorption spectra of the source 3C138 for comparison. We thank Eve Ostriker and Chang-Goo Kim for helpful discussions, and Ayesha Begum and Ben Klug for help with initial preparation and design for the survey. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. The Arecibo Observatory is operated by SRI International under a cooperative agreement with the National Science Foundation (AST-1100968), and in alliance with Ana G. Méndez-Universidad Metropolitana, and the Universities Space Research Association.

References

- Audit, E., & Hennebelle, P. 2005, *A&A*, 433, 1
- Begum, A., Stanimirović, S., Goss, W. M., et al. 2010, *ApJ*, 725, 1779
- Braun, R., & Walterbos, R. A. M. 1992, *ApJ*, 386, 120
- Brogan, C. L., Zauderer, B. A., Lazio, T. J., et al. 2005, *AJ*, 130, 698
- Bryan, G. 2007, in EAS Publications Series, Vol. 24, EAS Publications Series, ed. E. Emsellem, H. Wozniak, G. Massacrier, J.-F. Gonzalez, J. Devriendt, & N. Champavert, 77–88
- Carilli, C. L., Dwarakanath, K. S., & Goss, W. M. 1998, *ApJ*, 502, L79
- Chengalur, J. N., Kanekar, N., & Roy, N. 2013, *MNRAS*, 432, 3074

- Christensen, C., Quinn, T., Governato, F., et al. 2012, MNRAS, 425, 3058
- Condon, J., Cotton, W., Greisen, E., et al. 1998, AJ, 115, 1693
- Cox, D. P. 2005, ARA&A, 43, 337
- Crovisier, J., Kazes, I., & Aubry, D. 1978, A&AS, 32, 205
- Davies, R. D., & Cummings, E. R. 1975, MNRAS, 170, 95
- Dawson, J. R., McClure-Griffiths, N. M., Kawamura, A., et al. 2011, ApJ, 728, 127
- Deguchi, S., & Watson, W. D. 1985, ApJ, 290, 578
- Dickey, J. M., Salpeter, E. E., & Terzian, Y. 1977, ApJ, 211, L77
- Dickey, J., Kulkarni, S., van Gorkom, J., & Heiles, C. 1983, ApJS, 53, 591
- Dickey, J. M., McClure-Griffiths, N. M., Gaensler, B. M., & Green, A. J. 2003, ApJ, 585, 801
- Dickey, J. M., Strasser, S., Gaensler, B. M., et al. 2009, ApJ, 693, 1250
- Dickey, J. M., McClure-Griffiths, N., Gibson, S. J., et al. 2013, PASA, 30, 3
- Dwarakanath, K. S., Carilli, C. L., & Goss, W. M. 2002, ApJ, 567, 940
- Faison, M. D., Goss, W. M., Diamond, P. J., & Taylor, G. B. 1998, AJ, 116, 2916
- Field, G. B. 1958, Proceedings of the IRE, 46, 240
- Field, G. B., Goldsmith, D. W., & Habing, H. J. 1969, ApJ, 155, L149
- Governato, F., Brook, C., Mayer, L., et al. 2010, Nature, 463, 203
- Haslam, C. G. T., Salter, C. J., Stoffel, H., & Wilson, W. E. 1982, A&AS, 47, 1
- Heiles, C., & Troland, T. 2003a, ApJS, 145, 329
- Heiles, C., & Troland, T. H. 2003b, ApJ, 586, 1067
- Hill, A. S., Joung, M. R., Mac Low, M.-M., et al. 2012, ApJ, 750, 104
- Horne, J. H., & Baliunas, S. L. 1986, ApJ, 302, 757
- Kalberla, P., Burton, W., Hartmann, D., et al. 2005, A&A, 440, 775

- Kanekar, N., Subrahmanyan, R., Chengalur, J. N., & Safouris, V. 2003, MNRAS, 346, L57
- Kanekar, N., Braun, R., & Roy, N. 2011, ApJ, 737, L33
- Kim, C.-G., Ostriker, E. C., & Kim, W.-T. 2014, ApJ, 786, 64
- Koyama, H., & Inutsuka, S.-i. 2002, ApJ, 564, L97
- Lazareff, B. 1975, A&A, 42, 25
- Lindner, R. R., Vera-Ciro, C., Murray, C. E., et al. 2015, arXiv:1409.2840
- Liszt, H. 2001, A&A, 371, 698
- Mac Low, M.-M., Balsara, D. S., Kim, J., & de Avillez, M. A. 2005, ApJ, 626, 864
- McClure-Griffiths, N. M., Dickey, J. M., Gaensler, B. M., et al. 2005, ApJS, 158, 178
- McClure-Griffiths, N. M., Stanimirovic, S., Murray, C. E., et al. 2015, arXiv:1501.01130
- McKee, C., & Ostriker, J. 1977, ApJ, 218, 148
- Mebold, U., Winnberg, A., Kalberla, P. M. W., & Goss, W. M. 1982, A&A, 115, 223
- Mebold, U., Düsterberg, C., Dickey, J. M., Staveley-Smith, L., & Kalberla, P. 1997, ApJ, 490, L65
- Mohan, R., Dwarakanath, K., & Srinivasan, G. 2004, Journal of Astrophysics and Astronomy, 25, 143
- Ntormousi, E., Burkert, A., Fierlinger, K., & Heitsch, F. 2011, ApJ, 731, 13
- Ostriker, E. C., McKee, C. F., & Leroy, A. K. 2010, ApJ, 721, 975
- Payne, H., Dickey, J., Salpeter, E., & Terzian, Y. 1978, ApJ, 221, L95
- Peek, J., Heiles, C., Douglas, K., et al. 2011, ApJS, 194, 20
- Pikel'Ner, S. B. 1968, ARA&A, 6, 165
- Rohlfs, K., & Wilson, T. L. 2004, Tools of radio astronomy, 4th rev. and enl. ed., by K. Rohlfs and T.L. Wilson. Berlin: Springer, 2004, 107

- Roy, N., Chengalur, J. N., & Srianand, R. 2006, MNRAS, 365, L1
- Roy, N., Minter, A. H., Goss, W. M., Brogan, C. L., & Lazio, T. J. W. 2012, ApJ, 749, 144
- Roy, N., Kanekar, N., Braun, R., & Chengalur, J. N. 2013a, MNRAS, 436, 2352
- Roy, N., Kanekar, N., & Chengalur, J. N. 2013b, MNRAS, 436, 2366
- Stanimirović, S., & Heiles, C. 2005, ApJ, 631, 371
- Stanimirović, S., Putman, M., Heiles, C., et al. 2006, ApJ, 653, 1210
- Stil, J. M., Taylor, A. R., Dickey, J. M., et al. 2006, AJ, 132, 1158
- Taylor, A. R., Gibson, S. J., Peracaula, M., et al. 2003, AJ, 125, 3145
- Westmoquette, M. S., Exter, K. M., Smith, L. J., & Gallagher, J. S. 2007, MNRAS, 381, 894
- Wolfire, M., McKee, C., Hollenbach, D., & Tielens, A. 2003, ApJ, 587, 278

Chapter 3

Excitation temperature of the Warm Neutral Medium as a new probe of the Ly α radiation field

*A version of this chapter has previously appeared
in the Astrophysical Journal*

C. E. Murray, S. Stanimirović, R. R. Lindner, et al. 2014, ApJ, 781, L41

Abstract

We use the Karl G. Jansky Very Large Array (VLA) to conduct a high-sensitivity survey of neutral hydrogen (HI) absorption in the Milky Way. In combination with corresponding HI emission spectra obtained mostly with the Arecibo Observatory, we detect a widespread warm neutral medium (WNM) component with excitation temperature $\langle T_s \rangle = 7200_{-1200}^{+1800}$ K (68% confidence). This temperature lies above theoretical predictions based on collisional excitation alone, implying that Ly- α scattering, the most probable additional source of excitation, is more important in the interstellar medium (ISM) than previously assumed. Our results demonstrate that HI absorption can be used to constrain the Ly- α radiation field, a critical quantity for studying the energy balance in the ISM and intergalactic medium yet notoriously difficult to model because of its complicated radiative transfer, in and around galaxies nearby and at high redshift.

3.1 Introduction

Understanding physical conditions within the diffuse neutral interstellar medium (ISM) is essential for producing realistic models of star and galaxy formation. Ambient gas temperature and density are crucial input parameters for heating, cooling and feedback recipes on all astronomical scales. While numerical simulations are becoming increasingly complex, details regarding neutral gas temperature distributions, shielding properties, essential feedback sources, and excitation processes are still very much under debate (Bryan 2007; Christensen et al. 2012). The excitation processes of the 21-cm line are especially important for interpreting radio signals from early epochs of cosmic structure formation (e.g., the cosmic dark ages and subsequent epoch of reionization), when neutral hydrogen dominated the baryonic content of the Universe and facilitated the formation of the first stars and galaxies (Pritchard & Loeb 2012). To interpret 21-cm signals from the early universe, it is necessary to decouple astrophysical effects from cosmological effects which is likely best done by analyzing excitation processes in the local ISM.

Traditional ISM models contain two neutral phases, the cold neutral medium (CNM) and the warm neutral medium (WNM), individually in thermal and pressure equilibrium (Field et al. 1969; McKee & Ostriker 1977; Wolfire et al. 2003). Widely-accepted theoretical properties of these phases in the Milky Way include: a kinetic temperature of $T_k \sim 40\text{--}200\,\text{K}$ and a volume density of $n(\text{H}\text{I}) \sim 5\text{--}120\,\text{cm}^{-3}$ for the CNM, and $T_k \sim 4100\text{--}8800\,\text{K}$ and $n(\text{H}\text{I}) \sim 0.03\text{--}1.3\,\text{cm}^{-3}$ for the WNM (Wolfire et al. 2003).

A convenient tracer for neutral gas is the H I 21-cm line, originating from the hyperfine energy splitting caused by magnetic moment interactions between the hydrogen atom's electron and proton. The high optical depth of the CNM makes 21-cm absorption signatures easy to detect, even with low sensitivity observations (e.g., Lazareff 1975; Dickey et al. 1977; Crovisier et al. 1978; Payne et al. 1978; Dickey et al. 1983; Braun & Walterbos 1992; Heiles & Troland 2003; Kanekar et al. 2003; Mohan et al. 2004; Roy et al. 2006; Begum et al. 2010). The excitation

temperature (or spin temperature, T_s) of the CNM can be directly estimated by solving radiative transfer equations.

In contrast, the WNM is characterized by very-low peak optical depth and so measuring its spin temperature from absorption requires extremely high sensitivity, $\sigma_\tau \leq 10^{-3}$ and attention to systematic errors. Only two direct measurements of WNM spin temperature exist so far (Carilli et al. 1998; Dwarakanath et al. 2002). However, Carilli et al. (1998) observed absorption in the direction of Cygnus A, an exceptionally bright radio continuum source with a flux density of ~ 400 Jy, and achieved excellent sensitivity with orders of magnitude less integration time than is required for the majority of (> 3 Jy) radio continuum sources. Less accurate are upper limits on T_k (and thus T_s since $T_s \leq T_k$) estimated by spectral linewidths (Mebold et al. 1982; Kanekar et al. 2003; Heiles & Troland 2003; Roy et al. 2013b), and approximate T_s estimates obtained by assigning a single temperature to strongly-absorbing, complex HI profiles (Kanekar et al. 2011; Roy et al. 2013a).

Relating measured HI spin temperatures to model-predicted kinetic temperatures is further complicated by the uncertainty in the excitation mechanisms involved. The 21-cm transition in the high-density CNM is expected to be thermalized by collisions with electrons, ions, and other HI atoms, resulting in $T_s \sim T_k$. In contrast, low densities in the WNM imply that collisions cannot thermalize the 21-cm transition and therefore $T_s < T_k$ (Field 1958; Deguchi & Watson 1985; Liszt 2001). However, the Ly- α radiation field from Galactic and extragalactic sources can serve to thermalize the transition. This requires a very large optical depth and a large number of scatterings of Ly- α photons to bring the radiation field and the gas into local thermal equilibrium. While the underlying atomic physics is understood (Wouthuysen 1952; Field 1958; Pritchard & Loeb 2012), the details of Ly- α radiative transfer are complicated and depend on the topology and the strength of the Ly- α radiation field, which are complex and poorly constrained in the multi-phase ISM (Liszt 2001).

To measure the physical properties of the WNM and the coupling between neutral gas and

$\text{Ly-}\alpha$ radiation, we are conducting a large survey, 21-cm Spectral Line Observations of Neutral Gas with the EVLA (21 SPONGE), to obtain high-sensitivity Milky Way HI absorption spectra using the Karl G. Jansky Very Large Array (VLA). The recently upgraded capabilities of the VLA allow us to routinely achieve RMS noise levels in optical depth of $\sigma_\tau \sim 7 \times 10^{-4}$ per 0.42 km s^{-1} channel, which are among the most sensitive observations of HI absorption to date. Currently, 24 out of 58 sightlines are complete after over 200 hours of observing time. This paper summarizes our initial results from this project and the detection of the WNM using a newly developed analysis technique based on the spectral stacking of HI absorption and emission spectra. In Section 3.2, we summarize our observing and data processing strategies, results of the stacking analysis are provided in Section 3.3 and discussed in Section 3.4, and we present our conclusions in Section 3.5.

3.2 Observations and Data Processing

3.2.1 Observations

Each source, selected from the NRAO/VLA Sky Survey (Condon et al. 1998), was chosen to have a 1.4 GHz flux density $\geq 3 \text{ Jy}$ to avoid excessively long integration times with the VLA to reach the desired sensitivity. In addition, we select sources generally at high Galactic latitude ($|b| > 10^\circ$) to avoid complicated CNM profiles associated with the Galactic plane, and with angular sizes less than $1'$ to avoid resolving substantial flux density. All VLA observations use three separate, standard L-band configurations, each with one dual-polarization intermediate frequency band of width 500 kHz with 256 channels, allowing for a velocity coverage of 107.5 km s^{-1} and resolution of 0.42 km s^{-1} . We perform bandpass calibration via frequency switching, and all data were reduced using the Astronomical Image Processing System¹ (AIPS). The absorption spectra, $\tau(v)$, were extracted from the final cleaned data cubes following calibration detailed in the 21 SPONGE pilot paper (Begum et al. 2010).

In addition, for each sightline we obtain HI emission profiles which estimate the brightness

¹<http://www.aips.nrao.edu>

temperature ($T_B(v)$) in the absence of the radio continuum source. Of our 24 sightlines, 11 have emission profiles from the Millennium Arecibo 21-cm Absorption Line Survey (Heiles & Troland 2003), 10 have emission profiles from the Galactic Arecibo L-band Feed Array Survey in HI (GALFA-HI; Stanimirović et al. 2006; Peek et al. 2011), and for 3 sightlines which were not included in the Millennium survey or the GALFA-HI survey to date, we use emission spectra from the Leiden Argentine Bonn (LAB; Kalberla et al. 2005) survey. Arecibo HI emission spectra have not been corrected for contamination entering the telescope beam through distant sidelobes (so called stray radiation).

3.2.2 Derivation of T_s for individual components

To estimate T_s for individual spectral components, we follow the method of Heiles & Troland (2003). We first fit each absorption profile with Gaussian functions assuming thermal and turbulent broadening. We apply the statistical f -test to determine the best-fit number of components. We then fit the $T_B(v)$ profile simultaneously for additional Gaussian components and T_s for each absorption-detected component by solving the radiative transfer equations discussed at length in Heiles & Troland (2003). These results will be presented along with additional survey information in Murray et al. 2014 (in prep).

We next construct “residual” spectra to search for weak absorption features below our observational sensitivity by removing fitted models from the absorption and emission profiles. The best-fitting Gaussian components in absorption ($\tau_i(v)$ where i denotes the i^{th} component) are subtracted from each original absorption profile $\tau(v)$ to produce a residual absorption spectrum, $\tau_{\text{res}}(v) = \tau(v) - \sum \tau_i(v)$. The corresponding emission contribution from these components, $\Sigma T_{s,i}(1 - e^{-\tau_i(v)})$, is removed from each emission profile, to produce a residual emission spectrum $T_{B,\text{res}}(v) = T_B(v) - \Sigma T_{s,i}(1 - e^{-\tau_i(v)})$. Each residual emission spectrum is dominated by signals from the WNM purely detected in emission. Correspondingly, each residual absorption spectrum, $\tau_{\text{res}}(v)$, contains noise, model imperfections and weak absorption components below our detection threshold.

Table 3.1. Source Information

Name	σ_{τ}^a ($\times 10^3$)	$W = 1/\sigma_{\tau}^b$ ($\times 10^2$)	l (deg)	b (deg)
4C32.44	1.5	2.7	67.240	81.049
4C25.43	0.9	4.5	22.464	80.991
3C286	0.7	5.8	56.527	80.676
4C12.50	1.3	3.1	347.220	70.173
3C273	0.6	6.7	289.945	64.358
3C298	0.8	5.1	352.159	60.667
3C225A	1.5	2.7	219.866	44.025
3C225B	3	1.3	220.010	44.007
3C345	1.3	3.1	63.455	40.948
3C327.1	1.4	2.9	12.181	37.006
3C147*	0.4	10.1	161.686	10.298
3C154*	0.9	4.5	185.594	4.006
3C410	1.9	2.1	69.210	-3.768
B2050+36*	2.3	1.8	78.858	-5.124
P0531+19	0.5	8.1	186.760	-7.110
3C111*	0.7	5.8	161.675	-8.821
3C133	1.8	2.2	177.725	-9.914
3C138	0.9	4.5	187.403	-11.347
3C123*	0.7	5.8	170.581	-11.662
3C120	1.1	3.7	190.373	-27.397
3C48	0.7	5.8	133.961	-28.720
4C16.09	0.8	5.1	166.633	-33.598
3C454.3	0.9	4.5	86.108	-38.182
3C78	4.1	1.0	174.857	-44.514

Note. — *: Excluded sources (see Section 3.2.2).

^a: RMS noise in absorption profile calculated per 0.42 km s^{-1} channel. ^b: Normalized weighting factor (see Section 3); $\Sigma_n W_n = 1$.

To minimize the effects of model imperfections in our results, we exclude sources from our analysis which have strong model-subtraction errors using the following technique. We first calculate the expected noise profile of the absorption spectrum, $\sigma_\tau(v)$, for each source. Because the LAB survey uses antennas of comparable size to those of the VLA, we follow methods described by Roy et al. (2013a) to estimate $\sigma_\tau(v)$ using LAB HI emission. We next produce probability distribution functions (PDFs) of $\tau_{\text{res}}(v)/\sigma_\tau(v)$ and exclude from further analysis 5/24 sources whose PDFs deviate from a normal distribution at $\geq 97.5\%$ confidence. The excluded sources are significantly contaminated by model-subtraction artifacts, and lie at low Galactic latitude where lines of sight probe many velocity-blended HI clouds, which complicates Gaussian modeling. Table 1 lists the source name, σ_τ calculated per 0.42 km s^{-1} velocity channel in offline channels, a weighting factor described in Section 3, and Galactic coordinates.

3.3 Stacking analysis of HI absorption and emission spectra

We have performed a spectral “stacking” analysis on our 19 remaining residual spectra to search for extremely weak absorption signals from the diffuse WNM. We first apply a velocity shift to both the residual emission and absorption profiles to remove the effect of Galactic rotation and align any remaining signals at 0 km s^{-1} . The velocity shifts are computed using the first velocity moment of the residual emission spectrum for each source, so that $\Delta v = - \int T_{\text{B},\text{res}} v dv / \int T_{\text{B},\text{res}} dv$. To maximize the signal to noise ratio of the stacked absorption spectrum, the weight, W , for each profile is given by $W = \tau_{\text{res}}/\sigma_\tau^2$ (Treister et al. 2011). However, as discussed in the previous section, $\sigma_\tau \propto T_{\text{B}}$. For constant T_s , we have $\tau_{\text{res}} \propto T_{\text{B}}$, and so the weight simplifies to: $W = 1/\sigma_\tau$. We measure σ_τ in the offline channels of each profile, and list these values with the weights W , normalized by the sum over all $n = 19$ profiles, $\Sigma_n W_n$, in Table 1. The same weighting values were applied to the residual emission spectra.

The weighted profiles are averaged to produce the final stacked emission and stacked absorption spectra shown in Figure 3.1. The RMS noise in the stacked absorption spectrum is $\sigma_\tau = 2.6 \times 10^{-4}$, calculated over a fixed range of channels (20 to 30 and -30 to -20 km s^{-1}). This

is 4.2 times more sensitive than the median RMS noise calculated in the same velocity channels in the individual residual absorption profiles ($\sigma_\tau = 1.15 \times 10^{-3}$).

The enhanced sensitivity enabled by stacking allows us to detect a weak and broad absorption component which has a velocity width and centroid consistent with the stacked emission signal (Figure 3.1). This gives confidence that the stacked profiles trace physically related quantities. This broad absorption component is the weighted-mean, 0 km s^{-1} -centered

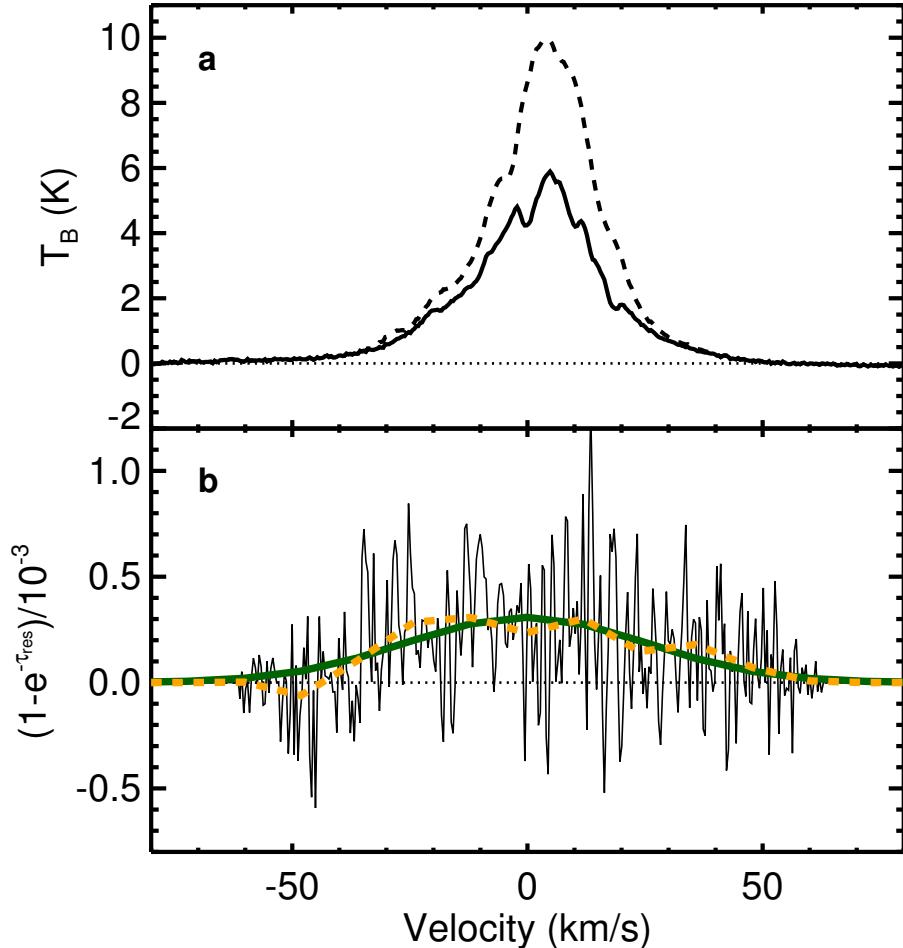


Figure 3.1 (a) The result of stacking the residual emission profiles $T_{B,\text{res}}(v)$ (solid line) for the 19 selected sources, and the result of stacking the total emission profiles, $T_B(v)$ (dashed line) for reference. (b) The result of stacking the residual VLA absorption profiles $\tau_{\text{res}}(v)$ (thin solid black) from the selected 19 sources. A smoothed version of the stack (10 km s^{-1} boxcar kernel) is overlaid (dashed orange), with a simple Gaussian fit to the profile overlaid (thick solid green). Due to the fact that all shifted profiles do not cover the same velocity range, we add zeroes to the edges of each profile so that they cover the same range. This results in fewer non-zero channels in the farthest velocity bins, which causes the noise level to be lower there.

absorption signal over all sightlines, and has an equivalent width (EW) within the velocity range common to all shifted profiles (-46 to 30 km s^{-1}) of $\text{EW} = \int_{-46}^{30} (1 - e^{-\tau_{\text{res}}}) dv = 0.018\text{ km s}^{-1}$.

We next conduct a bootstrapping (see, e.g., Wall & Jenkins 2003) Monte Carlo simulation to estimate the integrated strength of the stacked absorption signal (Figure 3.1b) and test for contamination by outlier spectra. We run the stacking analysis on a new sample of 19 sightlines randomly chosen from our original 19 sightlines with replacement. We repeat this trial 10^5 times, each time recomputing the EW. The resultant normalized PDF of EWs (Figure 3.2a, solid histogram) is nearly Gaussian, showing that the stacked EW signal is consistent with being drawn from a parent population of spectra with comparable means, and not due to a few outlier spectra. The peak of the EW distribution and the numerically-integrated 68% confidence limits are $\langle \text{EW} \rangle = 0.0182_{-0.0036}^{+0.0044}\text{ km s}^{-1}$, giving 1 in 2×10^6 chance of being spurious (5σ). Figure 3.2a also displays the result of repeating the bootstrapping simulation while inverting (or multiplying

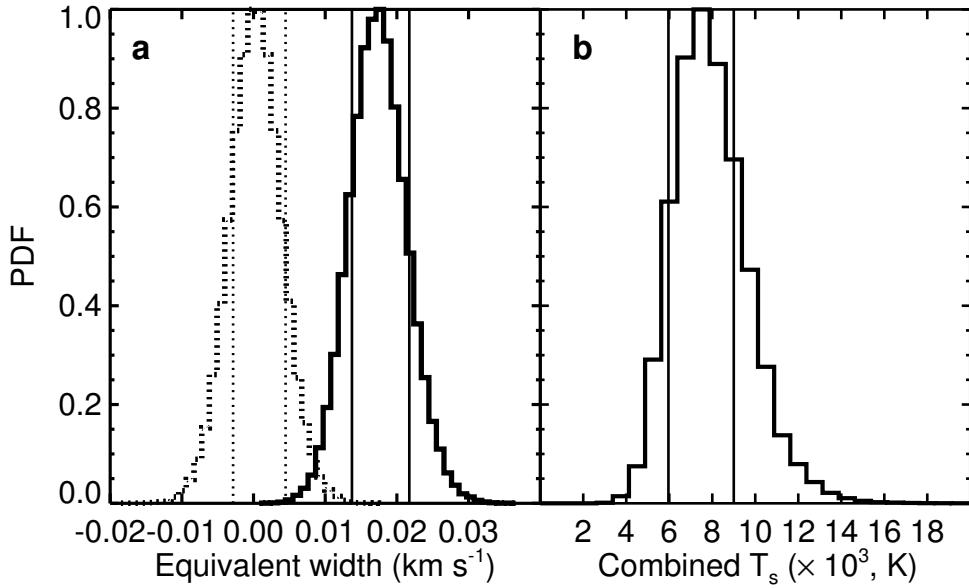


Figure 3.2 (a) PDF of stacked absorption EW following 10^5 iterations using a bootstrapping algorithm to select 19 sources from the total list of 19 with replacement, yielding $\langle \text{EW} \rangle = 0.0182_{-0.0036}^{+0.0044}\text{ km s}^{-1}$ (68% confidence). To test that the signal is real, we invert (or multiply by -1) a random selection of half of the absorption residual profiles in 10^5 additional trials, yielding a distribution (dotted histogram) consistent with zero signal, as expected. Vertical lines denote 1σ uncertainty limits. (b) Distribution of spin temperatures from combining the PDFs of three different estimations (see text) computed in all 10^5 bootstrapping trials. The derived spin temperature is $\langle T_s \rangle = 7200_{-1200}^{+1800}\text{ K}$ (68% confidence), with a 99% confidence lower limit of 4300 K .

by -1) a random selection of half of the profiles (dotted histogram). The result is fully consistent with zero signal, verifying that our stacking method does not produce spurious detections.

Using the Monte Carlo bootstrapping method, we also constrain the FWHM and peak optical depth of the stacked absorption feature by Gaussian fit as $\text{FWHM} = 50_{-7}^{+15} \text{ km s}^{-1}$ and $\tau_0 = 3.0_{-0.4}^{+1.0} \times 10^{-4}$ (at the central velocity of the feature, v_0), respectively. Both quantities have well-defined, single-peaked distributions, suggesting that the detection is tracing a single gas component rather than a blend of many narrow components. We therefore proceed to estimate the typical spin temperature T_s of the gas detected in the stacked absorption spectrum.

We use three methods: (1) $T_s = \int_{-46}^{30} T_{\text{B,res}} dv / \int_{-46}^{30} (1 - e^{-\tau_{\text{res}}}) dv$, (2) $T_s = T_{\text{B,res}}(v_o) / (1 - e^{-\tau_{\text{res}}})(v_o)$, where v_o is the velocity at the peak of the smoothed stacked absorption signal, and (3) by the methods of Heiles & Troland (2003), fitting a single Gaussian component to the absorption stack and solving for T_s by fitting this component to the emission stack without additional components. For each of the 10^5 bootstrapping trials, we compute all three T_s estimates and combine their distributions to minimize systematic errors associated with any individual approach (Figure 3.2b).

We estimate a spin temperature of the detected absorption feature of $\langle T_s \rangle = 7200_{-1200}^{+1800} \text{ K}$ (68% confidence), with a lower limit at 99% confidence of 4300 K. We note that by shifting the residual profiles by the second velocity moment or by the location of maximum $T_{\text{B,res}}$, we find consistent temperature estimates with similar significance. The possible contamination of stray radiation to the Arecibo HI emission spectra has no effect on the presence of the absorption stack, although it will tend to increase the EW of the emission stack, and therefore increase the estimated T_s of the stacked feature. For the GALFA-HI survey, Peek et al. (2011) estimated that the level of stray radiation contamination is $\leq 200\text{--}500 \text{ mK}$, leading to an overestimate in our computed T_s of at most 13%. Therefore, the presence of stray radiation in our Arecibo HI emission spectra would not change our results

3.4 Discussion

In Figure 3.3, we plot all previous WNM T_s detections (open symbols) against the RMS noise in off-line channels of the individual absorption spectra in which they were detected, with 1σ error bars derived from line fitting. We exclude literature measurements which are estimated as upper limits from T_k or which assume single temperatures for multi-component sightlines. Our result from the stacking analysis (green filled star) is in agreement with measurements by Carilli et al. (1998) (open triangles) obtained in the direction of ~ 400 Jy-bright Cygnus A. We emphasize

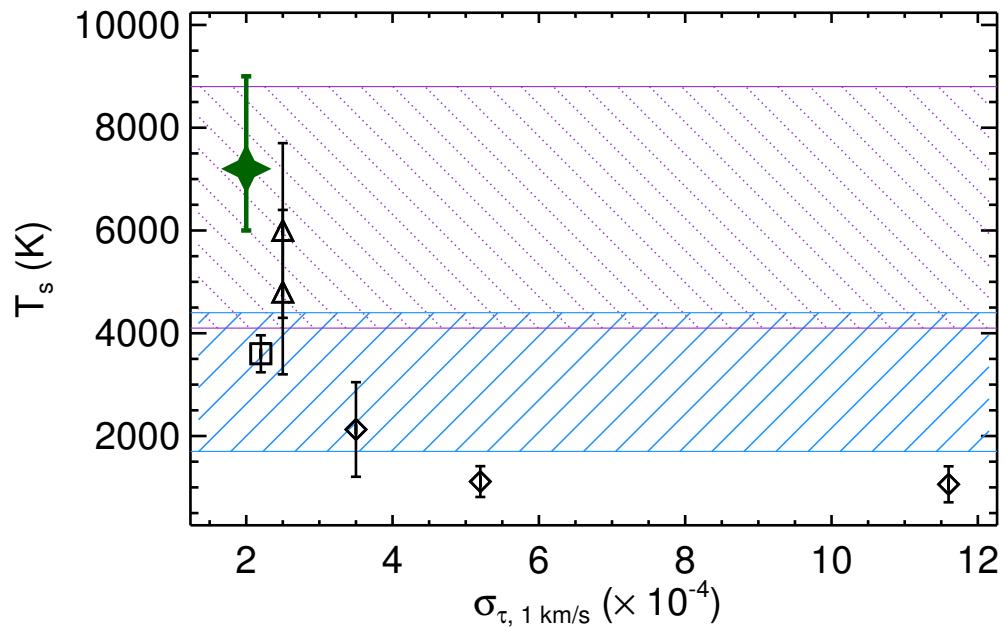


Figure 3.3 Comparison of previous T_s measurements for the WNM. Hollow symbols denote direct observational measurements of individual sightline components. Triangles: Absorption towards Cygnus A (Carilli et al. 1998), Square: Absorption towards 3C147 (Dwarakanath et al. 2002) Diamonds: 21 SPONGE absorption towards PKS0531+19, 3C298 and 3C133. All points are plotted versus the RMS noise in off-line channels in the absorption profile, per 1km s^{-1} channel ($\sigma_{\tau, 1 \text{ km/s}}$). The purple dotted hatched region denotes the kinetic temperature range from Wolfire et al. (2003) ($T_k \sim 4100\text{--}8800$ K), and the blue hatched region denotes the spin temperature range from Liszt (2001) for all possible ISM pressures ($T_s \sim 1800\text{--}4400$ K). The green filled star is the derived spin temperature from Figure 3.2b, $\langle T_s \rangle = 7200^{+1800}_{-1200}$ K (68% confidence), with a 99% confidence lower limit of 4300 K.

that our result samples a widespread population detected from 19 sightlines, while Carilli et al. (1998) examine only a single direction. Our T_s estimate is significantly higher than all other direct measurements shown here. The trend of increasing T_s with increasing observational sensitivity confirms the expectation that previous experiments with lower sensitivity ($\sigma_\tau \geq 5 \times 10^{-4}$) were unable to detect WNM with $T_s > 1000$ K due to its low optical depth.

The range of predicted kinetic temperatures from the most detailed ISM heating and cooling considerations, shown by the dotted purple hatched region in Figure 3.3, is $T_k \sim 4100\text{--}8800$ K (Wolfire et al. 2003). We use models considering only collisional excitation from Liszt (2001) (their Figure 2) to show that under all plausible ISM pressures, the range of spin temperatures implied by this T_k range from Wolfire et al. (2003), is $T_s \sim 1800\text{--}4400$ K (solid blue hatched region in Figure 3.3). Our measurement indicates that the mean spin temperature of the WNM is higher than expected theoretically for collisionally excited HI at 98% confidence.

Resonant scattering of Ly- α photons can contribute enough to 21-cm excitation to allow $T_s = T_k$, as long as a sufficient fraction of the Ly- α radiation permeates the WNM (Liszt 2001; Pritchard & Loeb 2012). However, the degree to which this can occur in a multi-phase ISM depends on several observationally unconstrained quantities: the local and external Ly- α field, interstellar pressure, interstellar turbulence, ionization fraction, and the topology of the ISM. Using models from Liszt (2001), which assume a column density of hydrogen nuclei equal to 10^{19} cm $^{-2}$ and the temperature of the Ly- α radiation field being equal to the kinetic temperature, our T_s measurement constrains the fraction of Galactic flux from early-type stars which permeates HI clouds to $> 1 \times 10^{-4}$.

Our work provides the first observational evidence for the Ly- α mechanism acting throughout the bulk of the Galactic WNM. By increasing observational sensitivity to T_s in comparison with expectations for T_k , stacking can be used to constrain the importance of non-collisional excitation on HI, as well as the origin and intensity of the Ly- α radiation field. In the future, 21 SPONGE will obtain more absorption sightlines to increase our sensitivity to weak underlying

signatures of the WNM, thereby further refining these temperature constraints, and sampling different Galactic environments.

3.5 Conclusions

We have presented the discovery of a weak ($\tau_0 = 3.0_{-0.4}^{+1.0} \times 10^{-4}$), broad (FWHM = $50_{-7}^{+15} \text{ km s}^{-1}$) WNM absorption feature in the stacked absorption spectrum of 19 independent Galactic sight-lines from the 21-SPONGE survey. Using Monte Carlo simulations, we have estimated the feature's spin temperature to be $\langle T_s \rangle = 7200_{-1200}^{+1800} \text{ K}$, which is significantly (98% confidence) higher than theoretical predictions based on collisional excitation alone, likely due to the thermalization of HI by resonant Ly- α scattering. This work provides the first observational evidence that the Ly- α excitation mechanism is acting throughout the bulk of the Galactic WNM, and demonstrates that the Ly- α radiation field, a quantity that is difficult to measure yet vitally important for interpreting HI signals from early epochs of cosmic structure formation, can be probed using measurements of the WNM.

We thank an anonymous referee for helpful comments and suggestions. This work was supported by the NSF Early Career Development (CAREER) Award AST-1056780. C. M. acknowledges support by the National Science Foundation Graduate Research Fellowship and the Wisconsin Space Grant Institution. S. S. thanks the Research Corporation for Science Advancement for their support. We thank A. Begum for help with initial survey observations, D. Able for developing initial data reduction strategy, and H. Liszt, J. S. Gallagher, N. Roy and T. Wong for helpful discussions. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. The Arecibo Observatory is operated by SRI International under a cooperative agreement with the National Science Foundation (AST-1100968), and in alliance with Ana G. Méndez-Universidad Metropolitana, and the Universities Space Research Association.

References

- Begum, A., Stanimirović, S., Goss, W., et al. 2010, *ApJ*, 725, 1779
- Braun, R., & Walterbos, R. A. M. 1992, *ApJ*, 386, 120
- Bryan, G. 2007, in EAS Publications Series, Vol. 24, EAS Publications Series, ed. E. Emsellem, H. Wozniak, G. Massacrier, J.-F. Gonzalez, J. Devriendt, & N. Champavert, 77–88
- Carilli, C. L., Dwarakanath, K. S., & Goss, W. M. 1998, *ApJ*, 502, L79
- Christensen, C., Quinn, T., Governato, F., et al. 2012, *MNRAS*, 425, 3058
- Condon, J., Cotton, W., Greisen, E., et al. 1998, *AJ*, 115, 1693
- Crovisier, J., Kazes, I., & Aubry, D. 1978, *A&AS*, 32, 205
- Davies, R. D., & Cummings, E. R. 1975, *MNRAS*, 170, 95
- Deguchi, S., & Watson, W. D. 1985, *ApJ*, 290, 578
- Dickey, J., Kulkarni, S., van Gorkom, J., & Heiles, C. 1983, *ApJS*, 53, 591
- Dickey, J. M., Salpeter, E. E., & Terzian, Y. 1977, *ApJ*, 211, L77
- Dwarakanath, K. S., Carilli, C. L., & Goss, W. M. 2002, *ApJ*, 567, 940
- Field, G. B. 1958, *Proceedings of the IRE*, 46, 240
- Field, G. B., Goldsmith, D. W., & Habing, H. J. 1969, *ApJ*, 155, L149
- Heiles, C., & Troland, T. 2003, *ApJS*, 145, 329
- Kalberla, P., Burton, W., Hartmann, D., et al. 2005, *A&A*, 440, 775
- Kanekar, N., Braun, R., & Roy, N. 2011, *ApJ*, 737, L33
- Kanekar, N., Subrahmanyan, R., Chengalur, J. N., & Safouris, V. 2003, *MNRAS*, 346, L57
- Lazareff, B. 1975, *A&A*, 42, 25
- Liszt, H. 2001, *A&A*, 371, 698

- McKee, C., & Ostriker, J. 1977, ApJ, 218, 148
- Mebold, U., Winnberg, A., Kalberla, P. M. W., & Goss, W. M. 1982, A&A, 115, 223
- Mohan, R., Dwarakanath, K., & Srinivasan, G. 2004, Journal of Astrophysics and Astronomy, 25, 143
- Payne, H., Dickey, J., Salpeter, E., & Terzian, Y. 1978, ApJ, 221, L95
- Peek, J., Heiles, C., Douglas, K., et al. 2011, ApJS, 194, 20
- Pritchard, J. R., & Loeb, A. 2012, Reports on Progress in Physics, 75, 086901
- Roy, N., Chengalur, J. N., & Srianand, R. 2006, MNRAS, 365, L1
- Roy, N., Kanekar, N., Braun, R., & Chengalur, J. 2013a, ArXiv e-prints, 1309.4098
- Roy, N., Kanekar, N., & Chengalur, J. 2013b, ArXiv e-prints, arXiv:1309.4099
- Stanimirović, S., Putman, M., Heiles, C., et al. 2006, ApJ, 653, 1210
- Treister, E., Schawinski, K., Volonteri, M., Natarajan, P., & Gawiser, E. 2011, Nature, 474, 356
- Wall, J. V., & Jenkins, C. R. 2003, Practical Statistics for Astronomers (Cambridge University Press)
- Wolfire, M., McKee, C., Hollenbach, D., & Tielens, A. 2003, ApJ, 587, 278
- Wouthuysen, S. A. 1952, AJ, 57, 31

Chapter 4

Recovering Interstellar Gas Properties with HI

Spectral Lines: A comparison between Synthetic Spectra and 21-SPONGE

*A version of this chapter has previously appeared
in the Astrophysical Journal*

C. E. Murray, S. Stanimirović, C.-G. Kim, et al. 2017, ApJ, 837, 55

Abstract

We analyze synthetic neutral hydrogen (H I) absorption and emission spectral lines from a high-resolution, three-dimensional hydrodynamical simulation to quantify how well observational methods recover the physical properties of interstellar gas. We present a new method for uniformly decomposing H I spectral lines and estimating the properties of associated gas using the Autonomous Gaussian Decomposition (AGD) algorithm. We find that H I spectral lines recover physical structures in the simulation with excellent completeness at high Galactic latitude, and this completeness declines with decreasing latitude due to strong velocity-blending of spectral lines. The temperature and column density inferred from our decomposition and radiative transfer method agree with the simulated values within a factor of < 2 for the majority of gas structures. We next compare synthetic spectra with observations from the 21-SPONGE survey at the Karl G. Jansky Very Large Array using AGD. We find more components per line of sight in 21-SPONGE than in synthetic spectra, which reflects insufficient simulated gas scale heights and the limitations of local box simulations. In addition, we find a significant population of low-optical depth, broad absorption components in the synthetic data which are not seen in 21-SPONGE. This population is not obvious in integrated or per-channel diagnostics, and reflects the benefit of studying velocity-resolved components. The discrepant components correspond to the highest spin temperatures ($1000 < T_s < 4000 \text{ K}$), which are not seen in 21-SPONGE despite sufficient observational sensitivity. We demonstrate that our analysis method is a powerful tool for diagnosing neutral ISM conditions, and future work is needed to improve observational statistics and implementation of simulated physics.

4.1 Introduction

Neutral hydrogen (H I) in the interstellar medium (ISM) plays a crucial role in the life cycles of galaxies. The atomic medium provides the main fuel reservoir for molecular gas and, ultimately, star formation. Furthermore, the structure of interstellar H I bears important clues to the nature of gas recycling via radiative and dynamical feedback and Galactic winds (e.g., McClure-Griffiths et al. 2013).

Throughout the ISM, H I exists in a “multi-phase” state, characterized by two thermally-stable phases in pressure equilibrium (Field et al. 1969; McKee & Ostriker 1977; Wolfire et al. 2003): the cold neutral medium (CNM) and warm neutral medium (WNM). An effective constraining observable for the balance between the CNM and WNM is the excitation temperature (a.k.a., spin temperature, T_s) of the gas. However, both emission and absorption by the 21 cm hyperfine transition of H I are required to measure T_s . Therefore, although the CNM ($T_s \sim 20 - 200 \text{ K}$) has been extensively analyzed with H I absorption (e.g., Crovisier et al. 1978; Dickey et al. 2003; Heiles & Troland 2003a), excellent sensitivity is required to constrain the temperature of the WNM ($T_s \sim 1000 - 7000 \text{ K}$), and few measurements exist (Carilli et al. 1998; Dwarakanath et al. 2002; Roy et al. 2013, , Chapter 3).

Furthermore, to understand the physical mechanisms responsible for observed H I properties, comparisons between observations and theory are necessary. Synthetic datasets from numerical simulations provide a means to (1) assess the power of observational diagnostics to reveal the inherent state of astronomical systems, and (2) test whether complex simulations recover all the properties of real systems. For example, the velocity structures of synthetic spectral lines provide important diagnostics of interstellar turbulence (e.g., Falgarone et al. 1994), and the nature of CNM dynamics (Hennebelle et al. 2007; Saury et al. 2014). Furthermore, synthetic observations have been used extensively to investigate molecule formation (Shetty et al. 2011; Smith et al. 2014; Duarte-Cabral et al. 2015; Duarte-Cabral & Dobbs 2016) and Galactic morphology (Douglas et al.

2010; Acreman et al. 2012; Pettitt et al. 2014). Important observational biases can be directly quantified using these comparisons. For example, considering correspondence between the true positions and observed velocities of molecular clouds, Beaumont et al. (2013) showed that the superposition of clouds along the line of sight introduces significant uncertainty to observational estimates of cloud mass, size and velocity dispersion.

However, numerical simulations with suitable dynamic range and resolution for describing the dynamics of both the CNM and WNM have only recently been performed. Kim et al. (2014) constructed a sample of synthetic H I absorption and emission spectral lines from their three-dimensional hydrodynamical simulations (Kim et al. 2013). Comparing conditions in the simulated data with properties inferred from synthetic spectra, Kim et al. (2014) found excellent agreement between “true” and “observed” per-channel and line of sight (LOS)-integrated properties such as column density and spin temperature. Furthermore, they showed that column densities computed in the optically-thin limit significantly underestimate the true column density when the H I optical depth is greater than $\tau \sim 1$. This agrees with previous comparisons of observed and simulated LOS column density by Chengalur et al. (2013), who used Monte Carlo simulated spectra to test the role of optically-thick H I. However, Kim et al. (2014) found the discrepancy factor to be much smaller than Chengalur et al. (2013), indicating that when proper dynamics are considered, spectral line blending due to overlapping cold clouds is not significant.

Although LOS-integrated ISM properties provide important diagnostics, interpretation is more complicated for multi-temperature, as opposed to isothermal, conditions. Observations show that even high galactic latitude lines of sight contain several components of varying spin temperature (e.g., Heiles & Troland 2003a; Begum et al. 2010; Roy et al. 2013, Chapter 2). The technique of Gaussian decomposition is one method that can identify individual spectral components from disparate H I phases, and has been used extensively to disentangle complex spectral lines (e.g., Lazareff 1975; Mebold et al. 1982; Dickey & Lockman 1990; Mohan et al. 2004). When applied to both H I absorption and emission spectra, Gaussian decomposition can be used

to estimate the spin temperatures of individual spectral features and the fraction of CNM along the line of sight (Heiles & Troland 2003a; Stanimirović et al. 2014, Chapter 2). However, the method suffers from non-uniqueness complications, as Gaussian functions do not form an orthogonal basis (e.g., Heiles & Troland 2003a) and the number of Gaussian functions used to produce reasonable spectral fits can vary significantly. For example, comparing the Gaussian decompositions of the Galactic H I absorption spectrum towards 3C138, there are no fitted components which agree between the 6 found in Chapter 2 and the 13 found by Roy et al. (2013). Furthermore, no quantitative estimates of how well Gaussian functions recover the properties of interstellar gas exist.

One of the major scientific goals for future H I observations with the Australian Square Kilometer Array Pathfinder (ASKAP) and the Square Kilometer Array (SKA) will be to understand the temperature distribution of the ISM and how it relates to the life cycles of galaxies using thousands of spectra (e.g., Dickey et al. 2013; McClure-Griffiths et al. 2013). However, the first step in this undertaking is to understand the biases and limitations of our observational

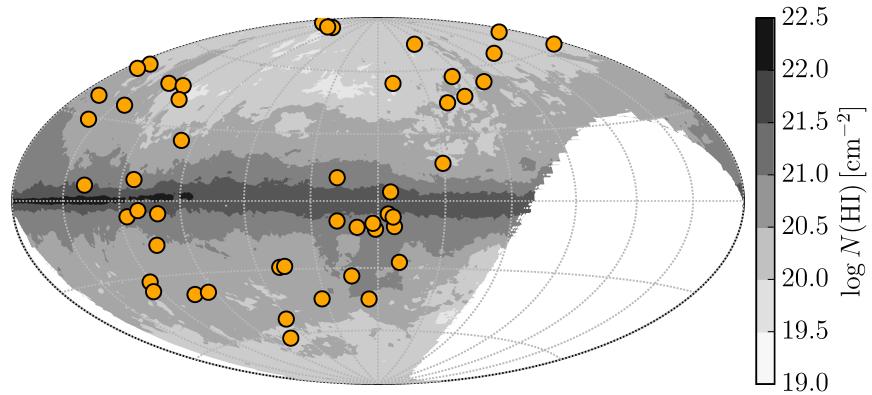


Figure 4.1 Galactic H I column density ($N(\text{H}\text{I})$) map from the Leiden Dwingeloo Survey (Hartmann & Burton 1997). The locations of the 52 21-SPONGE H I absorption line sources (orange circles) are overlaid.

and analysis methods in reproducing interstellar gas properties. This is best done by analyzing synthetic H I data from numerical simulations, which include realistic physical processes, and also provide full 3D information on simulated H I structures (e.g. density, temperature, velocity).

Accordingly, we begin this paper by quantifying how well Gaussian analysis of H I spectral lines via simple radiative transfer recovers “true” interstellar gas properties by analyzing synthetic 21 cm spectral profiles derived from 3D hydrodynamical simulations from Kim et al. (2013, 2014). To analyze the synthetic spectral profiles, we use the Autonomous Gaussian Decomposition (AGD) algorithm (Lindner et al. 2015). AGD implements derivative-based computer vision to perform Gaussian decomposition of spectral lines, enabling efficient, reproducible and objective spectral decomposition.

In the second half of the paper, we compare synthetic and observed H I spectra objectively using the same methodology. For this we use data from the 21-cm Spectral Line Observations of Neutral Gas with the VLA (21-SPONGE) survey, one of the most sensitive Galactic 21 cm surveys, as well as the Kim et al. synthetic H I spectra. We assess the ways in which the detailed statistical properties of synthetic spectra may agree or disagree with the statistics of observed spectra. This in turn reflects the influence of the star formation feedback mechanisms and other physics of the simulations. We especially focus on the importance of Ly α resonant scattering for H I excitation and the temperature distribution.

In Section 4.2, we describe the 21-SPONGE observations, and in Section 4.3 we describe Kim et al. simulations and synthetic data. We present and discuss our analysis method in Section 4.4. In Section 4.5, we compare the properties inferred from synthetic spectra with the simulated properties of gas along the same LOS. We then compare the synthetic spectra with 21-SPONGE observations in Section 4.6. Finally, we present our summary and conclusions in Section 4.7.

4.2 Observations

For observations of Galactic H I, we use data from the 21-SPONGE survey (Chapter 2). 21-SPONGE is the most sensitive survey of Galactic H I absorption at the Karl G. Jansky Very Large Array (VLA) to date. We target strong extragalactic radio continuum sources mostly at high Galactic latitudes ($|b| > 5^\circ$), and consistently reach RMS noise in H I optical depth of $\sigma_\tau < 10^{-3}$ per 0.4 km s^{-1} channels.

21-SPONGE utilizes H I emission observations along the same LOS from the Arecibo Observatory ($\sim 3.5'$ beam at 21 cm). The emission spectrum in the direction of each source is computed by observing a pattern of 16 off-source positions and interpolating across the target position (see, e.g., Heiles & Troland 2003a, hereafter HT03). We note that the Arecibo H I emission profiles are not corrected for an effect known as “stray radiation”, wherein radiation enters the main telescope beam through higher-order side lobes. Although stray radiation can be modeled and removed from H I emission data (e.g., LAB and GASS surveys; Kalberla et al. 2005; McClure-Griffiths et al. 2009; Kalberla et al. 2010), it is a complex process requiring stable beam shapes which are not achieved at Arecibo. Comparison between the GALFA-H I survey at Arecibo and the stray radiation-corrected LAB survey suggested that stray radiation likely does not contribute more than $\sim 500 \text{ mK}$ over $\sim 50 \text{ km s}^{-1}$ to observed H I brightness temperature (Peek et al. 2011). We emphasize that the effect is only significant for emission, not absorption.

Armed with high-sensitivity H I absorption and emission spectra along each line of sight, 21-SPONGE is sensitive to H I column densities and temperatures from all neutral ISM phases, including the CNM, WNM and thermally unstable medium.

In Chapter 2, we presented the survey design, analysis methods and initial results for 21-SPONGE. To derive physical properties of interstellar gas along each line of sight, we decomposed (by hand) H I absorption and emission spectral pairs simultaneously into Gaussian functions, and solved radiative transfer equations to derive the column density and spin temperature of individual spectral components, taking into account the presence of self-absorption and the order of features

along each line of sight (as done, e.g., in HT03, Stanimirović & Heiles 2005; Stanimirović et al. 2014). We found excellent agreement with previous H I absorption surveys. The high sensitivity of 21-SPONGE allowed us to extend the maximum H I spin temperatures detected directly in absorption and emission from 600 K in HT03 to ~ 1500 K (Chapter 2).

Figure 4.1 displays the positions of the 52 21-SPONGE sources overlaid on an H I column density map from the Leiden Dwingeloo Survey (LDS; Hartmann & Burton 1997). The targets probe a large range in Galactic latitude.

4.3 Numerical simulations

We analyze recent high-resolution Galactic ISM simulations by Kim et al. (2013, hereafter KOK13). These simulations include momentum feedback from supernovae, time-varying heating, interstellar cooling appropriate for warm/cold gas, galactic differential rotation, gaseous self-gravity and external gravity from dark matter and stars. We refer the reader to KOK13 for a full description of the numerical setup, and methods.

Using these simulations, Kim et al. (2014, hereafter KOK14) constructed a set of synthetic spectral lines sampling the local ISM. Assuming an observer sits in the center of the simulation, they selected 10^4 positions randomly distributed in Galactic latitude (l) and longitude (b) and extracted the number density (n), temperature (kinetic, T_k , and spin, T_s) and velocity (v) as functions of path length (s). These lines of sight (LOS) are restricted to $|b| > 4.9^\circ$ and $s \leq 3$ kpc so that the limited horizontal extent of the simulation does not adversely affect the results, as would be a concern at low Galactic latitude. For the particular KOK13 simulation we analyze in this work (the model denoted “QA10”), galactic rotation was applied with assumed angular velocity of $\Omega = 28$ km sec $^{-1}$ and gas surface density $\Sigma = 10$ M $_{\odot}$ pc $^{-2}$ (KOK13).

For each observed LOS, KOK14 produced synthetic H I 21 cm emission and absorption as functions of radial velocity using analytical radiative transfer and a simple prescription for line excitation prescriptions. The reader is referred to Section 2.3 of KOK14 for a complete description of the methods used to construct the synthetic spectra.

Table 4.1. AGD Summary

Source	LOS (number)	Absorption (total)	N_{AGD} (per LOS) ^a	Emission (N_{AGD})		Matched ^b (total)
				(total)	(per LOS) ^a	
KOK14	9355	14023	1.5 ± 1.7	23475	2.5 ± 1.4	9218
KOK14 (no WF)	9355	15468	1.7 ± 1.8	23519	2.5 ± 1.4	9490
21-SPONGE	52	237	4.6 ± 3.0	326	6.3 ± 2.9	88

Note. — ^a: Mean and standard deviation over all LOS; ^b: “Matched” statistics will be discussed in Section 6.

In particular, as part of their model for synthetic 21 cm level populations, KOK14 considered indirect radiative transitions due to resonant scattering by Ly α photons (the Wouthuysen-Field (WF) effect; Wouthuysen 1952; Field 1959) in addition to collisions and direct radiative transitions. They parameterized the WF effect following Field (1959) with a constant value for the Galactic Ly α photon number density, n_α , inferred from Liszt (2001) to be $n_\alpha = 10^{-6} \text{ cm}^{-3}$. This value of n_α is highly uncertain and difficult to constrain observationally or numerically. Given that observed LOS-integrated and per-channel properties are dominated by high-optical depth gas, wherein the 21 cm transition is already thermalized by collisions due to high densities, the WF effect does not significantly affect these values. The WF effect should be most important for the WNM, where generally $T_s \leq T_k$ due to the inefficiency of collisions at thermalizing the 21 cm transition (e.g., Liszt 2001). Indeed, at high T_s the WF effect is significant (c.f., Figures 9a, 10a of KOK14).

We note that the KOK13 simulations do not include chemistry and the H I-to-H₂ conversion. In addition, their implemented supernova feedback injects momentum and not thermal energy, resulting in the absence of a hot ($T \sim 10^5 - 10^7 \text{ K}$) medium. Although for the warm and cold medium these are secondary effects, and they are being addressed in ongoing simulations with thermal supernova feedback to create a hot ISM (Kim & Ostriker 2016), we need to keep these limitations in mind when considering properties of synthetic spectra from KOK14.

In this paper, we analyze components within the synthetic KOK14 H I spectral pairs to investigate how radiative transfer-based Gaussian fitting reproduces real physical quantities. From their catalog of 10^4 spectral pairs, we selected those without saturated (defined here as $\tau \geq 3$) or NaN-valued absorption lines, for a final catalog of 9355 H I spectral pairs. To simulate observational conditions, we added Gaussian-distributed noise with an amplitude of $\sigma_\tau = 10^{-3}$ to each absorption spectrum (equal to the median RMS noise in τ per channel from 21-SPONGE) and $\sigma_{T_B} = 0.2 \text{ K}$ to each emission spectrum (equal to the median RMS noise in T_B per channel in 21-SPONGE).

4.4 Gaussian Decomposition with AGD

To perform Gaussian fits to H I spectra (either real or synthetic), we use the Autonomous Gaussian Decomposition algorithm (AGD; Lindner et al. 2015). AGD implements derivative spectroscopy and machine learning techniques to efficiently and objectively provide initial guesses (i.e., amplitude, width, mean velocity) for multiple-Gaussian fits to spectral line data.

Before implementing AGD, we trained the algorithm to maximize the decomposition accuracy. We began by constructing a synthetic H I dataset from the Gaussian components detected by the Millennium Arecibo 21 cm Absorption Line Survey (HT03, Heiles & Troland 2003b). The synthetic training dataset construction and training are described fully in Lindner et al. (2015), and summarized here for clarity. We selected the number of components in each synthetic spectrum to be a uniform random integer ranging from the mean number of components in the survey (3) to the maximum number (8; HT03), and then drew the component parameters from the published HT03 amplitude, FWHM and mean velocity distributions with replacement. As done with the KOK14 synthetic spectra, we added Gaussian-distributed noise with an amplitude of $\sigma_\tau = 10^{-3}$ to each absorption spectrum (equal to the median RMS noise in τ per channel from 21-SPONGE) and $\sigma_{T_B} = 0.2$ K to each emission spectrum (equal to the median RMS noise in T_B per channel in 21-SPONGE). The synthetic training sets for absorption and emission consist of 20 spectra each.

After constructing the synthetic training dataset, we used the Python implementation of AGD, GaussPy, to decompose the synthetic training dataset for different values of the “two-phase” smoothing parameters α_1 and α_2 . These smoothing parameters serve to identify the types of spectral properties present in the data. Beginning with initial choices for α_1 and α_2 and a signal-to-noise (S/N) threshold below which the algorithm will not select components (S/N=3.0), GaussPy computes the accuracy of the decomposition (i.e. how closely the derived model parameters are to the true model parameters), for iteratively different values of α_1 and α_2 until it converges on minimal model residuals and maximum decomposition accuracy. After

training, we found accuracies of 80% and 70% for H I absorption and emission decompositions, respectively. The resulting values are $\alpha_1 = 1.12$ and $\alpha_2 = 2.73$ for absorption and $\alpha_1 = 1.70$ and $\alpha_2 = 3.75$ for emission.

With trained values of α_1 and α_2 in hand, we used GaussPy to apply the AGD algorithm identically (i.e. same values of α_1 , α_2 and S/N) to the observed 21-SPONGE and simulated KOK14 spectra to derive lists of Gaussian parameters for each dataset. Table 4.1 summarizes the decomposition results for the emission and absorption spectra from 21-SPONGE and KOK14 including the WF effect and without the WF effect (“no WF”). The typical uncertainties in the fitted parameters are $\sim 1 - 10\%$ from the least-squares fit applied by AGD.

4.5 Assessing the power of the Gaussian-fitting method with synthetic H I spectra

From the KOK14 simulations, we have information about the density and spin temperature as a function of distance ($n(s)$, $T_s(s)$), as well as the optical depth and brightness temperature as a function of velocity ($\tau(v)$, $T_B(v)$) for each line of sight. Therefore, following AGD analysis, we can compare inferred properties from spectral lines to the true gas properties within the simulated ISM. This will allow us to quantify the biases and limitations of Gaussian analysis in reproducing realistic physical properties.

4.5.1 Defining gas structures in position and velocity

Given that $\tau \propto n/T_s$, we define simulated gas structures by selecting peaks in n/T_s along each LOS. To select a threshold value, we consider the parameters of the simulated ISM from KOK13. The gas temperature and density PDFs in their Figure 8 display strong bi-modality indicative of multiple H I phases. The ratio n/T_s is high for the CNM and low for the WNM. In identifying gas “structures” along the LOS, we wish to mark concentrations using CNM-like peaks. We select $n \sim 2 \text{ cm}^{-3}$ and $T \sim 10^3 \text{ K}$ as representative values between the peaks of the published bi-modal PDFs (KOK13). These values correspond to a threshold of $(n/T_s)_{\text{thresh}} = 0.002 \text{ cm}^{-3}/\text{K}$.

We experimented with different values of this threshold, and the subsequent results do not change significantly.

In Figure 4.2, we display $(n/T_s)(s)$ (left), $\tau(v)$ (middle) and $T_B(v)$ (right) for 5 example LOS from KOK14. The positions of peaks above $(n/T_s)_{\text{thresh}} = 0.002 \text{ cm}^{-3}/\text{K}$ (“structures”) are plotted as colored circles. Across all 9355 synthetic LOS, there are 7582 structures with $(n/T_s) > 0.002 \text{ cm}^{-3}/\text{K}$.

To compare the properties of simulated gas structures with synthetic spectral lines, we first

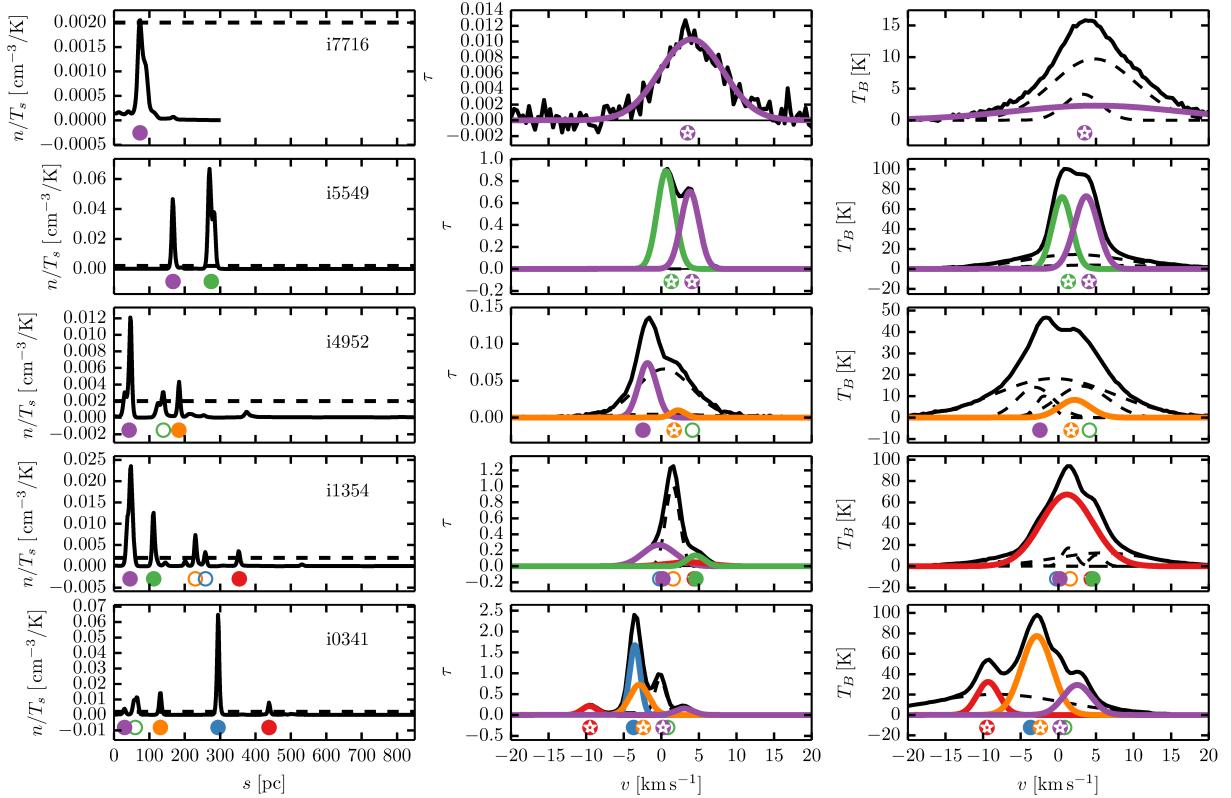


Figure 4.2 Example LOS from KOK14. Left: density over spin temperature (n/T_s) as a function of distance (s) along the LOS; Middle: synthetic optical depth as a function of velocity ($\tau(v)$); Right: synthetic brightness temperature as a function of velocity ($T_B(v)$). Gas structures defined by peaks above $(n/T_s)_{\text{thresh}} = 0.002 \text{ cm}^{-3}/\text{K}$ (dashed line, left column) are indicated by colored circles in each panel. If a structure from the left column matches with an AGD absorption component in the middle column according to Equations 4.6 and 4.8, the circles are filled and the matching AGD component is plotted in the corresponding color (unmatched components are plotted in dashed lines). If a structure matches with an AGD absorption and emission component according to Equations 4.9 and 4.10, the circle symbol has a white star within it.

determine the position and velocity range (i.e. line width) of each physical gas structure. We estimate the velocity of each gas structure, v_{sim} , by computing the average velocity ($v(s)$) of channels spanned by each peak in n/T_s weighted by their densities ($n(s)$), specifically,

$$v_{\text{sim}} = \frac{\int_{\text{structure}} n(s) v(s) ds}{\int_{\text{structure}} n(s) ds}, \quad (4.1)$$

where “structure” refers to all pixels above $(n/T_s)_{\text{thresh}} = 0.002 \text{ cm}^{-3}/\text{K}$. These values are plotted in the middle and right columns of Figure 4.2 as circles with colors corresponding to the labels in the left column (n/T_s). Next, we estimate the FWHM of the structure based on its thermal and turbulent properties. We compute the inferred thermal line width, Δv_{therm} , by solving (e.g., Eq. 9.31, Draine 2011),

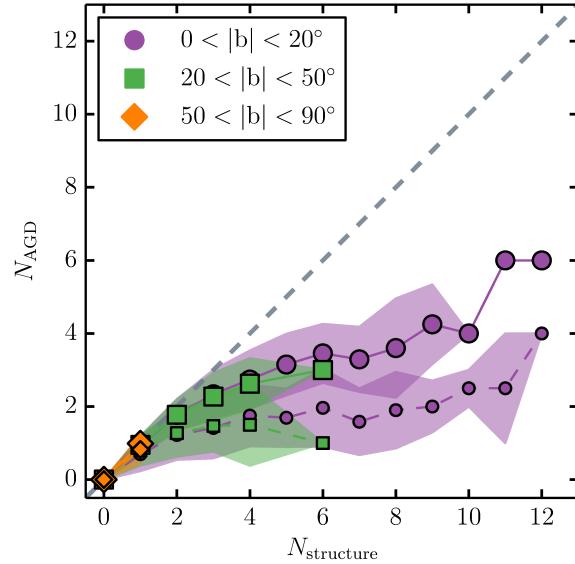


Figure 4.3 The number of gas structures defined by $n/T_s > 0.002 \text{ cm}^{-3}/\text{K}$ along an LOS ($N_{\text{structure}}$) compared with the number of AGD-fitted features in the synthetic line profile (N_{AGD}). Large symbols show N_{AGD} matched in absorption according to Equations 4.6 and 4.8, and small symbols show N_{AGD} matched in absorption and emission according to Equations 4.6, 4.8, 4.9 and 4.10. Symbols and shading indicate the mean and standard deviations of N_{AGD} for each unique $N_{\text{structure}}$. The samples are binned by latitude according to the inset legend.

$$\Delta v_{\text{therm}} = 2.15 \sqrt{\frac{T_{\text{mean}}/100 \text{ K}}{M/m_{\text{H}}}} = 0.190 \sqrt{T_{\text{mean}}} \text{ km s}^{-1} \quad (4.2)$$

where we assume $M = \mu m_{\text{H}}$ for mean molecular weight $\mu = 1.27$ (c.f., KOK13), and T_{mean} is the harmonic mean kinetic temperature of the gas spanned by each peak, given by,

$$T_{\text{mean}} = \frac{\int_{\text{structure}} n(s) ds}{\int_{\text{structure}} (n/T)(s) ds}. \quad (4.3)$$

We then estimate the contribution from turbulent line broadening, Δv_{turb} , by computing the standard deviation of the velocities spanned by each structure, multiplied by a factor of 2.355 to convert to a FWHM, or,

$$\Delta v_{\text{turb}} = 2.355 \sqrt{\frac{\int_{\text{structure}} n(s) (v(s) - v_{\text{sim}})^2 ds}{\int_{\text{structure}} n(s) ds}}. \quad (4.4)$$

The final estimate of the velocity FWHM of each structure, Δv_{sim} , is a quadratic sum of the thermal and turbulent contributions (Equations 4.2 and 4.4), or $\Delta v_{\text{sim}} = \sqrt{\Delta v_{\text{therm}}^2 + \Delta v_{\text{turb}}^2}$, and has values between $\sim 1 - 10 \text{ km s}^{-1}$ for a range of density threshold choices.

4.5.2 Matching gas structures with H I absorption lines

To match AGD-fitted Gaussian absorption lines with gas structures along each LOS, we use two matching criteria. First, we define δ_v to be the difference in mean velocity of a Gaussian component fitted by AGD in absorption (v_0) with the estimated velocity of the gas structure (v_{sim}) in terms of the measured FWHM from the AGD fit (Δv_0), or,

$$\delta_v \equiv \frac{|v_0 - v_{\text{sim}}|}{\Delta v_0 / 2.355}, \quad (4.5)$$

For a gas structure to match a Gaussian component, we require that their positions in velocity be less than one standard deviation away from each other, so that,

$$\delta_v \leq 1. \quad (4.6)$$

Second, we define R_{FWHM} to be the ratio of the FWHM of a component in absorption (Δv_0) and the estimated FWHM of the gas structure in velocity (Δv_{sim}), or,

$$R_{\text{FWHM}} \equiv \Delta v_{\text{sim}} / \Delta v_0. \quad (4.7)$$

For a structure to match a Gaussian component, we require that the structure's simulated velocity FWHM (Δv_{sim}), including the thermal and turbulent contributions, be similar to the FWHM of the Gaussian component (within a factor of 3), or,

$$0.3 \leq R_{\text{FWHM}} \leq 3. \quad (4.8)$$

We note that choices of cutoff values for R_{FWHM} does not significantly change the results, as the criterion described by Equation 4.6 dominates the matching. In addition, we emphasize that the matching criteria were designed to be as simple as possible to minimize imposed selection biases.

In Figure 4.2, the circle markers for structures which match with AGD absorption components according to Equations 4.6 and 4.8, are filled, and the matching AGD absorption component is plotted in the corresponding color in the middle panel. If a structure does not have an AGD match, the circle marker is unfilled. Of the 7582 total structures, there are 6097 structures with matches to AGD absorption components.

4.5.2.1 Matching H I absorption lines with H I emission lines

Both H I absorption ($\tau(v)$) and emission ($T_B(v)$) information are required to constrain the spin temperature (T_s) and column density ($N(\text{HI})$) of neutral gas in the ISM via radiative transfer. Therefore, to compare the density and temperature of simulated gas structures with properties inferred from observations, we need to determine the optical depth and brightness temperature of each structure.

To match H I absorption lines with H I emission lines fitted by AGD, we apply a similar set of match criteria as described by Equations 4.6 and 4.8. Specifically,

$$\delta_{v,\text{AGD}} \equiv \frac{|v_0 - v_{0,\text{em}}|}{\Delta v_0 / 2.355} \leq 1, \quad (4.9)$$

$$1 \leq R_{\text{FWHM,AGD}} \equiv \Delta v_{0,\text{em}} / \Delta v_0 \leq 3, \quad (4.10)$$

where $(v_{0,\text{em}}, \Delta v_{0,\text{em}})$ are the mean velocity and FWHM of an AGD component fitted to $T_B(v)$.

We impose the requirement that $R_{\text{FWHM,AGD}} \geq 1$ to ensure that the matched line width is larger in emission than absorption, and impose $R_{\text{FWHM,AGD}} \leq 3$ in order to ensure that the line widths are reasonably similar. We do not include a criterion for matching component amplitudes here, because the amplitude of an absorption feature in emission is determined by both its optical depth and its spin temperature, which are difficult to disentangle. Furthermore, we are interested in analyzing how well this simple approach recovers the spin temperatures of structures, which we analyze in Section 5.4.1, and therefore we do not impose any requirement that the component amplitudes match at this stage.

In Figure 4.2, if a structure (left panel) matches with an AGD absorption line (middle panel) and also matches with an AGD emission line (right panel) according to Equations 4.9 and 4.10, the circle marker contains a white star and the matching emission line is plotted in the corresponding color (right panel). Of the 6097 structures with AGD absorption line matches, 4228 also have a match to an AGD emission line.

4.5.3 Quantifying Match Completeness

For most examples in Figure 4.2, gas structures (left) are accounted for by the majority of the total optical depth along the LOS (middle). This suggests that the AGD absorption lines can be mapped to real structures. However, in the presence of strong line blending, as shown by the third row of Figure 4.2 (e.g., case i4952), several absorption lines have nearly the same central velocity and the majority of the absorption feature cannot be matched. Although the majority (90%) of LOS in KOK14 have < 2 fitted components, we selected the examples in Figure 4.2 to illustrate a range in complexity for our fitting and matching process. In particular, the more

complex LOS shown in the bottom three rows of Figure 4.2, with > 2 fitted components, are likely more representative of real observations. We discuss this issue further in Section 6.2.

In Figure 4.3 we compare the number of structures ($N_{\text{structure}}$) along each LOS with the number of matched AGD-fitted components of the synthetic line profile (N_{AGD}). For each unique value of $N_{\text{structure}}$, the mean (symbol) and standard deviation (shading) of N_{AGD} are shown. Furthermore, we break the sample into latitude bins according to the inset legend. Large symbols indicate the number of AGD absorption matches according Equations 4.6 and 4.8, and small symbols indicate the number of AGD absorption and emission matches following the subsequent application of Equations 4.9 and 4.10.

Using Figure 4.3, we quantify the completeness (C) of recovery by,

$$C = \frac{\sum N_{\text{AGD}}}{\sum N_{\text{structure}}}. \quad (4.11)$$

For the matches between gas structures and absorption lines only (large symbols), the completeness for the low ($0 < |b| < 20^\circ$), mid ($20 < |b| < 50^\circ$) and high latitude ($|b| > 50^\circ$) bins are $C_{\text{absorption}} = (0.53, 0.67, 0.99)$ respectively. The recovery is best at the highest latitudes where the number of gas structures and AGD components are smallest and the LOS complexity is minimized, allowing for simple and robust AGD fits. At low latitudes, the blending of gas velocities and AGD components makes it difficult to associate unambiguous spectral components with (n/T_s) peaks.

When matching gas structures to H I emission instead of absorption, the recovery completeness is $C_{\text{emission}} = (0.46, 0.53, 0.93)$ for low, mid and high latitudes respectively. The completeness is worse in emission than absorption. As observed in Figure 4.3, broad components associated with high-temperature H I are prominent, thereby making the match to corresponding gas structures more difficult. Although Gaussian analysis has been used extensively in the past to identify gas populations in local and external galaxies, this is among the first statistical quantification of the correspondence between Gaussian H I emission components and individual gas structures.

When matching structures to both absorption and emission (small symbols in Figure 4.3), the completeness for the low, mid and high latitude bins are: $C_{\text{both}} = (0.29, 0.38, 0.83)$ respectively. At all latitudes, the structure recovery completeness declines when the match between absorption and emission is performed. The bottom row of Figure 4.2 displays an example of this decline. Whereas 4/5 structures along the LOS are recovered by AGD absorption components, only 2/4 of those absorption components have matches in emission according to Equations 4.9 and 4.10. The structures selected by the matching process with both absorption and emission are biased towards unambiguous features in all three spaces. Nevertheless, the completeness of Gaussian decomposition for the multi-phase structures seen in both emission and absorption is good at $|b| > 50^\circ$. This is certainly promising for future large data sets – even with $\sim 50\%$ attrition a large number of structures can be retrieved over a wide range of interstellar conditions.

As discussed in the context of Figure 4.2, in some cases the fraction of total absorption or total emission that can be accounted for by gas structures along the LOS is low. This is especially true when $N_{\text{AGD}} \geq 2$, although those comprise the minority of cases in KOK14. In the future, improving our selection method for gas structures along the LOS beyond a single cutoff value in (n/T_s) will improve this completeness of structure recovery. In addition, developing additional criteria for structure-component matching based on their amplitudes and/or total column densities will enable us to better quantify the range of gas structures that can be recovered reliably by fitted spectral lines. The analysis presented here represents a first step in our ongoing investigation.

4.5.4 Observed vs. “True” Gas Properties

Given a sample of gas structures with matches to AGD components in absorption and emission, we compare the true temperatures in the simulation with the values inferred from AGD-fitting of the spectra.

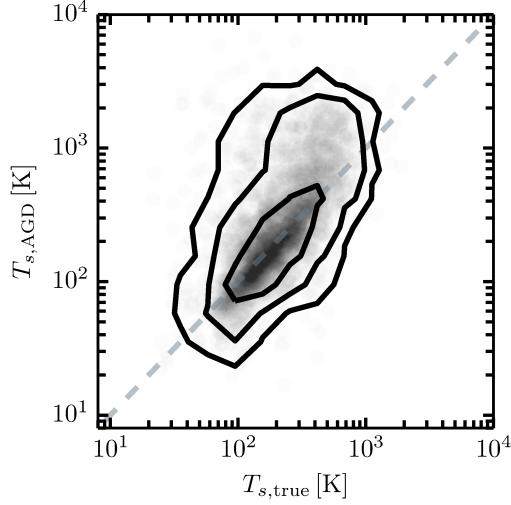


Figure 4.4 “True” simulated spin temperature ($T_{s,\text{true}}$, Equation 4.14) versus inferred spin temperature ($T_{s,\text{AGD}}$, Equation 4.13) for all structures which match fitted absorption and emission lines according to Equations 4.6, 4.8, 4.9 and 4.10. Contours indicate the 1, 2 and 3σ limits.

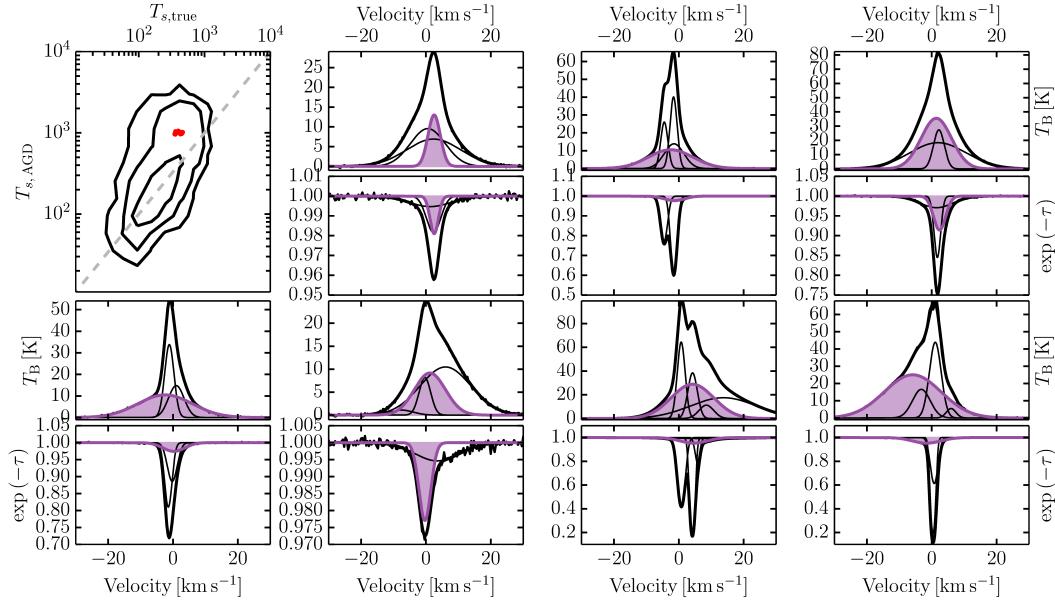


Figure 4.5 Example synthetic H I emission and absorption spectral pairs in which the AGD-derived spin temperature ($T_{s,\text{AGD}}$) overestimates the simulated spin temperature ($T_{s,\text{true}}$). The top-left panel is reproduced Figure 4.4, with 7 examples around $T_{s,\text{true}} = 400$ K and $T_{s,\text{AGD}} = 1000$ K highlighted in red. The matched H I absorption and emission pairs corresponding to these highlighted points are plotted and shaded in purple in the accompanying panels, along with the full AGD decomposition for each spectrum (black).

4.5.4.1 Spin Temperature

With the goal of estimating spin temperature automatically for a large number of spectra, we take a more simplified approach than what has been done in HT03 or Chapter 2. For each AGD match between absorption and emission, we start with the isothermal spin temperature as a function of velocity, $T_{s,\text{AGD}}(v)$, given by,

$$T_{s,\text{AGD}}(v) = \frac{T_{B,\text{AGD}}(v)}{1 - e^{-\tau_{\text{AGD}}(v)}}, \quad (4.12)$$

where $T_{B,\text{AGD}}(v)$ and $\tau_{\text{AGD}}(v)$ are the matched set of Gaussian functions fitted by AGD to $T_B(v)$ and $\tau(v)$ respectively. This method assumes a single temperature gas within each velocity channel. To estimate average spin temperature per AGD component, we compute the optical depth-weighted spin temperature per component,

$$T_{s,\text{AGD}} \equiv \frac{\int \tau_{\text{AGD}}(v) T_{s,\text{AGD}}(v) dv}{\int \tau_{\text{AGD}}(v) dv}. \quad (4.13)$$

This approach produces a weighted mean temperature for each component, given that $T_{s,\text{AGD}}(v)$ is smaller near the peak of $\tau_{\text{AGD}}(v)$ and larger away from it. We note that there are several possible ways to estimate mean temperature from $\tau(v)$ and $T_B(v)$ observations, and some discussion of the pros and cons of each method are given in HT03 and Dickey et al. (2003). It is important to note that Equation 4.13 works well if, within a multi-phase structure, the CNM and WNM are centered at a similar radial velocity. However, if the CNM is shifted in velocity relative to the WNM due to turbulent motions, so that the peaks of $T_{B,\text{AGD}}(v)$ and $\tau_{\text{AGD}}(v)$ are slightly offset, Equation 4.13 will overestimate $T_{s,\text{AGD}}$. HT03 and Chapter 2 have allowed for the CNM motion relative to the WNM in their temperature estimates by using a more complex fitting approach where spin temperature is fitted simultaneously with all WNM components. This is, however, computationally expensive for us to implement at this stage.

To estimate the temperature of a simulated gas structure, we compute the harmonic mean temperature, $T_{s,\text{true}}$, within the pixels spanned by each peak in n/T_s -space. Specifically,

$$T_{s,\text{true}} = \frac{\int_{\text{structure}} n(s) ds}{\int_{\text{structure}} (n/T_s)(s) ds}. \quad (4.14)$$

In Figure 4.4, we compare the simulated spin temperature with the inferred spin temperature derived using our radiative transfer approach for all structures with AGD matches in absorption and emission according to Equations 4.6, 4.8, 4.9 and 4.10. We recover nearly the full range of spin temperatures found in KOK14, which indicates that our structure selection method is likely not missing a significant gas population. Furthermore, the AGD and true estimates agree within the 1σ contours of Figure 4.4. However, at high temperatures, where $T_{s,\text{true}} > 400$ K, the AGD temperature overestimates the true spin temperature.

To understand why AGD overestimates spin temperature at high temperatures, Figure 4.5 displays a set of example matched spectral lines with $T_{s,\text{true}} \sim 400$ K and $T_{s,\text{AGD}} \sim 1000$ K. We reproduce the contours from Figure 4.4 in the top left panel of Figure 4.5, and include the corresponding matched H I emission and absorption components in the accompanying panels, highlighted with purple shading. In all highlighted cases, the emission component is slightly

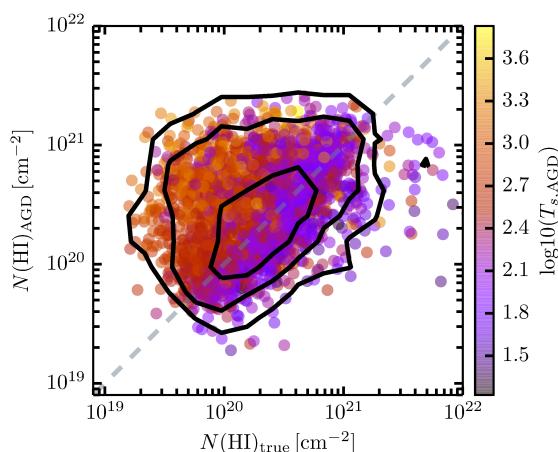


Figure 4.6 “True”, simulated column density ($N(\text{HI})_{\text{true}}$, Equation 4.17), versus “observed” column density ($N(\text{HI})_{\text{AGD}}$, Equation 4.16) for all peaks which match fitted absorption and emission components according to Equations 4.6, 4.8, 4.9 and 4.10. Contours indicate the 1, 2 and 3σ limits.

offset in velocity from the corresponding absorption component.

As discussed above, this offset is caused by interstellar turbulence. When Equation 4.13 is applied, the resulting spin temperature will be overestimated if we do not account for this velocity offset. For example, if we estimate the spin temperature using the peak brightness temperature and peak optical depth of highlighted components in Figure 4.5, we get a value that agrees much more closely with $T_{s,\text{true}}$, since $T_{B,\text{peak}}/(1 - e^{-\tau_{\text{peak}}}) \sim 400$ K. A more complex radiative transfer treatment such as the method of HT03 and Chapter 2, accounts for this effect, which is strongest for those components with the highest turbulent velocity offset between the CNM and WNM. We find that $T_{s,\text{AGD}}$ most strongly over-estimates $T_{s,\text{true}}$ when the velocity offset between absorption and emission is highest. However, most components are not affected. We will fine-tune our radiative transfer treatment in future work.

4.5.4.2 Column Density

The H I column density ($N(\text{HI})$) is given by,

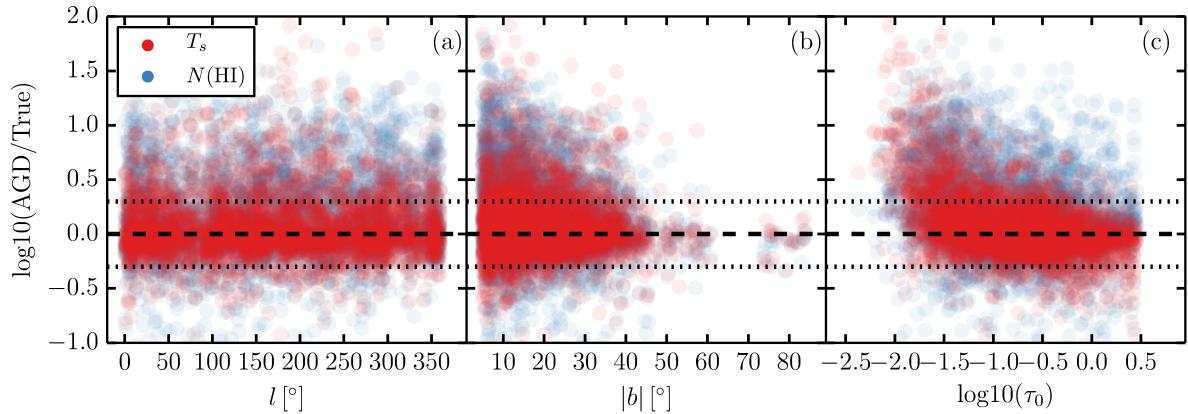


Figure 4.7 The ratio of inferred (“AGD”) to simulated (“True”) properties of gas structures matched with Gaussian spectral lines fitted to H I emission and absorption lines as a function of various LOS parameters. (a): Galactic longitude, $l [\circ]$; (b): absolute Galactic latitude, $|b| [\circ]$; (c): peak optical depth of the matched absorption line, τ_0 ; Spin temperature, $T_{s,\text{AGD}}/T_{s,\text{true}}$ (red) and column density, $N(\text{HI})_{\text{AGD}}/N(\text{HI})_{\text{true}}$ (blue). Dotted lines indicate factors of 2.

$$N(\text{HI}) = C_0 \int T_s(v) \tau(v) dv, \quad (4.15)$$

where $C_0 = 1.813 \times 10^{18} \text{ cm}^{-2} \text{ K}^{-1} (\text{km s}^{-1})$. For a pair of matched AGD absorption and emission lines, we compute the column density per component ($N(\text{HI})_{\text{AGD}}$) as,

$$N(\text{HI})_{\text{AGD}} = C_0 T_{s, \text{AGD}} \int \tau_{\text{AGD}}(v) dv, \quad (4.16)$$

where $T_{s, \text{AGD}}$ is computed using Equation 4.13 and $\tau_{\text{AGD}}(v)$ is the Gaussian function fitted by AGD to $\tau(v)$.

The simulated column density of each gas structure is given by,

$$N(\text{HI})_{\text{true}} = \int_{\text{structure}} n(s) ds. \quad (4.17)$$

In Figure 4.6, we compare the “true” and inferred column density for all structures with AGD matches in absorption and emission according to Equations 4.6, 4.8, 4.9 and 4.10. As in Figure 4.4, the AGD and true estimates agree within the 1σ contours. However, outside the 1σ contours, $N(\text{HI})_{\text{AGD}}$ overestimates $N(\text{HI})_{\text{true}}$, in part because the uncertainty in $T_{s, \text{AGD}}$ is propagated to $N(\text{HI})$ via Equation 4.16. We color the points in Figure 4.6 by $T_{s, \text{AGD}}$ to illustrate this. The most discrepant points correspond to the highest values of $T_{s, \text{AGD}}$, where the uncertainty in the matching process is highest (c.f., Figure 4.4).

In Figure 4.7, we investigate the scatter present in Figures 4.4 and 4.6 by plotting the ratios of the inferred (“AGD”) to direct (“true”) estimates of spin temperature and column density as a function of various LOS parameters. These include Galactic longitude ($|l|$; a), absolute Galactic latitude ($|b|$; b), and the peak optical depth of the matched AGD line (τ_0 ; c). In each panel, the data points are colored according to $T_{s, \text{AGD}}/T_{s, \text{true}}$ (red) and $N(\text{HI})_{\text{AGD}}/N(\text{HI})_{\text{true}}$ (blue).

In Figure 4.7, the ratio of AGD to true spin temperature and column density falls within a factor of 2 for the majority (68%, 1σ contours) of structures at all longitudes (a), latitudes (b) and peak optical depths (c) probed. This indicates that the AGD method is able to recover

these properties reasonably well. The scatter in the ratio of AGD to true column density is also larger than for spin temperature, due to the fact that any uncertainty in $T_{s,\text{AGD}}$ is propagated to $N(\text{H}\text{i})_{\text{AGD}}$ via Equation 4.16.

The scatter in AGD/True appears to be constant with longitude (l), yet increases at low latitudes ($|b|$) and low peak optical depths ($\tau_0 < 0.1$). This is caused by the increase in LOS complexity at low latitudes, whose effect was noted in Figure 4.3, and the increased likelihood for a low- τ_0 component to be affected by velocity blending. Examples of strong line blending are shown in Figure 4.5.

However, the agreement between AGD and true spin temperature and column density for the majority of simulated gas structures given our simple structure selection and matching prescription is encouraging, and indicates that automatic routines for identifying and analyzing spectral components from H I observations – essential for future large observed and simulated datasets – can be successful in recovering true properties for a large fraction of interstellar gas structures.

4.6 Comparing Real and Synthetic H I Spectra

After analyzing the biases of Gaussian analysis in recovering gas structures and their properties using synthetically spectra from simulations, we proceed to compare observed (21-SPONGE) and simulated (KOK14) H I spectra via AGD-fitted Gaussian parameters.

With the AGD decompositions of H I emission and absorption for all 21-SPONGE and KOK14 LOS, we apply the criteria described by Equations 4.9 and 4.10 to match as many AGD lines between absorption and emission as possible. For KOK14, this does not take into account matching with gas structures along the LOS as described in Section 4.5, in the interest of eliminating as many biases as possible in our comparison between the matching statistics of KOK14 and 21-SPONGE (which does not have LOS density and temperature information for defining structures).

Figure 4.8 displays the matches between H I absorption and H I emission for a set of

5 example observed LOS from 21-SPONGE. All components fitted by AGD to each LOS are shown in dashed black, and the components which satisfy the matching criteria described by Equations 4.9 and 4.10 are shown in colors.

4.6.1 Number of components along LOS

In Table 4.1 we list the total number of AGD components in each H I absorption and emission dataset, as well as the mean and standard deviation of the number of components (N_{AGD}) per LOS. In addition, we list the total number of matched components between absorption and emission, and the mean and standard deviation number of matches per LOS.

From Table 4.1, the mean value of N_{AGD} is more than a factor of two greater for the observed 21-SPONGE H I absorption and emission spectra than KOK14, despite large scatter. Figure 4.9 displays histograms of N_{AGD} for 21-SPONGE (top left panel) and KOK14 (bottom left panel) absorption (black solid) and emission (orange dashed) components. In agreement with the statistics shown in Table 4.1, the maximum number of components fitted to 21-SPONGE

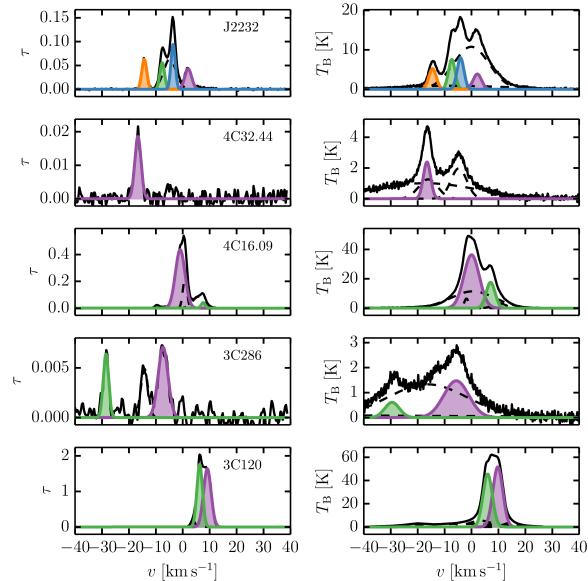


Figure 4.8 Example LOS H I absorption spectra (left) and corresponding H I emission spectra (right) from 21-SPONGE, with AGD-fitted Gaussian decompositions overlaid (dashed black). Components which match between absorption (left) and emission (right) according to Equations 4.9 and 4.10 are indicated by matching colors between the left and right columns.

absorption (12) is a factor of two higher than KOK14 (6).

In the right panels of Figure 4.9, we account for the effect of different viewing angles by multiplying N_{AGD} by $\sin |b|$ for all 21-SPONGE and KOK14 components. This quantity is the effective number of components in the vertical direction of the simulated or observed volume. Since $N_{\text{AGD}} \times \sin |b|$ is still larger for the 21-SPONGE than KOK14, it suggests that the discrepancy between is not a $\sin |b|$ effect.

The larger number of components found in the 21-SPONGE spectra, after correcting for observing angle, indicates that the velocity range used to produce the KOK14 synthetic spectra is smaller than what is sampled by observations. The KOK13 simulations are known to have a relatively low vertical velocity dispersion ($\sim 5 - 7 \text{ km s}^{-1}$), somewhat smaller than observed values. The lower velocity dispersion also yields a scale height somewhat smaller than observations. We will return to this effect in Section 6.2.2. In more recent simulations (Kim & Ostriker 2016) with a better treatment for supernovae, velocity dispersions are in fact larger, and it will be interesting to test whether this will lead to an increase in the number of AGD features per LOS.

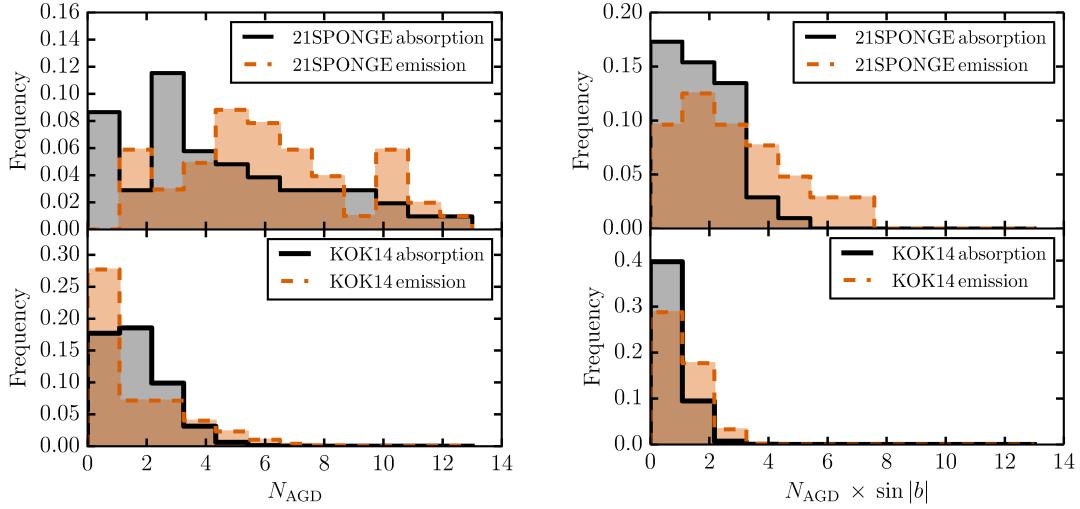


Figure 4.9 Left: Histograms displaying number of AGD lines (N_{AGD}) fit to H I absorption (black) and emission (orange) observations from 21-SPONGE (top) and synthetic observations by KOK14 (bottom). Right: Histograms displaying number of AGD lines per unit path length in the “vertical” direction ($N_{\text{AGD}} \times \sin |b|$) for H I absorption (black) and emission (orange) observations from 21-SPONGE (top) and synthetic observations by KOK14 (bottom).

We emphasize that the results shown in Figure 4.9 are derived from identical implementation of AGD to real and synthetic H I spectra, and thus the comparison is unaffected by biases introduced in spectral line analysis. Therefore, although the caveats described above are known from external analysis of the KOK13 simulations, Figure 4.9 suggests that N_{AGD} reflects the total velocity range and path length.

Furthermore, from Table 4.1, although the number of matches per LOS is consistent with the number of fitted lines per LOS (N_{AGD}) in KOK14, there are comparatively fewer matches per LOS than N_{AGD} in 21-SPONGE. This difference likely comes from the so-called mismatch of angular resolution. In 21-SPONGE, the angular resolution of the H I absorption measurements is determined by the size of the background source, not the telescope beam (and is therefore $< 1 - 40''$). However, the H I emission spectrum has an angular resolution of $3.5'$. Therefore, the H I emission spectrum may not sample the same structures seen in absorption, especially if there is significant emission structure on angular scales below the resolution limit. This mismatch complicates the matching process and causes a larger attrition rate for observations. Simulations, on the other hand, do not suffer from this problem, as emission and absorption are derived using the same angular resolution. In the future, we plan to quantify this effect by smoothing simulated spectra, and we further compare the 21-SPONGE and KOK14 emission properties in detail in Section 6.3.

4.6.2 Properties of H I absorption components

To compare 21-SPONGE and KOK14, Figure 4.10 displays cumulative distribution functions (CDFs) of Gaussian parameters fitted by AGD to H I absorption spectra observed by 21-SPONGE (black) and simulated by KOK14 with (blue) and without (orange) the WF effect, normalized by the total number of components (see Table 1). These parameters include amplitude (τ_0 ; left panel), FWHM (Δv_0 in km s^{-1} ; center panel) and mean velocity (v_0 in km s^{-1} ; right panel). For each dataset, we also plot the CDFs of 1000 bootstrapped samples (shown in lighter-shaded colors according to the legend) to illustrate the effect of sample size and outliers

on the shape of the CDF.

4.6.2.1 Comparison with previous studies

We include the results of the by-hand Gaussian decomposition of the first 31/52 21-SPONGE sources (dashed purple; Murray et al. 2015, “DR1”) and the Millennium Arecibo 21cm Absorption Line Survey (dashed green; HT03) in Figure 4.10. With lower sensitivity in optical depth, the HT03 distribution contains fewer $\tau_0 < 10^{-2}$ components than are found in the 21-SPONGE or KOK14 AGD decompositions. However, the 21-SPONGE DR1 τ_0 distribution agrees very well with the 21-SPONGE AGD distribution, which indicates that although the AGD algorithm was trained using component parameters from HT03, it is successfully able to recover lower- τ amplitudes found in the higher-sensitivity 21-SPONGE and KOK14 spectra. This agreement was also noted in the comparison between by-hand and AGD analysis of a subset of the 21-SPONGE sample (Lindner et al. 2015). In addition, the 21-SPONGE AGD Δv_0 distribution agrees very well with DR1 and HT03, indicating that for a wide range in optical depth sensitivity, a similar range in Gaussian spectral line widths can be recovered.

4.6.2.2 Influence of local box simulations

From the righthand panel of Figure 4.10, the observed 21-SPONGE (AGD and DR1) and HT03 absolute mean velocities ($|v_0|$) agree very well. However, the KOK14 spectra are dominated by components with $v_0 < 10 \text{ km s}^{-1}$. This difference may be caused by the fact that the KOK14 spectra are constructed with a limited path length ($s < 3 \text{ kpc}$) based on their local box simulations, in which the sources of H I lines are limited by nearby gas with small variations of galactic rotation velocity. The limited range of v_0 in KOK14 spectra causes more components to have similar central velocity.

To test the influence of the local box and lack of global rotation effects in the simulation, we consider the effect of latitude on the matching statistics. For high latitude LOS ($|b| > 50^\circ$) in KOK14, the number of absorption fits, emission fits and matches per LOS are consistent with the

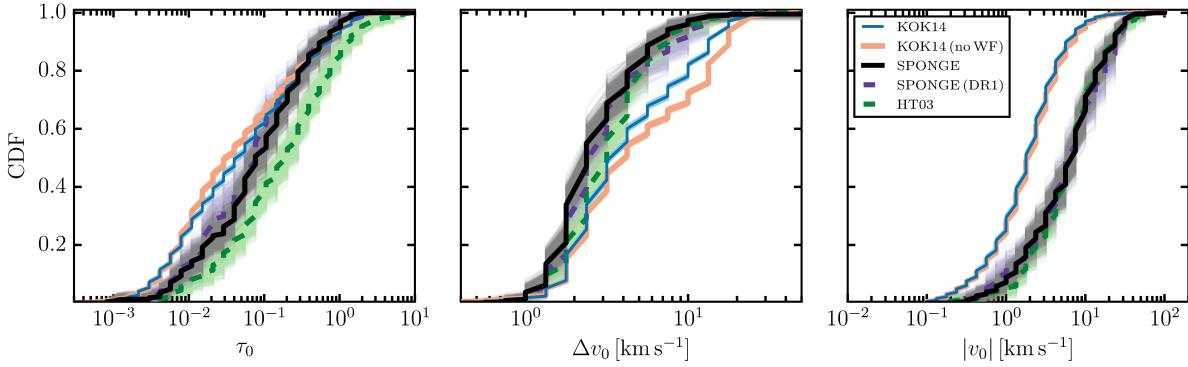


Figure 4.10 Cumulative distribution functions (CDFs) of Gaussian parameters, including optical depth amplitude (τ_0), FWHM (Δv_0) and absolute mean velocity ($|v_0|$), of the components fitted by AGD to observed (21-SPONGE), synthetic (KOK14) H I absorption spectra, including previous by-hand results from 21-SPONGE (purple, DR1; Murray et al. 2015) and HT03 (green) for comparison. For each dataset, we plot the CDFs of 1000 bootstrapped samples (shown in lighter-shaded colors according to the legend) to illustrate the effect of sample size and outliers on the CDF.

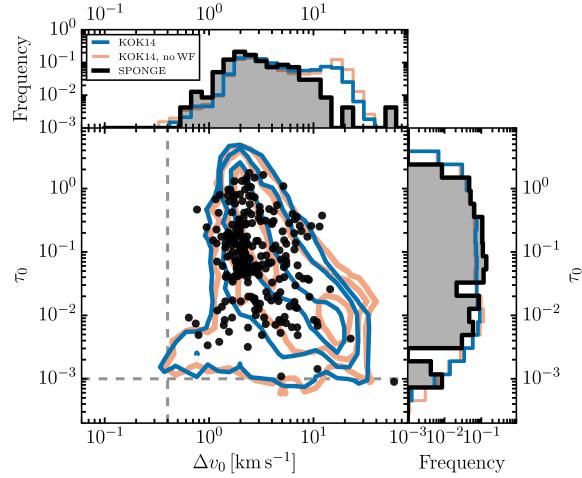


Figure 4.11 Parameters of Gaussian components (Δv_0 , τ_0) fitted by AGD to H I absorption spectra from 21-SPONGE (black) and KOK14 with the WF effect (blue) and without the WF effect (orange). Contours indicate the 1, 2, and 3 σ limits for the KOK14 distributions. Marginal histograms display the same results according to the legend.

full KOK14 sample (i.e., all latitudes). However, for 21-SPONGE, at high latitudes ($|b| > 50^\circ$) there are 1.9 ± 1.1 absorption fits per LOS, 4.4 ± 1.7 emission fits per LOS, and 0.9 ± 0.9 matches per LOS, which are more consistent with the full KOK14 sample than with the full 21-SPONGE sample (c.f., Table 4.1). With increasing latitude, the effect of galactic rotation on spectral line properties declines. The consistency in matching statistics between 21-SPONGE and KOK14 at high latitudes suggests that simulated Galactic rotation dynamics play an important role in the differences between observed and simulated spectral line properties at low latitudes. Therefore, the Gaussian fitting process and match between emission and absorption (Equations 4.9 and 4.10) in KOK14 likely suffer from more completeness issues than in observed 21-SPONGE spectra at low latitudes. This also suggests that our completeness statistics in Section 4.5.3 will be improved in future simulations, in which a larger box, and more realistic supernova feedback are implemented to push gas to higher scale heights and a wider range in velocity.

4.6.2.3 Minimum CNM temperature?

To highlight the comparison between 21-SPONGE and KOK14 we plot τ_0 vs. Δv_0 in Figure 4.11, including marginal histograms of both parameters. The median 1σ sensitivity limit in τ is indicated by the dashed horizontal line ($\sigma_\tau = 10^{-3}$), and the 21-SPONGE velocity resolution of 0.4 km s^{-1} is indicated by the dashed vertical line.

From the top panel of Figure 4.11, we observe a sharp cutoff in $\Delta v_0 \sim 1 - 2 \text{ km s}^{-1}$ in 21-SPONGE and KOK14.¹ If we assume a limiting line width of $\sim 1 - 2 \text{ km s}^{-1}$, in the case of no turbulent broadening, the corresponding CNM kinetic temperature is $\sim 20 - 30 \text{ K}$, which is also equal to the spin temperature. The fact that 21-SPONGE and KOK14 agree in their lower limit to Δv_0 , together with the fact that the AGD method is a good measure of $T_{s,\text{true}}$ at similar temperatures (c.f., Figure 4), suggest that the simulation and observations have a similar lower limit for the CNM temperature of $\sim 20 - 30 \text{ K}$.

¹We note that the components with $\Delta v_0 < 1 \text{ km s}^{-1}$ in 21-SPONGE and KOK14 are likely spurious fits, given that the accuracy of the AGD decomposition is known to be 80% in absorption (c.f., Section 4.4).

As shown in Figure 4.11, the peak optical depth spans the whole parameter space all the way to our sensitivity limit with no obvious evidence for the existence of a minimum optical depth for the CNM. In addition, only a small fraction, < 10%, of components have $\tau_0 > 1$ in KOK14 and 21-SPONGE.

4.6.2.4 Role of the WF effect

As seen most clearly in the main panel of Figure 4.11, the KOK14 spectra with and without the WF effect include significant populations of components with $\Delta v_0 > 10 \text{ km s}^{-1}$ and $0.001 < \tau_0 < 0.01$. Although this region is located well above the 21-SPONGE median sensitivity in optical depth, we find very few 21-SPONGE components there. In addition, in KOK14, these components are often found without narrow (CNM-like) components superimposed along the same LOS. This is a type of profile not seen in 21-SPONGE observations. There are no observational biases that would prevent us from seeing simple Gaussian line profiles with a peak optical depth of ~ 0.01 and a velocity FWHM of 10 km/sec. In addition, the lack of isolated (devoid of CNM), broad, WNM-like features in observations is supported by additional high-sensitivity H I absorption studies (e.g., HT03, Roy et al. 2013). An example isolated, broad absorption line

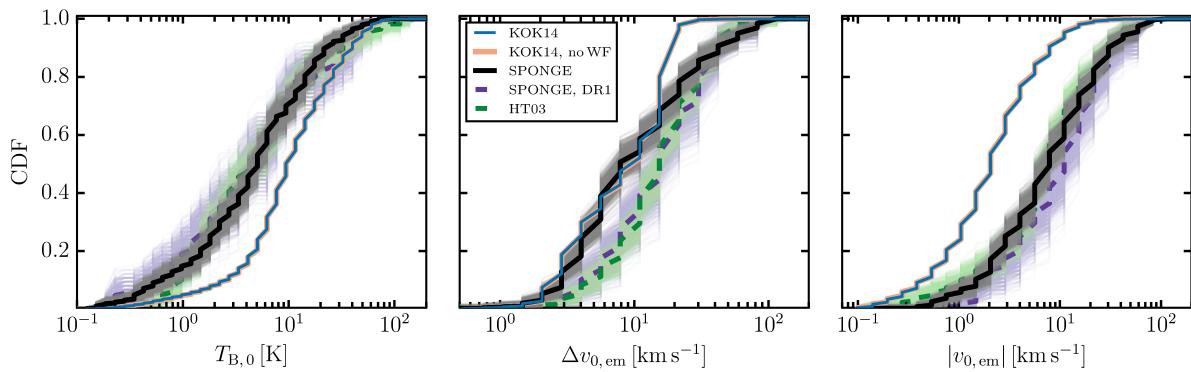


Figure 4.12 Cumulative distribution functions (CDFs) of Gaussian parameters, including brightness temperature amplitude (T_B), FWHM (Δv_0) and mean velocity (v_0), of the components fitted to observed (21-SPONGE) and synthetic (KOK14) H I emission spectra, including previous by-hand results from 21-SPONGE (purple, DR1; Murray et al. 2015) and HT03 (green) for comparison. For each dataset, we plot the CDFs of 1000 bootstrapped samples drawn from the full sample with replacement (shown in lighter-shaded colors according to the legend) to illustrate the effect of sample size and outliers on the CDF.

from KOK14 is shown in the top row of Figure 4.2.

We note that these broad, low- τ features appear in the synthetic KOK14 spectra regardless of our treatment of the WF effect, although agreement with observations is somewhat improved when this is included (c.f., Figure 11). The origin of these low- τ_0 components is not well understood and future comparisons with synthetic 21 cm profiles from simulations will explore the effect of a more realistic feedback treatment (e.g., Kim & Ostriker 2016). The present results already suggest that H I absorption spectra may be able to provide discriminating tests of the input physics in simulations.

4.6.3 Properties of H I emission components

Before comparing the properties of H I emission components in the same manner as in our comparison of absorption components in Section 4.6.2, we emphasize again that the 21-SPONGE and KOK14 emission profiles are derived from different angular scales. The angular resolution of the KOK14 emission spectra is the same as for the absorption lines. However, the 21-SPONGE emission spectra were observed with the Arecibo radio telescope, with a $\sim 3.5'$ beam at 21 cm

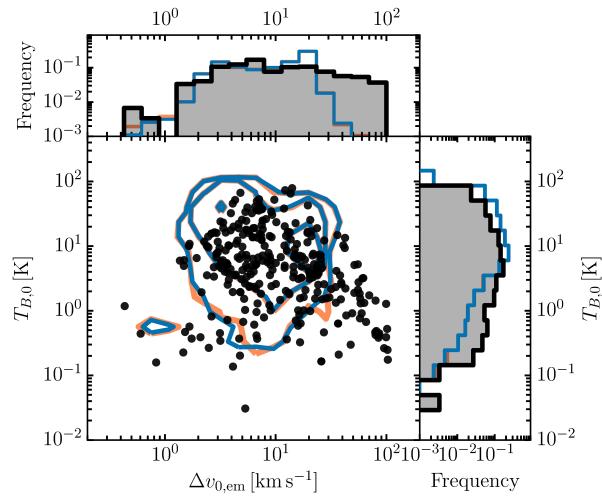


Figure 4.13 Parameters of Gaussian components (FWHM, $T_{B,0}$) fitted by AGD to H I emission spectra from 21-SPONGE (black) and KOK14 with the WF effect (blue) and without the WF effect (orange). Contours indicate the 1, 2, and 3σ limits for the KOK14 distributions. Marginal histograms display the same results according to the legend.

derived from off-target positions, and therefore have different angular resolution than the 21-SPONGE VLA absorption spectra. In the case of the KOK14 spectra, which are derived from a simulation with a physical resolution of 2 pc, the path length would need to be longer than ~ 2 kpc to achieve better angular resolution than 21-SPONGE. Therefore, for shorter LOS at predominately higher latitudes, the angular resolution of synthetic emission spectra is actually worse than 21-SPONGE observations. We stress that absorption lines do not have the resolution problem and this also the reason why they are more suitable for comparison with simulations.

Figure 4.12 displays CDFs of Gaussian parameters from the fits to H I emission spectra observed by 21-SPONGE (black), simulated by KOK14 with the WF effect (blue) and without the WF effect (orange), in addition to the results of HT03 (green) and 21-SPONGE DR1 (purple). These parameters include amplitude in brightness temperature ($T_{B,0}$ in K; left), FWHM ($\Delta v_{0,\text{em}}$ in km s^{-1} ; center) and mean velocity ($v_{0,\text{em}}$ in km s^{-1} ; right). To illustrate the comparison between 21-SPONGE and KOK14 further, we display $T_{B,0}$ vs. $\Delta v_{0,\text{em}}$ for 21-SPONGE and KOK14 in Figure 4.13, with marginal histograms for both parameters. We note that the WF effect does not make a difference to components fitted to H I emission spectra (i.e., orange and blue lines are indistinguishable in Figure 4.12). This indicates that future testing of the implementation of the WF effect should use absorption, rather than emission spectra.

4.6.3.1 Boundaries in brightness temperature

In the left panel of Figure 4.12 (as well as Figure 4.13), the amplitudes fitted by AGD to 21-SPONGE agree well with 21-SPONGE DR1 and HT03. All three datasets were obtained using the Arecibo radio telescope, and therefore they have similar angular resolution. However, all three observed distributions are shifted to slightly lower amplitude in brightness temperature ($T_{B,0} < 10$ K) relative to KOK14 in Figure 4.12. For the KOK14 LOS with lower effective angular resolution than 21-SPONGE, the synthetic brightness temperature spectrum averages any simulated emission over larger solid angles, and therefore the KOK14 $T_{B,0}$ should tend to be smaller than the 21-SPONGE values derived from smaller angular scales. However, we observe

more low- $T_{B,0}$ components in 21-SPONGE than KOK14.

The slight excess of components with high $T_{B,0}$ and low Δv_0 in KOK14 may be caused by the lack of chemistry and H I-to-H₂ transition in the simulation. Furthermore, as noted previously, the relative lack of high- Δv_0 components in KOK14 may be partly attributed to the reduced velocity dispersion in the simulations compared to observations. Our observations may even suggest an upper-limit on $T_{B,0}$ of $\sim 60 - 70$ K. In addition, on the low-end of the distribution the absence of a large filling-factor ($\sim 50\%$) hot ($T \sim 10^5 - 10^7$ K) medium in the simulations produces too much neutral gas per LOS, and therefore too few LOS with $T_{B,0} <$ a few K. The presence of a hot medium occupying a large fraction of the volume would also tend to reduce the incidence of detectable low- τ_0 features, since many of the LOS without CNM would be primarily hot rather than primarily warm medium.

4.6.3.2 Broad emission components

In the middle panel of Figure 4.12, the 21-SPONGE and KOK14 $\Delta v_{0,\text{em}}$ distributions agree very well below $\Delta v_{0,\text{em}} \sim 10 \text{ km s}^{-1}$. However, the 21-SPONGE AGD, DR1 and HT03 distributions extend to higher values of $\Delta v_{0,\text{em}}$ than KOK14. This is especially noticeable in Figure 4.13 where 21-SPONGE components with large $\Delta v_{0,\text{em}}$ and small $T_{B,0}$ form a prominent tail of the distribution. As discussed previously, stray radiation the 21-SPONGE H I emission observations have not been corrected for stray radiation. which would appear in the form of weak and broad spectral components. To test how many 21-SPONGE components could be affected by stray radiation, we extracted H I emission spectra from the stray-radiation-corrected LAB survey (Kalberla et al. 2005) at the positions of our sources and implemented AGD in same manner to decompose the LAB spectra into Gaussian components. Consequently, LAB data contain many components with $\Delta v_{0,\text{em}} \sim 10 - 70 \text{ km s}^{-1}$ (c.f., Appendix 4.7). However, components with $\Delta v_{0,\text{em}} \gtrsim 70 \text{ km s}^{-1}$ are not seen in LAB data, and therefore those components in 21-SPONGE are likely caused by stray radiation. There are 9 such components and they are all located in the large $\Delta v_{0,\text{em}}$ and small $T_{B,0}$ tail. Peek et al. (2011), did a similar comparison between LAB

and Arecibo data and concluded that the effect is unlikely to exceed 500 mK, in agreement with our conclusion that the broadest 21-SPONGE components with $T_{B,0} < 0.5$ K are likely caused by stray radiation.

This leaves a population of broad ($\Delta v_{0,\text{em}} \gtrsim 20 - 70 \text{ km s}^{-1}$), shallow ($T_{B,0} < 10$ K) components that are prominent in 21-SPONGE but absent in KOK14. Possibly a higher velocity dispersion and the inclusion of Galactic fountain in the simulation (see e.g., Kim & Ostriker 2016) may be able to reproduce such broad lines. Alternatively, this could signify the presence of the WNM at temperature > 4000 K. For example, in Chapter 3 we presented a residual absorption component with a width of 50 km/s and $T_s \sim 7000$ K by stacking the first third of the 21-SPONGE data. Further work is clearly needed to understand the origin of such broad emission components, as well as to remove stray radiation from Arecibo spectra.

Finally, in the right panel of Figure 4.12, all observed H I emission spectra have absolute mean velocities up to $\sim 50 \text{ km s}^{-1}$, whereas the KOK14 mean velocities appear limited to $v_{0,\text{em}} < 10 \text{ km s}^{-1}$ (c.f., right panel Figure 4.10). In addition to the nature of the local box reducing $|v_{0,\text{em}}|$, the KOK13 simulations did not include a Galactic fountain of WNM with velocities up to tens of km s^{-1} . The inclusion of this important mechanism may produce relatively more structures with larger $|v_{0,\text{em}}|$ and $\Delta v_{0,\text{em}}$ and improve similarities between the 21-SPONGE and simulation results.

4.6.4 Observed Spin Temperature

We now compare distribution functions of the inferred spin temperature, from 21-SPONGE spectra and KOK14 synthetic spectra. As we have shown in Section 4.5.4.1, for the majority of cases the inferred spin temperature using our AGD and radiative transfer approach, $T_{s,\text{AGD}}$, is in agreement with the true simulated temperature, $T_{s,\text{true}}$. We also discussed how at > 400 K, $T_{s,\text{AGD}}$ over-estimates $T_{s,\text{true}}$. However, this bias will affect both observations and simulations in the same way, as we apply the same AGD fit and radiative transfer method to KOK14 and 21-SPONGE. In addition, our main focus in this paper is on the shape of distribution functions,

not the exact fractions. In future work, we will focus on the fractions of H I in CNM, WNM and unstable phases using updated simulations.

Figure 4.14 displays histograms and CDFs of $T_{s,\text{AGD}}$ (Equation 4.13) for 21-SPONGE (black) and KOK14 with (blue) and without (orange) the WF effect. The observed and simulated $T_{s,\text{AGD}}$ distributions follow each other well until $T_{s,\text{AGD}} \sim 400 - 500$ K, when they start to diverge. Although the 21-SPONGE observations appear to have a higher relative fraction of components at low $T_{s,\text{AGD}} \sim 20 - 30$ K, these bins are determined by small number of components, and the fractions agree within uncertainties as illustrated by the bootstrapped samples shown in the CDF

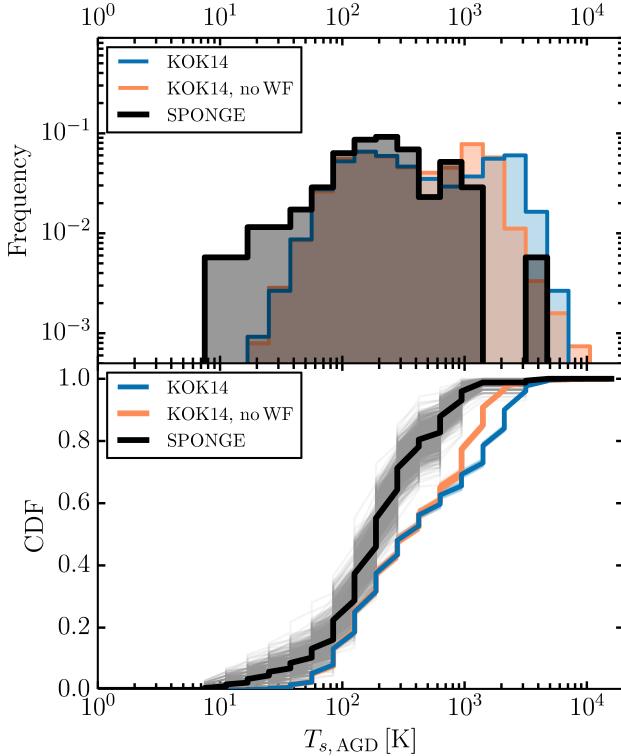


Figure 4.14 Top: histograms of $T_{s,\text{AGD}}$ (Equation 4.13) for all components which “match” between H I emission and absorption (Equations 4.9 and 4.10) following the AGD analysis of 52 21-SPONGE H I spectral pairs (thick, solid black) and 9355 KOK14 H I spectral pairs with the WF effect (thin, solid blue) and without the WF effect (thin, solid orange). Bottom: CDFs of the same results, with 1000 bootstrapped samples drawn from the full sample with replacement (shown in lighter-shaded colors according to the legend) to illustrate the effect of sample size and outliers on the CDF.

panel. As discussed in Section 6.2, the observations and simulations display consistent cutoff in CNM line width, and Figure 4.14 indicates that the fractions of material at corresponding spin temperatures of 20 – 30 K are consistent. HT03 detected a similar population of cold CNM components with $\lesssim 20$ K (17% by number, and 4% by mass), and suggested that this is evidence for the absence of photoelectric heating by dust (Wolfire et al. 1995). We note that the simulations have uniform heating throughout and constant metallicity, while in reality the photoelectric heating would be reduced in high-column regions.

However, the KOK14 spectra show more H I with $T_{s,\text{AGD}} = 300\text{--}3000$ K than 21-SPONGE. To illustrate the types of components with these temperatures, in Figure 4.15 we reproduce Figure 11, including the total contours with WF from KOK14 and all matched components from KOK14 (small circles) and 21-SPONGE (large circles with black outlines), colored by $T_{s,\text{AGD}}$. The AGD components with $T_{s,\text{AGD}} \sim 300\text{--}3000$ K in the KOK14 have lowest optical depth and the largest FWHM. These are exactly the components discussed in Section 4.6.2 to have simple, broad profiles without overlapping CNM components – a type of spectral profile which is not found often in 21-SPONGE.

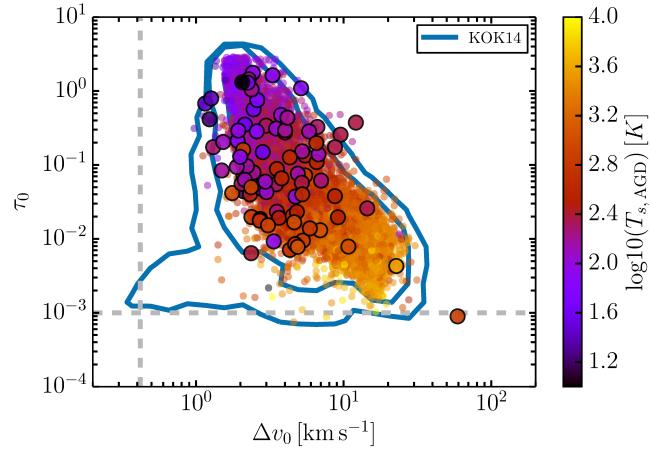


Figure 4.15 AGD absorption properties (τ_0 , Δv_0) for all components fitted to KOK14 with the WF effect (blue contours), and all components which match between absorption and emission according to Equations 4.9 and 4.10 for KOK14 (small circles) and 21-SPONGE (large circles with black outlines), colored by the AGD spin temperature ($T_{s,\text{AGD}}$).

Are the differences shown in Figures 4.14 and 4.15 caused by observational sensitivity? 21-SPONGE was designed with goal of detecting warm H I in absorption and has excellent optical depth sensitivity (RMS noise $\sigma_\tau < 10^{-3}$ per channel; Chapter 2). To estimate the temperatures we are sensitive to in 21-SPONGE observations, we assume a WNM column density of a few $\times 10^{20}$ (e.g., Stanimirović et al. 2014, , KOK14), a FWHM of $10 - 20 \text{ km s}^{-1}$ and a conservative RMS sensitivity in optical depth of $10^{-2} - 3 \times 10^{-3}$ (per 0.4 km/s velocity channels), which results in $T_s \sim 1000 - 6000 \text{ K}$. This is the range of what is expected by KOK14 and also Liszt (2001). Therefore, 21-SPONGE has the observational sensitivity to detect spectral components with $T_s > 1000 \text{ K}$. In Section 4.5.4.1, we demonstrated that our AGD and radiative transfer implementation are not biased against WNM, and in fact above $\sim 400 \text{ K}$ tends to over-estimate the

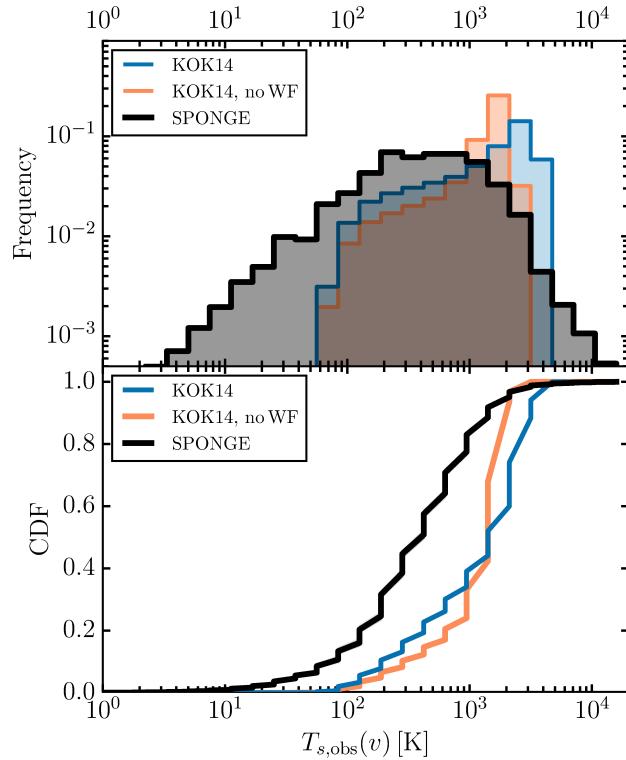


Figure 4.16 Top: histograms of per-channel spin temperature ($T_s(v)$) for all 52 21-SPONGE H I spectral pairs (thick, solid black) and 9355 KOK14 H I spectral pairs with the WF effect (thin, solid blue) and without the WF effect (thin, solid orange). Bottom: CDFs of the same results.

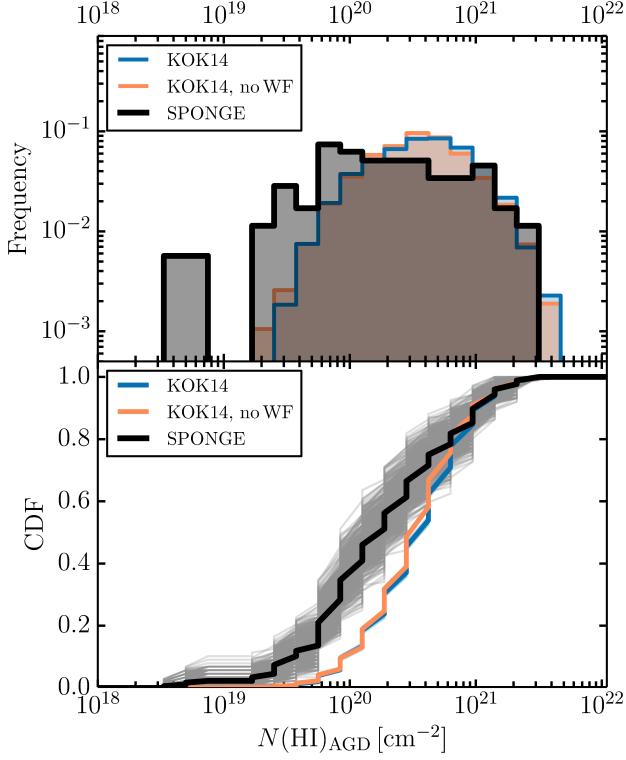


Figure 4.17 Top: histograms of $N(\text{HI})_{\text{AGD}}$ (Equation 4.16) for all components which “match” between H I emission and absorption (Equations 4.9 and 4.10) following the AGD analysis of 52 21-SPONGE H I spectral pairs (thick, solid black) and 9355 KOK14 H I spectral pairs with the WF effect (thin, solid blue) and without the WF effect (thin, solid orange). Bottom: CDFs of the same results, with 1000 bootstrapped samples drawn from the full sample with replacement (shown in lighter-shaded colors according to the legend) to illustrate the effect of sample size and outliers on the CDF.

true spin temperature. Therefore, we infer that the lack of observed components with $T_s > 1000$ K is not affected by sensitivity or analysis method. In addition, we emphasize again that the $T_s > 1000$ K components in simulations have isolated, simple spectral profiles that we do not find in observations.

To further test the effect of observational sensitivity on the inferred spin temperature, in Chapter 3 we investigated methods for improving the 21-SPONGE sensitivity to broad, shallow WNM-like absorption with high temperature. We stacked residual H I absorption lines of a subset of sources following by-hand Gaussian decomposition and found a residual H I absorption signal

at 5σ significance with an inferred excitation temperature of $T_s = 7200^{+1800}_{-1200}$ K, with a FWHM of 50 km s^{-1} , and an H I column density of $2 \times 10^{20} \text{ cm}^{-2}$ (Chapter 3). This temperature is higher than analytical predictions for collisional excitation of H I, and indicated that additional H I excitation mechanisms (e.g., the WF effect) may be more important for coupling the hyperfine spin states of H I to the local thermodynamic temperature than previously thought.

If the spin temperature of the WNM is actually $T_s \sim 7000$ K (Chapter 3), rather than $2000 < T_s < 4000$ K as inferred currently from analytical models of collisional excitation (e.g., Liszt 2001), then it is possible that individual 21-SPONGE spectra may still lack sensitivity for detecting the WNM. Even more sensitive observations, as well as further stacking analysis will be essential from an observational standpoint. This result shows that understanding the implementation of the WF effect in numerical simulations, and in particular the highly uncertain n_α , its spatial variations across the Milky Way, as well as the effect of turbulence, is essential to reconcile H I observations and theory.

4.6.4.1 Per-channel T_s

Instead of using Gaussian-based temperature estimates, KOK14 considered per-channel spin temperature in their analysis of simulated H I properties. They found that this quantity agrees well with the true per-channel temperature extremely well (within a factor of 1.5) for all channels with $\tau \lesssim 1$ (KOK14). Several observational studies have also used per-channel temperature estimates to analyze H I phases (e.g., Roy et al. 2013). To compare with their results, we derive per-channel spin temperature, $T_{s,\text{obs}}(v)$, by applying an equation similar to Equation 4.12 to the full $\tau(v)$ and $T_B(v)$ spectra from 21-SPONGE and KOK14, where,

$$T_{s,\text{obs}}(v) = \frac{T_B(v)}{1 - e^{-\tau(v)}}. \quad (4.18)$$

For each LOS, we compute $T_{s,\text{obs}}(v)$ for only those channels with optical depths greater than 3×10^{-3} , to conservatively exclude all channels with low S/N.

In Figure 4.16, we display histograms and CDFs of $T_{s,\text{obs}}(v)$ for 21-SPONGE (black) and

KOK14 with (blue) and without (orange) the WF effect. There is a similar discrepancy between observations and simulations in Figure 4.16 as seen in Figure 4.14. The KOK14 distributions are shifted to higher temperatures, while 21-SPONGE spectra contain more channels with low temperature ($< 200\text{ K}$). In addition, the 21-SPONGE CDF has a more gradual rise, while simulated data show an abrupt jump near $T_{s,\text{obs}}(v) \sim 1000\text{ K}$ suggesting that most simulated LOSs are dominated by higher $T_{s,\text{obs}}(v)$ derived from spectral channels without detectable absorption. It is important to keep in mind, however, that observed and simulated spectra probe different LOS lengths, as discussed in Section 4.6.2, which could affect the shape of per-channel CDFs. In addition, the absence of a hot medium in the present simulations may lead to too many LOS with $T_s \sim 1000\text{ K}$.

4.6.5 Observed Column Density

As a further benefit of AGD analysis, by resolving the properties of individual spectral components along each LOS, we can analyze the column densities of individual gas structures in contrast with the total LOS column density. In Figure 4.17, we display histograms and CDFs of $N(\text{H}\text{I})_{\text{AGD}}$ for individual matched spectral components from the KOK14 (blue) and 21-SPONGE (thick black) H I emission and absorption spectral pairs. The column density distributions shown in Figure 4.17 agree well at high- $N(\text{H}\text{I})$ ($> 10^{20}\text{ cm}^{-2}$), although the 21-SPONGE distribution extends further below $N(\text{H}\text{I}) = 10^{19}\text{ cm}^{-2}$.

The absence of low-column lines in the KOK14 spectra may be caused by insufficient angular resolution for detecting small CNM features, or, as has been mentioned throughout, the absence of a hot medium in the KOK13 simulations which would serve to reduce the observed column densities of the matched lines. The discrepancy around 10^{20} cm^{-2} is likely caused by the discrepancy in $T_{s,\text{AGD}}$ discussed above in Section 4.6.4. It is interesting to note that the application of the WF effect does not significantly affect the $N(\text{H}\text{I})_{\text{AGD}}$ distribution. The WF effect influences the optical depth and spin temperature in opposite ways, i.e., increases T_s and decreases τ . These quantities are the main ingredients of $N(\text{H}\text{I})_{\text{AGD}}$ (c.f., Equation 4.16), and

therefore the two results of the WF effect may cancel each other out when computing component-based column density. We conclude that the complexity of factors incorporated into the H I column density make it a less useful tool for isolating the importance of the WF effect.

4.7 Summary and Conclusions

Detailed comparisons between observations and simulations are crucial for understanding the physics behind the observed properties of the ISM. Armed with synthetic 21 cm emission and absorption profile data created from the 3D hydrodynamical simulations from KOK14 and high-sensitivity H I observations from 21-SPONGE, we address two main questions: (1) how well do H I spectral lines and our analysis methods recover simulated properties of interstellar gas structures? (2) how do simulated H I spectra compare with real observations? To analyze 9355 synthetic and 52 real observations in an unbiased and uniform way, we apply the Autonomous Gaussian Decomposition (AGD) algorithm (Lindner et al. 2015) identically to both datasets. With these fits in hand, we compare simulated properties of gas structures along each LOS with observed properties of the Gaussian components.

We summarize the main results:

1. For gas structures defined by peaks in n/T_s along random LOS in the KOK13 simulations, Gaussian fits by AGD to synthetic H I absorption lines are able to recover gas structures successfully (Figure 4.3). The recovery completeness (Equation 4.11) is 99% for high-latitude LOS ($|b| > 50^\circ$), 67% for mid-latitude LOS ($20^\circ < |b| < 50^\circ$) and 53% for low-latitude LOS ($0^\circ < |b| < 20^\circ$). The completeness declines with decreasing latitude because the LOS complexity is highest at the lowest latitudes. When these structures are matched to spectral line components in both H I absorption and emission, the completeness is 83%, 38% and 29% for high, mid and low latitudes respectively. The decline in recovery completeness when matches between gas structures and both H I absorption and emission components are required reflects the difficulty in associating unambiguous spectral features in the presence

of line blending and turbulence.

2. We use AGD fits to synthetic lines and simple radiative transfer to compute observational estimates of spin temperature ($T_{s,\text{AGD}}$) and column density ($N(\text{H}\text{i})_{\text{AGD}}$) for matched pairs of H I absorption and emission lines. We compare these estimates with the simulated spin temperatures ($T_{s,\text{true}}$) and column densities ($N(\text{H}\text{i})_{\text{true}}$) of corresponding structures in the simulation. The observed and simulated spin temperatures agree within a factor of 2 for the majority of structures (68%; Figure 4.4). At high temperatures, $T_{s,\text{AGD}}$ overestimates the $T_{s,\text{true}}$ due to velocity offsets between H I absorption and emission lines caused by turbulent motions (Figure 4.5).

The observed and simulated H I column densities also agree well for the majority of structures. However, the scatter is slightly larger than in the case of spin temperature, because $N(\text{H}\text{i})_{\text{AGD}}$ incorporates all uncertainty in $T_{s,\text{AGD}}$ (Figure 4.6). Furthermore, the agreement between inferred and true properties declines at low Galactic latitude and for low- τ components, where LOS-blending components hinder clear associations between emission and absorption spectral lines (Figure 4.7).

Overall, the agreement between temperature and column densities inferred from synthetic spectra and computed from physical conditions in the simulation is encouraging. Future comparisons with next-generation simulations will allow us to construct “correction functions” for observed spin temperature and column density distributions.

3. We find more fitted absorption and emission lines per LOS (N_{AGD}) in the 21-SPONGE observations than the KOK14 synthetic observations (Table 4.1). This difference reflects the fact that the simulated scale heights of the CNM and WNM in the KOK13 simulations are lower than in observations, due to velocity dispersions lower than seen in observations ($\sim 5 - 7 \text{ km s}^{-1}$). These results are derived from identical implementation of AGD to 21-SPONGE and KOK14, and thus the comparison is unaffected by biases introduced by AGD

analysis. The discrepancy reflects the limitations of local box simulations, with a simplified treatment of supernova feedback, in producing realistic synthetic spectral lines.

In addition, there are comparatively fewer matches per LOS in 21-SPONGE than KOK14. The so-called mismatch in angular resolution between 21-SPONGE H I emission ($\sim 3.5'$) and absorption ($1 - 40''$) complicates the matching process described by Equations 4.9 and 4.10. In the future, we plan to quantify this effect and correct the observational results by smoothing simulated spectra.

4. Using AGD, we objectively compare the properties of spectral lines fitted by AGD to H I absorption from 21-SPONGE and KOK14. The 21-SPONGE spectra have a wider range in mean velocity (right panel, Figure 4.10), due to the limited horizontal box size of the simulation. Furthermore, at high Galactic latitudes where the influence of the global effect is weakest, N_{AGD} and matching statistics agree between 21-SPONGE and KOK14. This indicates that simulated Galactic rotation plays an important role in observed H I properties, and improved implementations of global effects will improve the completeness statistics of H I structure recovery, and improve the completeness characterization discussed in Summary points 1, 2 and 3.
5. We find that KOK14 spectra include more low- τ and high- Δv absorption lines than are seen in 21-SPONGE (Figures 4.10 and 4.11), despite being well above the 21-SPONGE sensitivity and resolution limits. These broad spectral lines are often found without narrower, blended lines, which is a profile not seen in observations by 21-SPONGE or previous H I absorption line surveys (e.g., HT03, Roy et al. 2013). These are likely the result of the absence of a hot, large filling-factor gas phase in the KOK13 simulations (which would increase the number of LOS without detectably-absorbing neutral gas), or possibly some aspect of the simple treatment of the WF effect. In particular, we find that excluding the WF effect enhances the population of these discrepant H I absorption components, and suggests that the WF effect is important for realistic spectral line properties. These features are not obvious

in comparisons of integrated or per-channel properties, and reflect the utility of studying velocity-resolved spectral components.

6. We find that H I absorption spectra are more useful probes of ISM physics in comparison with simulations than H I emission spectra. Properties of components fitted to H I emission profiles are affected by angular resolution mismatch and stray radiation, and are not sensitive to the implementation of the WF effect.
7. The AGD-derived spin temperature from KOK14 has more high-temperature gas ($1000 < T_{s,\text{AGD}} < 4000$ K) relative to 21-SPONGE. The AGD method is not biased against high temperatures, and the 21-SPONGE observations have sufficient sensitivity for detecting gas at similar values. In KOK14, the highest-temperature components have large line widths and low optical depths, which is a type of profile not seen in 21-SPONGE (Figure 4.15). We conclude that the lack of observed components with $T_s > 1000$ K is not affected by sensitivity or the analysis method. We suggest that this gas may have even higher temperatures ($T_s \sim 7000$ K) than what 21-SPONGE is sensitive to and than what the KOK14 implementation of the WF effect allows (i.e., $1000 < T_s < 4000$ K). Again, future work is required to understand the importance of the WF effect in more detail. It is also important to test whether more realistic treatment of star formation feedback, including a hot ISM (see Kim & Ostriker 2016), reduces the incidence of features with $T_{s,\text{AGD}} \sim 1000 - 4000$ K in simulations.

Overall, we are encouraged that the AGD analysis and radiative transfer method presented here is a useful tool for diagnosing important physical conditions within ISM simulations. In the future, we will apply the strategies described here to updated simulations in order to derive correction functions for observational biases in real H I data. Upcoming large H I absorption surveys such as GASKAP at the Australian Square Kilometer Array Pathfinder (ASKAP) telescope (Dickey et al. 2013) will contribute many more sources and improve the observational

statistics. The objective and efficient nature of the AGD analysis strategy presented here is well-suited for future large observed and simulated datasets, and will be important for understanding the balance of CNM and WNM in the local and extragalactic ISM.

This work was supported by the NSF Early Career Development (CAREER) Award AST-1056780. C. E. M. acknowledges support by the National Science Foundation Graduate Research Fellowship and the Wisconsin Space Grant Institution. S. S. thanks the Research Corporation for Science Advancement for their support. The work of E.C.O. and C.-G. Kim was supported by NSF grant AST-1312006. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. The Arecibo Observatory is operated by SRI International under a cooperative agreement with the National Science Foundation (AST-1100968), and in alliance with Ana G. Méndez-Universidad Metropolitana, and the Universities Space Research Association.

References

- Acreman, D. M., Dobbs, C. L., Brunt, C. M., & Douglas, K. A. 2012, MNRAS, 422, 241
- Beaumont, C. N., Offner, S. S. R., Shetty, R., Glover, S. C. O., & Goodman, A. A. 2013, ApJ, 777, 173
- Begum, A., Stanimirović, S., Goss, W. M., et al. 2010, ApJ, 725, 1779
- Carilli, C. L., Dwarakanath, K. S., & Goss, W. M. 1998, ApJ, 502, L79
- Chengalur, J. N., Kanekar, N., & Roy, N. 2013, MNRAS, 432, 3074
- Crovisier, J., Kazes, I., & Aubry, D. 1978, A&AS, 32, 205
- Dickey, J. M., & Lockman, F. J. 1990, ARA&A, 28, 215
- Dickey, J. M., McClure-Griffiths, N. M., Gaensler, B. M., & Green, A. J. 2003, ApJ, 585, 801
- Dickey, J. M., McClure-Griffiths, N., Gibson, S. J., et al. 2013, PASA, 30, e003
- Douglas, K. A., Acreman, D. M., Dobbs, C. L., & Brunt, C. M. 2010, MNRAS, 407, 405
- Draine, B. T. 2011, Physics of the Interstellar and Intergalactic Medium (Princeton University Press)
- Duarte-Cabral, A., Acreman, D. M., Dobbs, C. L., et al. 2015, MNRAS, 447, 2144
- Duarte-Cabral, A., & Dobbs, C. L. 2016, MNRAS, 458, 3667

- Dwarakanath, K. S., Carilli, C. L., & Goss, W. M. 2002, ApJ, 567, 940
- Falgarone, E., Lis, D. C., Phillips, T. G., et al. 1994, ApJ, 436, 728
- Field, G. B. 1959, ApJ, 129, 536
- Field, G. B., Goldsmith, D. W., & Habing, H. J. 1969, ApJ, 155, L149
- Hartmann, D., & Burton, W. B. 1997, *Atlas of Galactic Neutral Hydrogen* (Cambridge University Press)
- Heiles, C., & Troland, T. H. 2003a, ApJS, 145, 329
- . 2003b, ApJ, 586, 1067
- Hennebelle, P., Audit, E., & Miville-Deschénes, M.-A. 2007, A&A, 465, 445
- Kalberla, P. M. W., Burton, W. B., Hartmann, D., et al. 2005, A&A, 440, 775
- Kalberla, P. M. W., McClure-Griffiths, N. M., Pisano, D. J., et al. 2010, A&A, 521, A17
- Kim, C.-G., & Ostriker, E. C. 2016, ArXiv e-prints, arXiv:1612.03918
- Kim, C.-G., Ostriker, E. C., & Kim, W.-T. 2013, ApJ, 776, 1
- . 2014, ApJ, 786, 64
- Lazareff, B. 1975, A&A, 42, 25
- Lindner, R. R., Vera-Ciro, C., Murray, C. E., et al. 2015, AJ, 149, 138
- Liszt, H. 2001, A&A, 371, 698
- McClure-Griffiths, N. M., Green, J. A., Hill, A. S., et al. 2013, ApJ, 770, L4
- McClure-Griffiths, N. M., Pisano, D. J., Calabretta, M. R., et al. 2009, ApJS, 181, 398
- McKee, C. F., & Ostriker, J. P. 1977, ApJ, 218, 148
- Mebold, U., Winnberg, A., Kalberla, P. M. W., & Goss, W. M. 1982, A&A, 115, 223
- Mohan, R., Dwarakanath, K. S., & Srinivasan, G. 2004, Journal of Astrophysics and Astronomy, 25, 185
- Murray, C. E., Lindner, R. R., Stanimirović, S., et al. 2014, ApJ, 781, L41
- Murray, C. E., Stanimirović, S., Goss, W. M., et al. 2015, ApJ, 804, 89
- Peek, J. E. G., Heiles, C., Douglas, K. A., et al. 2011, ApJS, 194, 20
- Pettitt, A. R., Dobbs, C. L., Acreman, D. M., & Price, D. J. 2014, MNRAS, 444, 919
- Roy, N., Kanekar, N., & Chengalur, J. N. 2013, MNRAS, 436, 2366
- Sauty, E., Miville-Deschénes, M.-A., Hennebelle, P., Audit, E., & Schmidt, W. 2014, A&A, 567, A16
- Shetty, R., Glover, S. C., Dullemond, C. P., & Klessen, R. S. 2011, MNRAS, 412, 1686

- Smith, R. J., Glover, S. C. O., Clark, P. C., Klessen, R. S., & Springel, V. 2014, MNRAS, 441, 1628
- Stanimirović, S., & Heiles, C. 2005, ApJ, 631, 371
- Stanimirović, S., Murray, C. E., Lee, M.-Y., Heiles, C., & Miller, J. 2014, ApJ, 793, 132
- 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
- Wouthuysen, S. A. 1952, AJ, 57, 31

To test for the effect of stray radiation on H I emission spectra, we analyze data from the LAB survey (Kalberla et al. 2005). The LAB spectra were measured with a $\sim 35'$ beam, whose shape was carefully modeled in order to remove contamination from stray radiation. We apply the AGD algorithm in the same manner as described in Section 6.3 to LAB spectra in the direction of the 52 21-SPONGE sources. In Figure 4.18, we reproduce Figure 4.13 including the LAB decomposition in red.

From Figure 4.18, the LAB $\Delta v_{0,\text{em}}$ distribution is shifted to larger values relative to KOK14 and 21-SPONGE. As a result of the much larger angular resolution of the LAB survey, the AGD decomposition of LAB data in Figure 4.18 feature fewer narrow velocity components ($\Delta v_{0,\text{em}} < 2 \text{ km s}^{-1}$), likely because the H I emission is smoothed over larger angular scales than in the 21-SPONGE spectra (from Arecibo Observatory, with $\sim 3.5'$ angular resolution).

Furthermore, the LAB decomposition contains large-line width ($\Delta v_{0,\text{em}} > 30 \text{ km s}^{-1}$), low-brightness temperature ($T_{B,0} < 1 \text{ K}$) components similar to 21-SPONGE. However, there is a population of 21-SPONGE components with $\Delta v_{0,\text{em}} \sim 50 - 100 \text{ km s}^{-1}$ outside the LAB distribution. Given that stray radiation has been removed from the LAB spectra, this type of spectral feature may be indicative of stray radiation.

Future work is needed to understand and remove the stray radiation contamination from Arecibo spectra. Kalberla et al. (2010) found that stray radiation contributed up to 35% to some GASS H I spectra, however the effect is typically $< 10\%$ and is not significant at Galactic

latitudes below $\sim 60^\circ$ (e.g., McClure-Griffiths et al. 2009). In Chapter 2, we compared 21-SPONGE emission spectra with the GALFA-H I (Peek et al. 2011) and LAB surveys (Kalberla et al. 2005) and concluded that difference are generally within 3σ uncertainties. In a future paper, we will compare the KOK14 synthetic H I emission spectra and data from the LAB and GALFA-HI surveys objectively using AGD in order to statistically quantify the effect of stray radiation on observational data.

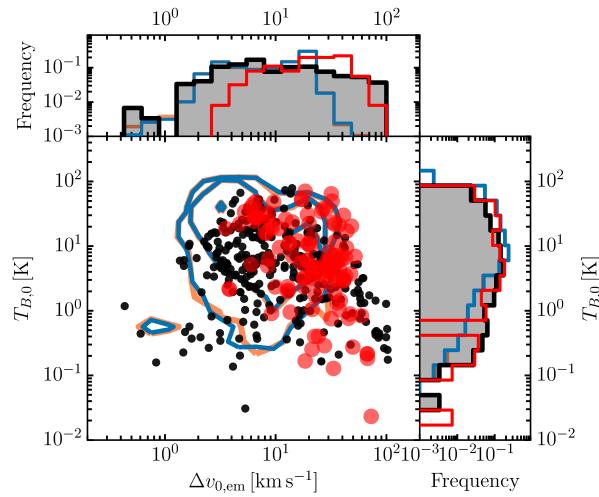


Figure 4.18 Parameters of Gaussian components ($\Delta v_{0,\text{em}}$, $T_{B,0}$) fitted by AGD to H I emission spectra from 21-SPONGE (black) and KOK14 synthetic observations including the WF effect (blue) and without the WF effect (orange). Contours indicate the 1, 2, and 3σ limits for the KOK14 distribution. We also include the AGD decomposition of H I emission spectra in the direction of the 21-SPONGE sources from the Leiden Argentine Bonn (LAB) survey (Kalberla et al. 2005) in red.

Chapter 5

The 21-SPONGE HI Absorption Line Survey II: The mass distribution of Galactic HI as a function of temperature

C. E. Murray, S. Stanimirović, et al. 2017, *in preparation*

Abstract

We present the final data products of 21-cm Spectral Line Observations of Neutral Gas with the EVLA (21-SPONGE): the largest, high-sensitivity survey for Galactic H I absorption at the Karl G. Jansky Very Large Array (VLA). Following 573 hours of VLA time targeting 59 bright, compact continuum sources at mostly high latitudes, 21-SPONGE achieved exceptional optical depth sensitivity ($\sigma_{\tau} < 0.001$ per 0.42 km s^{-1} channels). In combination with matching H I emission from the Arecibo Observatory, we are sensitive to 21 cm emission and absorption from the cold neutral medium (CNM), warm neutral medium (WNM) and thermally unstable medium (UNM). We present a new method for automatically and efficiently decomposing H I emission and absorption line pairs and deriving spin temperature and column density for individual gas structures along the line of sight, building from Heiles & Troland (2003a) and Murray et al. (2017). Via comparison with synthetic H I spectra from numerical simulations, we produce a bias-corrected mass distribution of H I as a function of temperature. We find significant fraction of mass in the CNM ($\sim 68\%$). We prove that with high sensitivity, CNM is detected ubiquitously (only 3/57 non-detections). We find very little evidence for WNM with temperatures in the predicted range of $T_s \sim 2000 - 4000 \text{ K}$, implying that either the WNM is all unstable, or it is found at even higher temperatures. Furthermore, we find a minimum of $\sim 28\%$ thermally-unstable gas by mass in agreement with previous indirect studies based on emission alone. This emphasizes the importance of dynamical and time-dependent physical processes (i.e., turbulence, supernovae), and calls into question the validity of steady-state ISM models.

5.1 Introduction

Shortly following the first astronomical observations of absorption and emission by the 21 cm transition of H I (Muller & Oort 1951; Ewen & Purcell 1951; Hagen et al. 1955), it became clear that H I exists in multiple phases of temperature and density within the interstellar medium (ISM). Simple analysis of observed velocity structure reveal clear differences between H I absorption and emission, including velocity width and extent, caused by variations in temperature and density of the gas along the line of sight (e.g., Clark 1965; Dickey et al. 1978). Theoretical models of steady-state ISM heating and cooling quantified the nature of this phase structure, predicting two thermally-stable phases: the cold neutral medium (CNM) and warm neutral medium (WNM), with density and kinetic temperatures of $(n, T_k) = (5\text{--}120 \text{ cm}^{-3}, 40\text{--}200 \text{ K})$, and $(n, T_k) = (0.03\text{--}1.3 \text{ cm}^{-3}, 4100\text{--}8800 \text{ K})$ respectively (McKee & Ostriker 1977; Wolfire et al. 2003). However, subsequent analysis and numerical simulations determined that time-dependent or dynamical processes such as turbulence and supernovae are likely very important, and will generate a significant amount of thermally unstable gas (UNM) in the intervening parameter space between CNM and WNM, thus throwing into question the validity of the steady-state paradigm of the ISM (e.g., Dalgarno & McCray 1972; Vázquez-Semadeni et al. 2000; Audit & Hennebelle 2005).

However, the physical properties, mass fractions and ionization state of the diffuse neutral gas phases (i.e., WNM and UNM) are particularly sensitive to both macro- and microphysical heating and cooling processes (Heiles & Troland 2003b). These include magnetic wave dissipation (Ferrière et al. 1988), magnetic reconnection (Vishniac & Lazarian 1999), turbulence (e.g., Audit & Hennebelle 2005), supernovae (e.g., Mac Low et al. 2005), diffusion of photons from HII regions, and diffusion of low-energy cosmic rays and X-rays from time-dependent stellar phenomena. As these processes originate from sources on a huge range of physical and temporal scales, understanding them as part of a self-consistent model of the ISM has proven challenging.

Furthermore, observational constraints for the properties of the UNM and WNM have been historically limited by insufficient observational capabilities. To constrain the optical depth and excitation (or spin) temperature of H I — crucial for determining the thermodynamic state of the gas — measurements of both emission and absorption are necessary. Due to their low densities, detecting absorption by the WNM and UNM requires extremely high sensitivity to H I absorption. Only a handful of such measurements exist (Carilli et al. 1998; Dwarakanath et al. 2002).

Most often in the literature, UNM and WNM properties are indirectly estimated from H I emission alone. These studies have argued that a significant fraction of the H I mass is thermally unstable. The Gaussian line widths of decomposed 21 cm emission profiles indicate that $\sim 50\%$ of the detected H I has thermally-unstable kinetic temperatures of ~ 3000 K (e.g., Verschuur & Magnani 1994; Haud & Kalberla 2007). The Millennium Arecibo 21 cm Absorption-Line Survey detected absorption from gas with excitation temperatures of $\sim 10 - 600$ K, and inferred from the emission that $\sim 48\%$ of the remaining material detected in emission alone is thermally unstable (Heiles & Troland 2003b). From a high-sensitivity survey of 35 sources, Roy et al. (2013a) estimated that at most 28% of H I is unstable, in tension with previous results. However this result is based on assumed H I spin temperatures and H I emission data from the Leiden Argentine Bonn (LAB) survey, whose 36' resolution probes much larger scales and thus different H I populations than their sub-arcminute interferometric measurements. Tackling the properties of the diffuse H I and constraining the UNM mass fraction requires expanded samples of absorption lines and careful attention to systematic uncertainties in both data and analysis techniques.

In this chapter, I present the final data products from the largest survey for high-sensitivity Galactic H I absorption to date at the Karl G. Jansky Very Large Array (VLA), titled 21-cm Observations of Neutral Gas with the EVLA (21-SPONGE). Together with H I emission from the Arecibo Observatory, we are uniquely sensitive to all CNM, UNM and WNM temperatures ($10 < T_s < 6000$ K) in the Galactic ISM (Murray et al. 2015). Furthermore, we have developed

an objective, efficient method for characterizing H I structures within simulations and synthetic observations (Murray et al. 2017). This enables us to quantify and correct for the biases associated with our analysis method in the 21-SPONGE results for the first time.

The chapter is organized as follows. In Section 5.2, we discuss the observations and data reduction, referring heavily to the techniques presented in Chapter 2. In Section 5.3 we present the integrated spectral line properties, and in Section 5.4 we discuss our revised Gaussian decomposition and radiative transfer approach, derived from Heiles & Troland (2003a, hereafter HT03) and Chapter 4. In Section 5.5, we present the results of our analysis of 21-SPONGE, including corrections for observational bias via comparison with synthetic spectra from Kim et al. (2014). In Section 5.6 we discuss our results in the context of previous work, and we summarize our conclusions in Section 5.7.

5.2 Data

5.2.1 Observing and Data Reduction

The observing strategy for 21-SPONGE is described in Chapter 2 and summarized here for clarity. Our targets are bright ($S_{1.4\text{ GHz}} > 3\text{ Jy}$), radio continuum sources from the NRAO/VLA Sky Survey (NVSS; Condon et al. 1998) at high Galactic latitude (generally, $|b| > 10^\circ$) with compact angular sizes ($< 1'$) to minimize the complexity of observed H I profiles, and to avoid resolving substantial continuum flux.

For all VLA observations, we use three separate 500 kHz bands with 1.95 kHz channel spacing centered on the H I line (1.42040575 GHz) and ± 1.5 MHz respectively. We use the off-line bands to avoid strong H I lines at Galactic velocities in the directions of our calibrator sources to perform bandpass calibration via frequency switching. Bandpass calibration is of particular importance for 21-SPONGE, as our primary interest is to detect broad, shallow absorption lines associated with high-temperature H I. For details on our bandpass calibration strategies, please see Appendix A.

Table 5.1. VLA Observation Information

Source (name)	RA (J2000) (hh:mm:ss)	Dec (J2000) (dd:mm:ss)	l (°)	b (°)	$S_{1.4\text{ GHz}}$ (Jy)	σ_τ ($\times 10^{-3}$)	τ_{peak}	$\int \tau \text{ dv}$	$N(\text{HI})_{\text{obs}}$ (10^{20} cm^{-2})
J0022	00:22:25.4	+00:14:56.2	107.462	-61.748	3.01	0.8	0.025 ± 0.001	0.15 ± 0.03	3.13 ± 0.07
3C018A	00:40:50.7	+10:03:05.0	118.623	-52.732	4.60	1.5	0.623 ± 0.004	2.39 ± 0.07	7.30 ± 0.02
3C018B	00:40:49.5	+10:03:50.0	118.616	-52.719	—	2.4	0.642 ± 0.006	2.36 ± 0.10	7.28 ± 0.03
3C041A	01:26:44.8	+33:13:02.3	131.379	-29.075	3.71	2.4	0.039 ± 0.004	0.43 ± 0.09	6.11 ± 0.14
3C041B	01:26:43.8	+33:13:21.8	131.374	-29.070	—	3.8	0.057 ± 0.006	0.39 ± 0.14	6.12 ± 0.22
3C48	01:37:41.3	+33:09:35.1	133.963	-28.719	16.02	0.7	0.050 ± 0.001	0.40 ± 0.03	4.99 ± 0.04
4C15.05	02:04:50.4	+15:14:11.0	147.930	-44.043	4.07	0.6	0.086 ± 0.001	0.74 ± 0.03	5.24 ± 0.02
3C78	03:08:26.2	+04:06:39.0	174.858	-44.514	5.75	2.2	1.366 ± 0.006	4.98 ± 0.10	12.82 ± 0.03
4C16.09	03:18:57.8	+16:28:32.7	166.636	-33.596	8.03	0.6	0.539 ± 0.002	3.01 ± 0.03	12.36 ± 0.02
3C111A	04:18:21.3	+38:01:35.8	161.676	-8.820	7.73	1.7	0.888 ± 0.006	10.20 ± 0.10	31.73 ± 0.06
3C111B	04:18:30.4	+38:02:30.4	161.686	-8.788	4.31	1.6	1.150 ± 0.006	10.14 ± 0.10	31.95 ± 0.06
3C111C	04:18:15.5	+38:00:48.2	161.671	-8.843	2.92	3.0	1.125 ± 0.010	11.39 ± 0.17	32.64 ± 0.07
3C120	04:33:11.1	+05:21:15.6	190.373	-27.397	3.44	0.9	2.033 ± 0.003	10.64 ± 0.04	18.40 ± 0.01
3C123A	04:37:04.9	+29:40:10.2	170.584	-11.660	49.73	0.6	1.750 ± 0.002	9.05 ± 0.03	22.89 ± 0.04
3C123B	04:37:04.0	+29:40:28.2	170.578	-11.659	—	0.7	1.783 ± 0.003	8.81 ± 0.04	22.76 ± 0.04
3C132	04:56:43.5	+22:49:16.3	178.862	-12.522	3.43	1.1	1.614 ± 0.005	7.55 ± 0.07	30.07 ± 0.03
3C133	05:02:58.1	+25:16:26.6	177.725	-9.913	5.77	2.7	1.645 ± 0.010	9.13 ± 0.16	34.90 ± 0.05
3C138	05:21:09.9	+16:38:22.1	187.405	-11.343	8.60	1.0	1.115 ± 0.004	5.99 ± 0.06	24.77 ± 0.02
PKS0531	05:34:44.5	+19:27:21.4	186.762	-7.108	7.02	0.5	0.535 ± 0.002	3.33 ± 0.03	30.14 ± 0.03
3C147	05:42:36.1	+49:51:07.2	161.686	10.298	22.88	0.5	0.796 ± 0.001	5.00 ± 0.03	20.06 ± 0.02
3C154	06:13:49.0	+26:04:36.7	185.592	4.003	5.00	0.7	1.704 ± 0.004	14.77 ± 0.06	53.94 ± 0.02
PKS0742	07:45:33.1	+10:11:12.7	209.797	16.592	3.51	0.6	0.011 ± 0.001	0.03 ± 0.03	3.61 ± 0.31
3C225A	09:42:15.3	+13:45:51.3	220.010	44.008	3.34	1.2	0.827 ± 0.002	1.49 ± 0.05	4.24 ± 0.01
3C225B	09:42:15.6	+13:45:49.3	220.011	44.009	—	2.3	0.790 ± 0.004	1.51 ± 0.09	4.22 ± 0.02
3C236	10:06:01.8	+34:54:10.4	190.065	53.980	3.24	0.6	0.003 ± 0.001	0.00 ± 0.03	1.02 ± 1.09
3C237	10:08:00.0	+07:30:16.6	232.117	46.627	6.52	1.0	0.410 ± 0.002	0.68 ± 0.04	2.54 ± 0.02
3C245A	10:42:44.6	+12:03:31.3	233.124	56.300	3.31	1.3	0.016 ± 0.002	0.04 ± 0.05	2.62 ± 0.37
3C245B	10:42:44.3	+12:03:31.6	233.123	56.299	—	4.2	0.024 ± 0.006	0.13 ± 0.15	2.62 ± 0.29
1055+018	10:58:29.6	+01:33:58.8	251.511	52.774	3.22	0.9	0.008 ± 0.001	0.07 ± 0.04	3.37 ± 0.18
3C263.1	11:43:25.1	+22:06:56.1	227.201	73.766	3.13	0.7	0.020 ± 0.001	0.07 ± 0.03	2.16 ± 0.10
3C273	12:29:06.1	+02:03:08.6	289.945	64.359	54.99	0.4	0.026 ± 0.001	0.08 ± 0.02	2.45 ± 0.06
4C32.44	13:26:16.5	+31:54:09.5	67.234	81.048	4.86	0.7	0.020 ± 0.001	0.07 ± 0.03	1.51 ± 0.06
4C25.43	13:30:37.7	+25:09:11.0	22.468	80.988	7.05	1.1	0.004 ± 0.001	0.02 ± 0.04	1.36 ± 0.38
3C286	13:31:08.3	+30:30:33.0	56.524	80.675	14.90	0.4	0.007 ± 0.001	0.08 ± 0.02	1.42 ± 0.03
4C12.50	13:47:33.4	+12:17:24.2	347.223	70.172	5.40	0.9	0.091 ± 0.002	0.32 ± 0.04	2.50 ± 0.03
3C298	14:19:08.2	+06:28:34.8	352.160	60.666	6.10	0.6	0.020 ± 0.001	0.07 ± 0.03	2.48 ± 0.10
UGC09799	15:16:44.5	+07:01:17.8	9.417	50.120	5.50	6.9	0.066 ± 0.011	0.16 ± 0.24	3.43 ± 0.51
4C04.51	15:21:14.4	+04:30:22.0	7.292	47.747	3.93	0.8	0.068 ± 0.001	0.33 ± 0.03	4.62 ± 0.05
3C327.1A	16:04:44.9	+01:17:52.8	12.181	37.006	4.08	3.2	0.505 ± 0.008	2.29 ± 0.13	8.87 ± 0.05
3C327.1B	16:04:45.6	+01:17:47.6	12.182	37.003	—	3.0	0.448 ± 0.008	2.20 ± 0.12	8.82 ± 0.05
PKS1607	16:09:13.3	+26:41:29.0	44.171	46.203	4.91	0.6	0.177 ± 0.001	0.93 ± 0.03	4.52 ± 0.02
J1613	16:13:41.1	+34:12:47.9	55.151	46.379	4.02	1.0	0.005 ± 0.001	0.01 ± 0.04	1.93 ± 1.28
3C345	16:42:58.8	+39:48:37.0	63.455	40.949	7.10	0.9	0.008 ± 0.001	0.00 ± 0.04	1.00 ± 1.96
3C346	16:43:48.6	+17:15:49.3	35.332	35.769	3.66	2.0	0.288 ± 0.004	1.21 ± 0.08	5.93 ± 0.04
3C390	18:45:37.6	+09:53:45.0	41.112	5.773	4.51	1.2	0.157 ± 0.004	2.68 ± 0.07	29.93 ± 0.10
4C33.48	19:24:17.5	+33:29:29.7	66.389	8.371	3.77	2.9	0.408 ± 0.007	2.57 ± 0.13	15.78 ± 0.08
3C409A	20:14:27.5	+23:34:55.4	63.398	-6.121	13.68	1.4	1.190 ± 0.005	8.69 ± 0.09	31.62 ± 0.05
3C409B	20:14:27.7	+23:34:50.2	63.398	-6.122	—	1.3	1.303 ± 0.005	8.50 ± 0.08	31.52 ± 0.05
3C410A	20:20:06.6	+29:42:14.8	69.212	-3.769	2.88	1.4	3.501 ± 0.007	17.81 ± 0.11	61.55 ± 0.17
3C410B	20:20:06.7	+29:42:09.6	69.211	-3.770	6.39	2.0	3.146 ± 0.010	17.10 ± 0.15	60.67 ± 0.17
B2050	20:52:52.1	+36:35:35.3	78.858	-5.124	5.14	0.8	0.331 ± 0.003	2.68 ± 0.05	26.71 ± 0.06
3C433	21:23:44.6	+25:04:02.2	74.475	-17.697	10.33	3.0	0.467 ± 0.008	1.98 ± 0.12	10.30 ± 0.06

We reduce all 21-SPONGE data using the Astronomical Image Processing System (AIPS¹).

For a full description of the data reduction strategy, see Chapter 2. For each source, we produce a cleaned, calibrated data cube and continuum image. We then extract the absorption spectrum from the pixel of maximum flux density, and divide by the continuum flux density at the pixel location to compute $\exp(-\tau(v))$. Our channel spacing of 1.95 kHz at the H I frequency corresponds to a channel spacing in velocity of 0.42 km s^{-1} , which corresponds to a velocity resolution of 0.5 km s^{-1} (Rohlfs & Wilson 2004).

Of the original 59 target sources, 11 were removed upon inspection of preliminary data products for being overly resolved (4 sources), displaying saturated absorption (2 sources), or for not receiving any integration time following the conclusion of the observing program (5 sources). An additional 8 sources were resolved into multiple continuum peaks, thereby providing additional sources for extracting H I absorption, albeit at degraded sensitivity due to the loss of continuum flux density. A gallery of continuum images for the 48 separate targets, demonstrating the range of complexity in source structure, is included in Appendix B, Figure B.1. Overall, we have extracted 57 total H I spectra. Table 5.1 displays information for the 57 final lines of sight (LOS), including coordinates, NVSS flux density (Condon et al. 1998), and RMS noise in optical depth (σ_τ) computed within offline channels.

¹<http://www.aips.nrao.edu/>

Table 5.1 (cont'd)

Source (name)	RA (J2000) (hh:mm:ss)	Dec (J2000) (dd:mm:ss)	<i>l</i> (°)	<i>b</i> (°)	$S_{1.4 \text{ GHz}}$ (Jy)	σ_τ ($\times 10^{-3}$)	τ_{peak}	$\int \tau \text{ dv}$	$N(\text{HI})_{\text{obs}}$ (10^{20} cm^{-2})
PKS2127	21:30:32.9	+05:02:17.5	58.652	-31.815	4.10	0.7	0.128 ± 0.001	0.56 ± 0.03	5.55 ± 0.04
J2136	21:36:38.6	+00:41:54.2	55.473	-35.578	3.47	1.2	0.143 ± 0.002	0.91 ± 0.05	4.84 ± 0.03
J2232	22:32:36.4	+11:43:50.9	77.438	-38.582	7.20	1.0	0.156 ± 0.002	1.10 ± 0.04	5.76 ± 0.03
3C454.3	22:53:58.0	+16:08:52.4	86.112	-38.185	12.66	1.1	0.317 ± 0.002	1.87 ± 0.05	8.18 ± 0.02
3C459	23:16:35.2	+04:05:18.1	83.040	-51.285	4.68	0.9	0.142 ± 0.002	1.16 ± 0.04	6.52 ± 0.03

Note. — Col. (1): Source name. Cols. (2) through (5): R.A. and Dec, *l* and *b* coordinates. Col. (6): Flux density at 1.4 GHz (Condon et al. 1998). Col. (7): RMS uncertainty in optical depth, measured in off-line channels. Col. (8): Peak optical depth. Col. (9): Integrated optical depth. Col. (10): Observed H I column density (Equation 5.8).

In comparison with Table 2.2, we have improved σ_τ for many sources by including additional integration time and/or re-processing the original files. We display a histogram of σ_τ in Figure 5.1. Generally, $\sigma_\tau < 1 \times 10^{-3}$ (median value = 9×10^{-4}) which makes 21-SPONGE among the highest-sensitivity surveys for H I absorption ever undertaken, and the largest of these by almost a factor of two.

In Chapter 2 (Figure 2.1), we demonstrated excellent agreement between 21-SPONGE and other H I absorption studies by comparing the integrated H I optical depths for sources which overlap with the Millennium Arecibo H I Absorption Line Survey (Heiles & Troland 2003a, hereafter HT03), Stanimirović & Heiles (2005) and Roy et al. (2013b). Of our 48 targets, we overlap with 22/78 from HT03, 9/35 from Roy et al. (2013b) and 9/104 from Mohan et al. (2004) (Table 2.1 summarizes these and other external surveys). We find consistent agreement with these studies at the level of our uncertainties.

5.2.2 Matching H I Emission Profiles

To estimate temperature and column density for H I structures detected in 21-SPONGE using radiative transfer calculations, we need information about the brightness temperature (i.e., emission) of H I probed by our VLA absorption profiles. Observing H I emission on the same angular scale as the H I absorption measurement is ideal, however these measurements are prohibitively expensive to conduct at an interferometric facility such as the VLA. Therefore, we obtain H I emission profiles from the 305 m Arecibo Observatory, whose $3'.5$ beam at 1.420 GHz allows us to minimize the effects of mismatched beam sizes on interpreting H I spectra. 31 sources were obtained from Arecibo observing project A2770, and 11 sources were obtained from publicly-available data from HT03 (see Section 2.2.4 for details on Arecibo H I emission profiles). The remaining 5 sources, which lie outside of the Arecibo field of view, were obtained from the next-highest resolution survey available: the Effelsberg-Bonn H I Survey (EBHIS, $9'$ resolution at 21 cm; Winkel et al. 2016).

5.2.3 Absorption Noise Spectra

The uncertainty in each spectral channel depends on the system temperature, which can be significantly increased by strong brightness temperature at Galactic velocities. To determine the frequency-dependent “noise spectra” for each line of sight, we follow the methods described in Section 2.3.2, which were derived following Roy et al. (2013b).

5.3 Integrated Properties

5.3.1 Line of Sight Properties

In Table 5.1, we also include LOS parameters for all 21-SPONGE spectra. First, we list the peak optical depth (τ_{peak}), with uncertainties equal to the value of the absorption noise spectrum (see Section 5.2.3) at the velocity of the peak optical depth. Figure 5.1 (right panel) displays a histogram of τ_{peak} . We find a median $\tau_{\text{peak}} = 0.32$ and a maximum $\tau_{\text{peak}} = 0.61$. We also list the integrated optical depth ($\int \tau \, dv$), with uncertainties computed by adding the uncertainty in each spectral channel in quadrature.

Next, we list the total observed column density ($N(\text{HI})_{\text{obs}}$) given by,

$$N(\text{HI})_{\text{obs}} = C_0 \int \frac{\tau(v) T_B(v)}{1 - e^{-\tau(v)}} \, dv, \quad (5.1)$$

where $C_0 = 1.823 \times 10^{18} \text{ cm}^{-2}/(\text{km s}^{-1} \text{ K})$, and dv is measured in km s^{-1} . This quantity is also

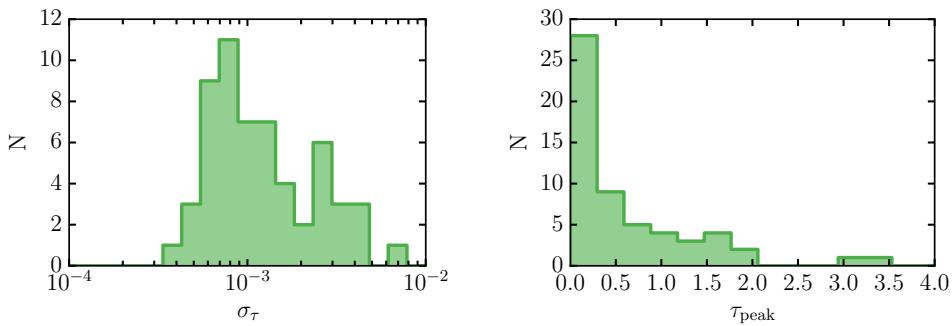


Figure 5.1 Left: RMS uncertainties in optical depth (σ_τ); Right: Peak optical depth (τ_{peak}).

known as the “isothermal” estimate of the total H I column density (e.g., Dickey & Benson 1982; Chengalur et al. 2013). Our observations probe $N(\text{HI})_{\text{obs}} \geq 10^{20} \text{ cm}^{-2}$.

5.4 Gaussian Fitting

To derive physical properties of individual H I structures along each LOS, we decompose all H I emission and absorption spectral line pairs into Gaussian functions. As shown via detailed analysis of synthetic H I spectra in Chapter 4, Gaussian components can successfully recover the true physical properties of H I structures, including spin temperature and column density.

We begin by decomposing the VLA H I absorption spectra uniformly using the Autonomous Gaussian Decomposition algorithm (AGD; Lindner et al. 2015). AGD implements derivative spectroscopy and machine learning to produce objective, efficient guesses for the basic parameters of Gaussian functions, including the number of components, and their amplitudes, positions and widths. Following the method described in Lindner et al. (2015) and employed in Chapter 4, we train the algorithm using a synthetic absorption line dataset constructed from spectral line parameters from HT03. From the training process, we determine optimal values of the two-phase smoothing parameters, $\alpha_1 = 1.12$ and $\alpha_2 = 2.75$, required by AGD to compute spectral line parameter guesses. We then use the Python implementation of AGD² to decompose the 21-SPONGE absorption lines using these values and a signal to noise ratio of $S/N = 5$. As shown in Lindner et al. (2015) and Chapter 4, the resulting parameters of the decomposition are statistically indistinguishable from those found in the by-hand analysis of the first 31 21-SPONGE sources in Chapter 2. We emphasize the benefits of the AGD decomposition: to eliminate subjective biases of human-derived guesses, and to ensure that the fits are completely reproducible.

After decomposing each H I absorption spectra using AGD into N components, we produce a model for the optical depth along the LOS:

²GaussPy; <https://github.com/gausspy/gausspy>

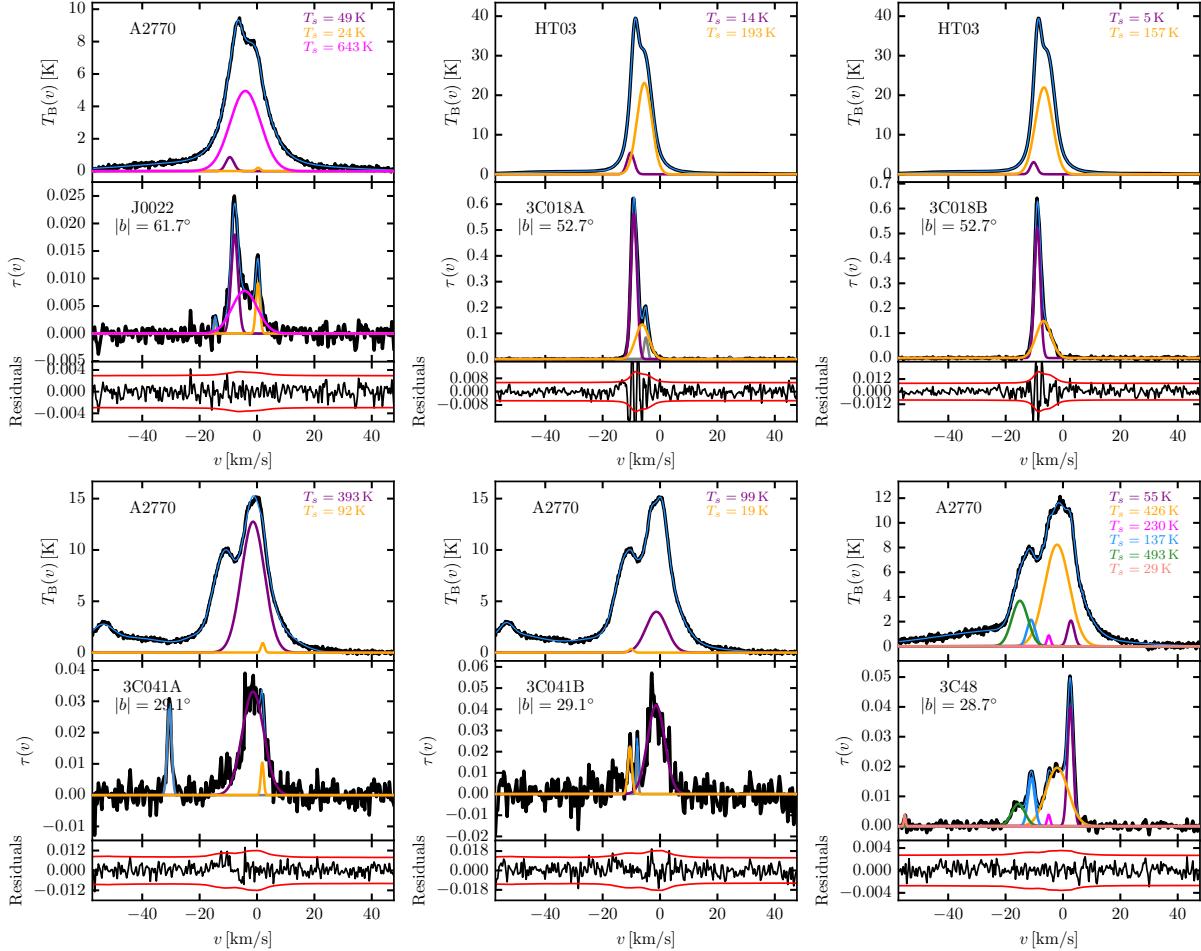
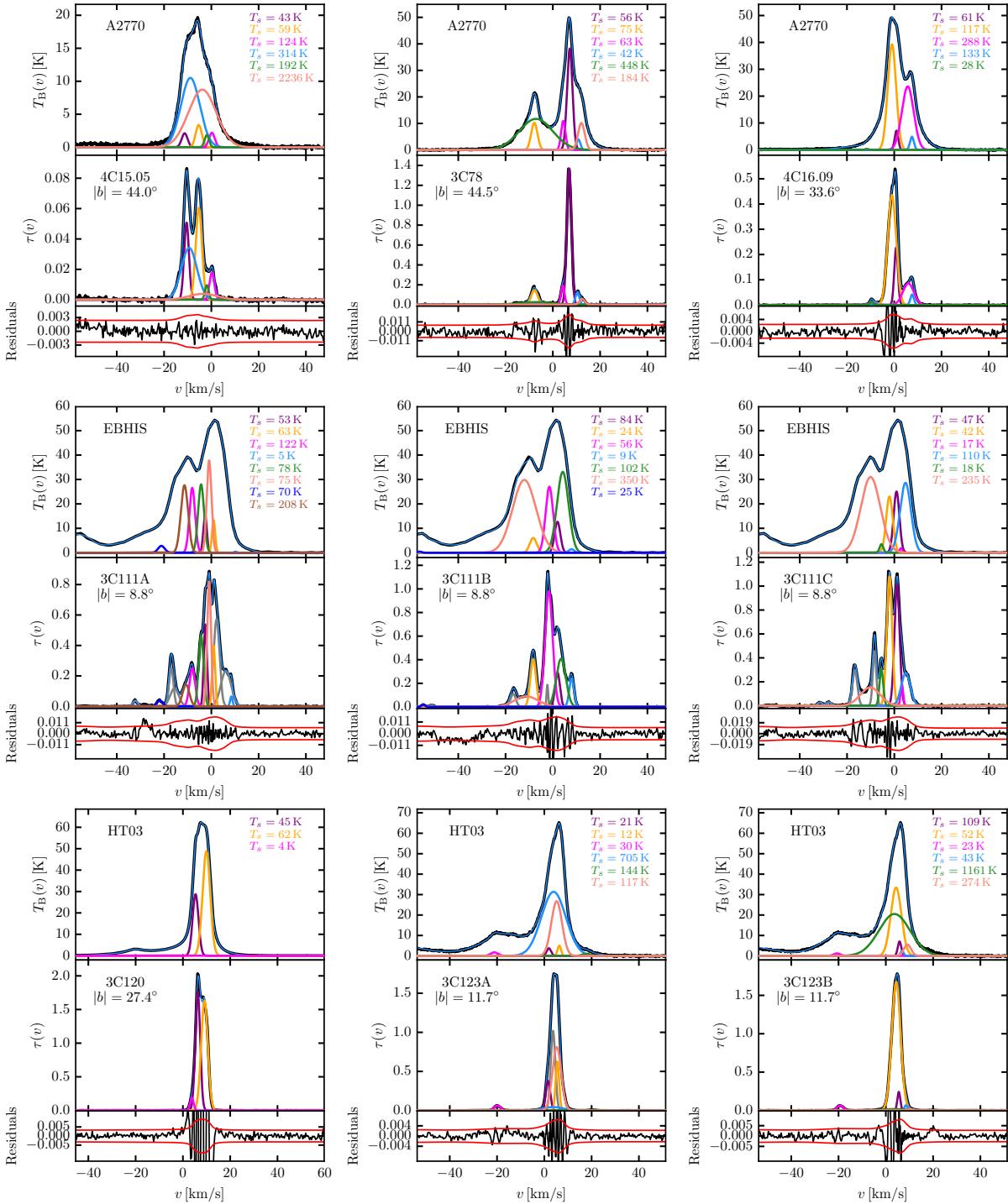


Figure 5.2 A summary of the Gaussian fits to 21-SPONGE H I emission and absorption spectral line pairs described in Section 5.4. In each panel, we plot the $T_B(v)$ (top), $\tau(v)$ (middle) and residual spectrum following the absorption fit (bottom), including $\pm 3 \times \sigma_\tau$ (i.e., the absorption noise spectra; red). We plot all fitted absorption components in the middle panel. Components whose derived spin temperatures (Equation 5.7) are unphysical (i.e., $\leq 3\text{ K}$) are plotted in grey, and components with $T_s > 3\text{ K}$ are plotted in matching colors in the middle and top panels. The total fits to $T_B(v)$ and $\tau(v)$ are displayed in thin, light blue lines. The source of $T_B(v)$, whether from Arecibo (A2770 or HT03) or EBHIS is printed in the top panels, along with the T_s values for each component. In the middle panels, we print the source name and the absolute Galactic latitude ($|b|$).

**Figure 5.1** (contd.)

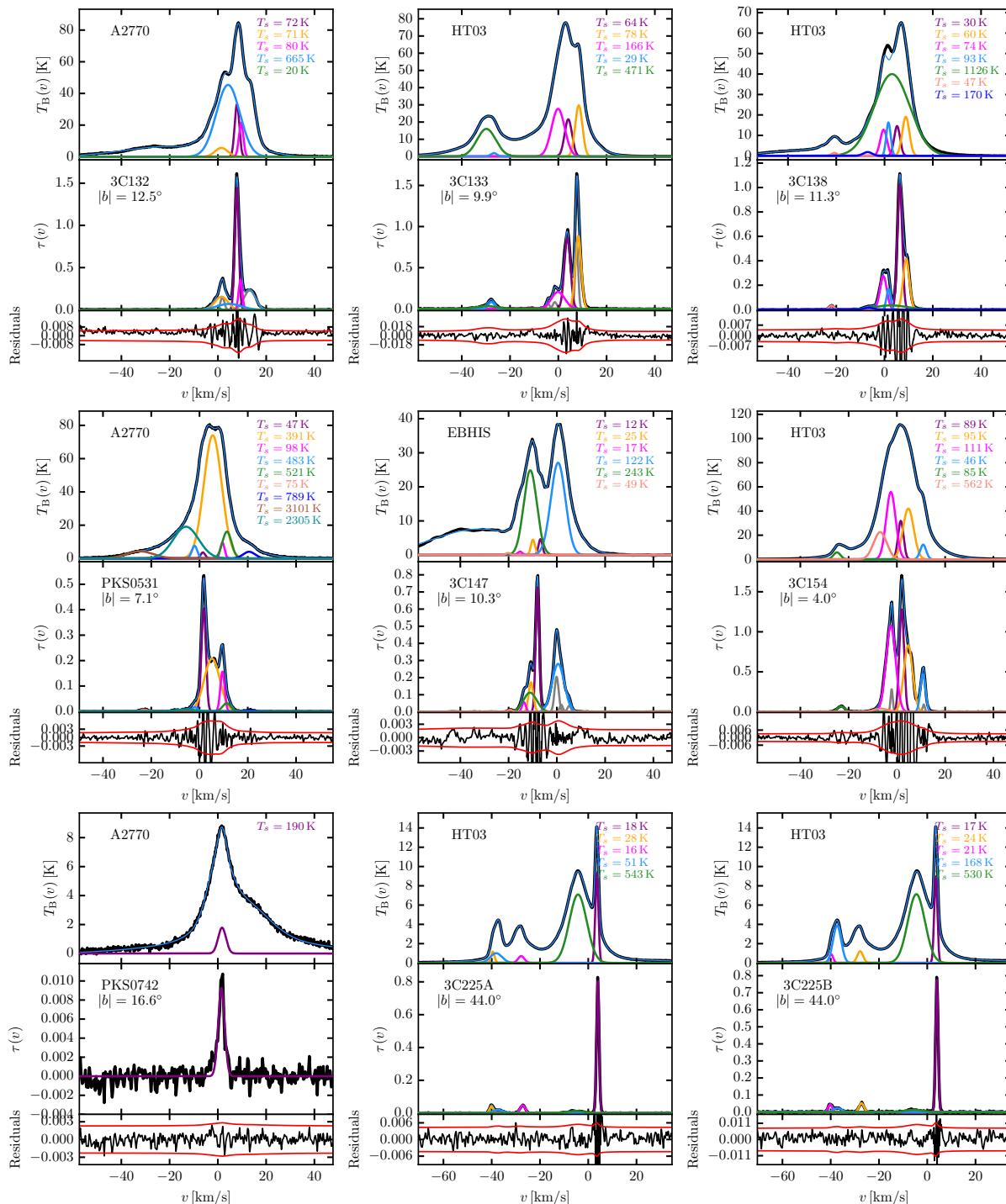


Figure 5.1 (contd.)

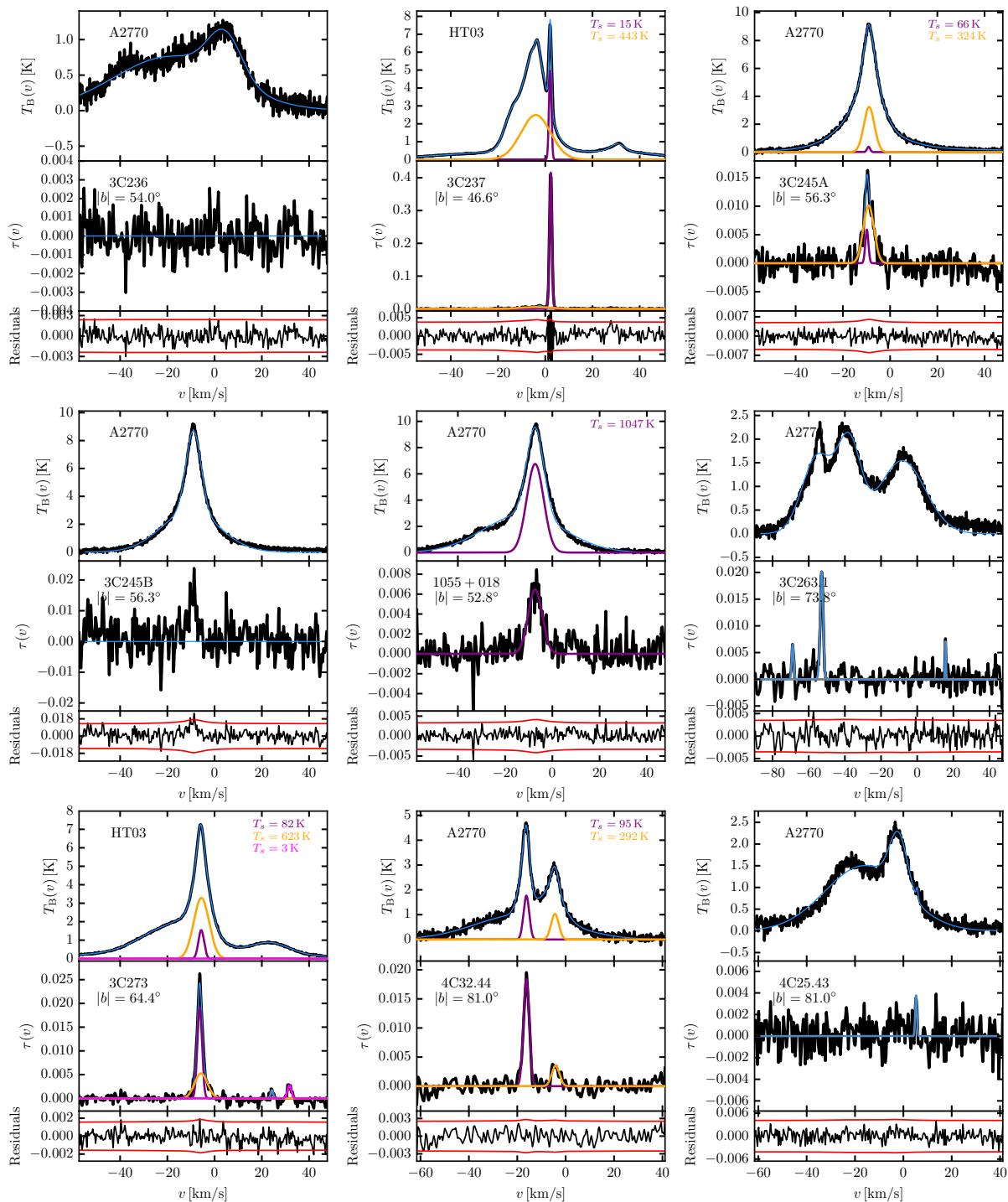


Figure 5.1 (contd.)

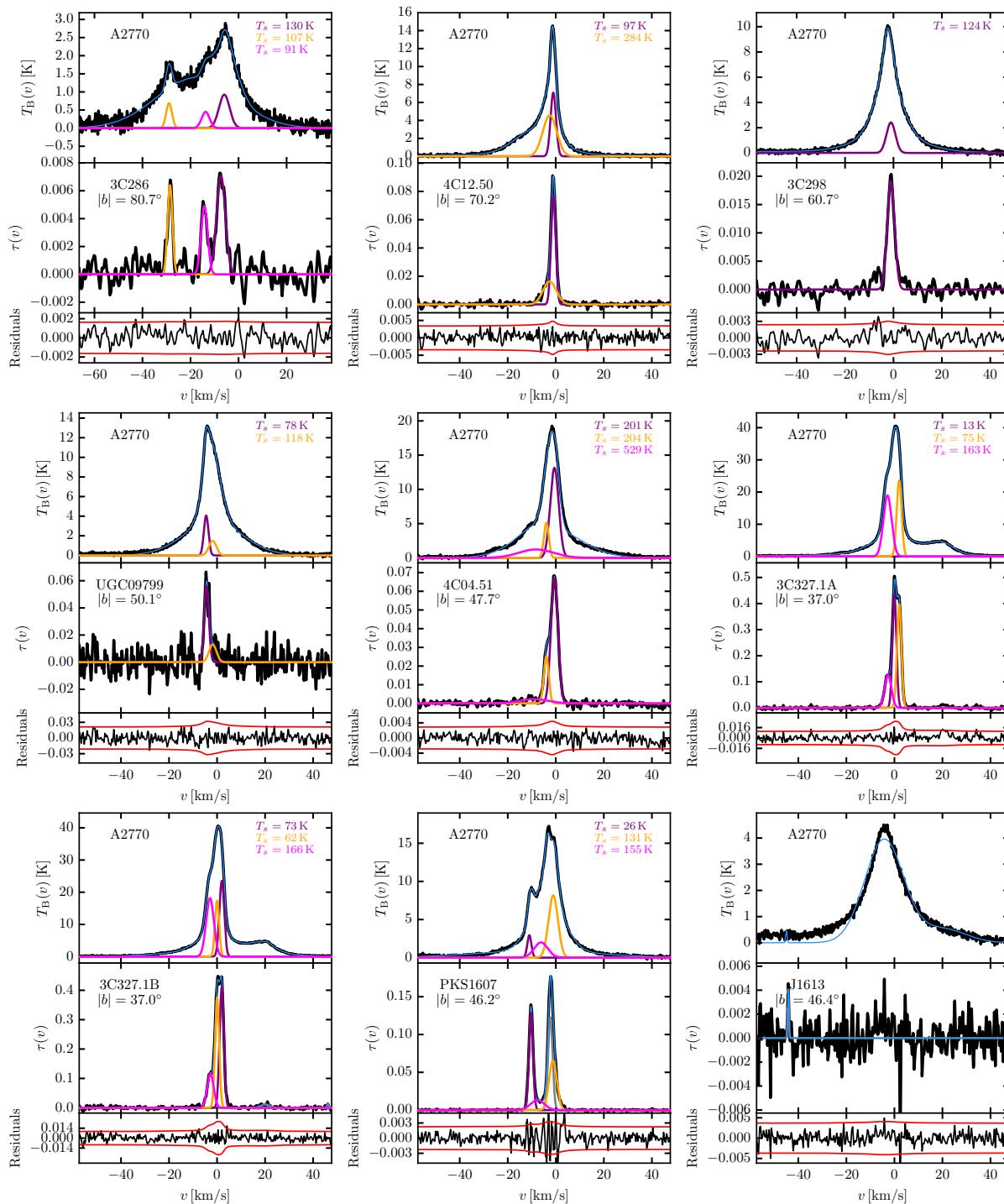
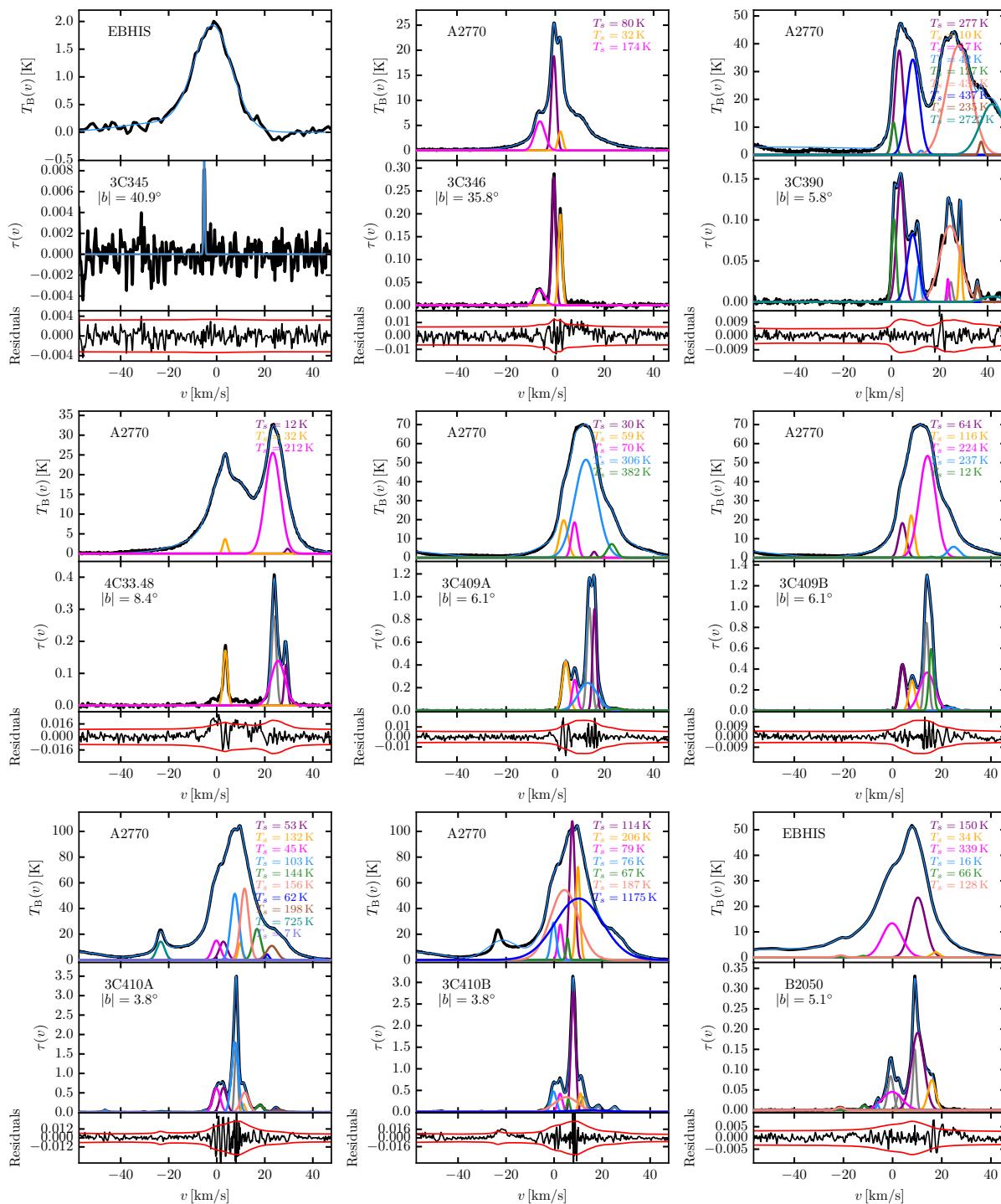


Figure 5.1 (contd.)

**Figure 5.1** (contd.)

$$\tau(v) = \sum_{n=0}^{N-1} \tau_{0,n} \times e^{-4 \ln 2 (v - v_{0,n})^2 / \Delta v_n^2} \quad (5.2)$$

where $(\tau_{0,n}, v_{0,n}, \Delta v_n)$ are the amplitude, mean velocity and full width at half maximum of the n^{th} component.

To fit these components to the brightness temperature along the LOS ($T_B(v)$), we assume a two-component H I medium, wherein some clouds contribute both opacity and brightness temperature (i.e., detected in emission and absorption), and some clouds are dominated by the WNM and contribute only brightness temperature to the LOS (e.g., Mebold et al. 1997; Dickey

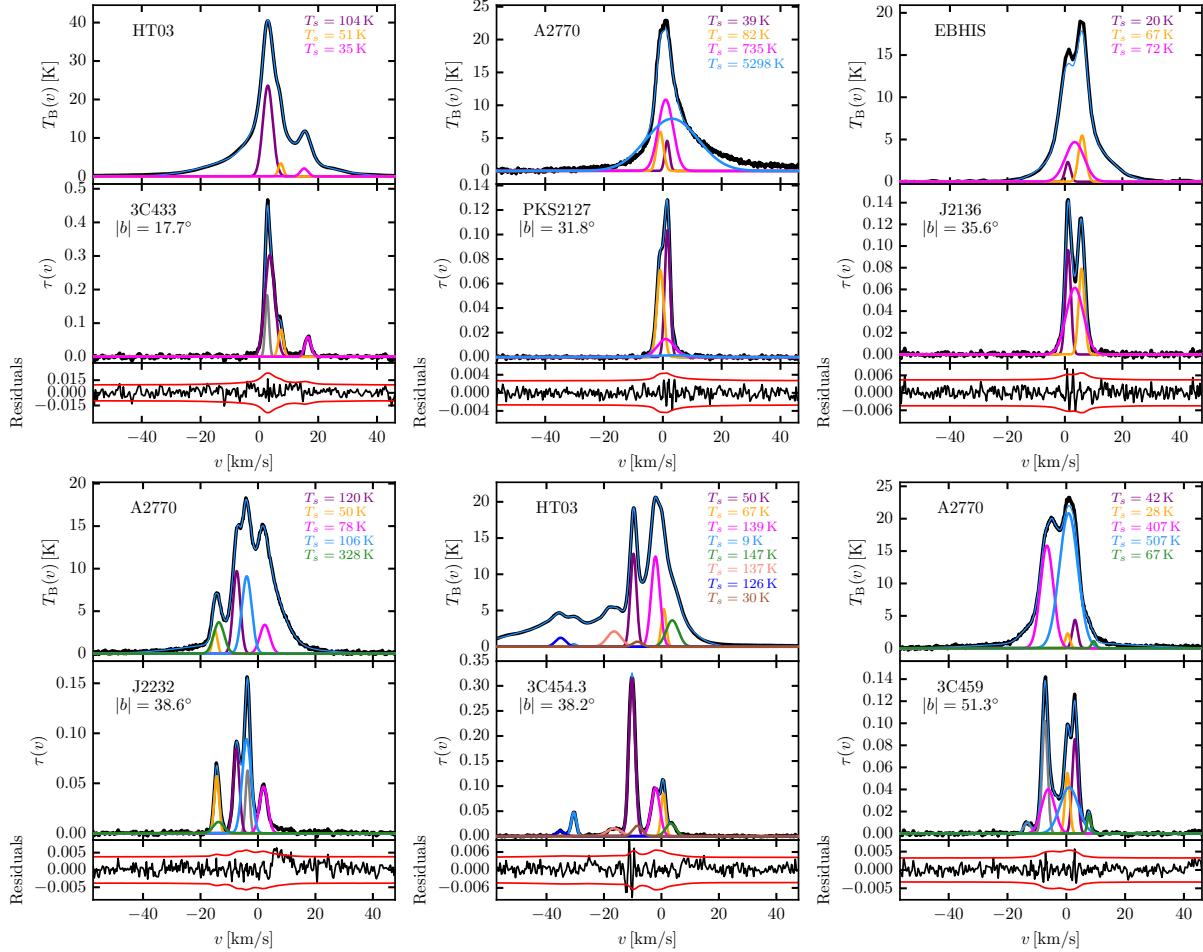


Figure 5.1 (contd.)

et al. 2000, HT03). Therefore, we solve,

$$T_B(v) = T_{B,\text{abs}}(v) + T_{B,\text{em}}(v). \quad (5.3)$$

To determine the contributions of absorption-detected and emission-only components to $T_B(v)$, we developed a method to fit all components simultaneously to $T_B(v)$ using GaussPy, following the strategies described in HT03. The method is described with the following steps:

1. Fit all N absorption-detected components to $T_B(v)$ via a least-squares fit. The mean velocities and widths are allowed to vary by $\pm 10\%$ and their amplitudes are constrained so that $0 < T_{B,n} = T_{s,n} \times (1 - e^{-\tau_{0,n}})$ and $T_{s,n} \leq T_{k,\text{max},n} = 21.866 \times \Delta v_n^2$, to produce realistic spin temperatures.
2. Subtract the best fit model in step (1) from $T_B(v)$ to produce a residual emission spectrum, which contains only emission not detected in absorption.
3. Apply GaussPy to fit K new components to the residual emission spectrum from (2), using the trained one-phase value of $\alpha = 3.75$ and $S/N = 5$ from previous analysis of HT03 emission spectra. We remove any components whose mean velocities agree with previously-detected absorption components within 1 spectral channel (i.e., 0.42 km s^{-1}).
4. Combine the $N + K$ Gaussian components from steps (1) and (3) and execute a final least squares fit to $T_B(v)$. In this final fit, we allow all mean velocities and widths to vary by 10%, and constrain all amplitudes such that $T_B > 0$.

Given a final list of $N+K$ Gaussian components fitted to $T_B(v)$ from the procedure described above, we solve Equation 5.3 for all possible orderings of the N absorption components along the LOS, and for varying absorption properties of the K emission-only components, following HT03. In detail, we solve:

$$T_{B,\text{abs}}(v) = \sum_{n=0}^{N-1} T_{s,n}(1 - e^{-\tau_n(v)}) e^{-\sum_{m=0}^M \tau_m(v)}, \quad (5.4)$$

where the subscript “ m ” refers to all components which lie in front of the n^{th} component, and,

$$T_{B,\text{em}}(v) = \sum_{k=0}^{K-1} [\mathcal{F}_k + (1 - \mathcal{F}_k)e^{-\tau(v)}] T_{0,k} e^{\frac{-4 \ln 2(v - v_{0,k})^2}{\Delta v_k^2}}, \quad (5.5)$$

where $(T_{0,k}, v_{0,k}, \Delta v_k)$ are the amplitude, mean velocity and full width at half maximum of the k^{th} component fitted only in emission, and \mathcal{F}_k is the fraction of each component lying in front of all absorption components. Previous analysis has shown that \mathcal{F} is a highly uncertain (e.g., HT03; Stanimirović et al. 2014; Murray et al. 2015), and yet it has a significant effect on the derived spin temperatures. Therefore, to minimize our fit iterations and to test the extreme cases, we allow $\mathcal{F}_k = 1$ or $\mathcal{F}_k = 0$ for all K emission-only components. In other words, we assume that all emission components lie in front of or behind all absorption components. Furthermore, the order of the absorption components along the LOS will only affect $T_{B,\text{abs}}(v)$ in the cases of components which overlap significantly in velocity (HT03). For each LOS, there are a maximum of $N!$ possible orderings, but we select only the unique orderings corresponding to components which overlap by more than $3\sigma_\tau$. Therefore, there are a total of $2 \times N!$ possible iterations for the final fit to Equation 5.3, but in practice there are many fewer for each LOS.

We select the iteration with the smallest residuals as the “best fit”, and determine a final estimate for the spin temperature of each absorption-detected component by computing the mean and standard deviation over all ordering trials. We fit a total of 289 absorption-detected components and 358 emission-only components to the 57 lines of sight. Figure 5.2 displays all 57 21-SPONGE sources and the results of the Gaussian fitting process.

5.4.1 Synthetic H I Fits

We use the same methodology described above to fit the sample of synthetic H I absorption and emission spectral line pairs presented in Kim et al. (2014, hereafter KOK14), derived from

the simulation of Kim et al. (2013, hereafter KOK13). From the full set of 10^4 synthetic lines of sight, we selected the 9355 spectra with non-saturated absorption (defined as $\tau \leq 3$). We analyzed the same synthetic dataset in Chapter 4, wherein we presented a method for matching Gaussian spectral lines to “real” gas structures found in the simulation. However, the new method presented in this chapter improves the statistics of components for which we can derive T_s and $N(\text{HI})$. The method described in Chapter 4 selected only those components with unambiguous signatures in $\tau(v)$ and $T_B(v)$, and resulted in fewer detected components per LOS. We fit a total of 14954 components for the 9355 synthetic absorption lines.

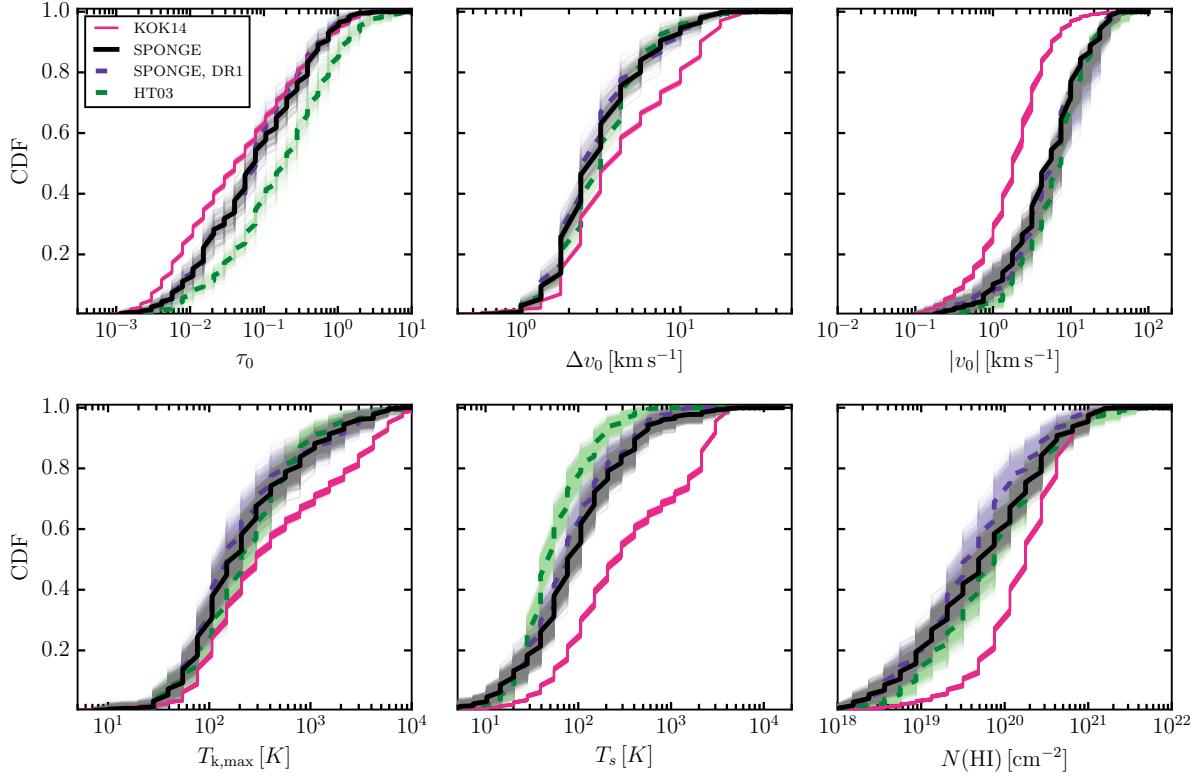


Figure 5.3 Parameters derived from the Gaussian fits to the 57 21-SPONGE absorption lines (289 components) via AGD (Section 5.4; black). These include: optical depth amplitude (τ_0), FWHM (Δv_0) and mean velocity ($|v_0|$), maximum kinetic temperature ($T_{k,\text{max}}$), spin temperature (T_s) and column density ($N(\text{HI})$). For comparison, we include the results of HT03 (green dashed), Chapter 2 (purple dashed) and a fit to KOK14 synthetic H I spectral line pairs using the same methodology (pink).

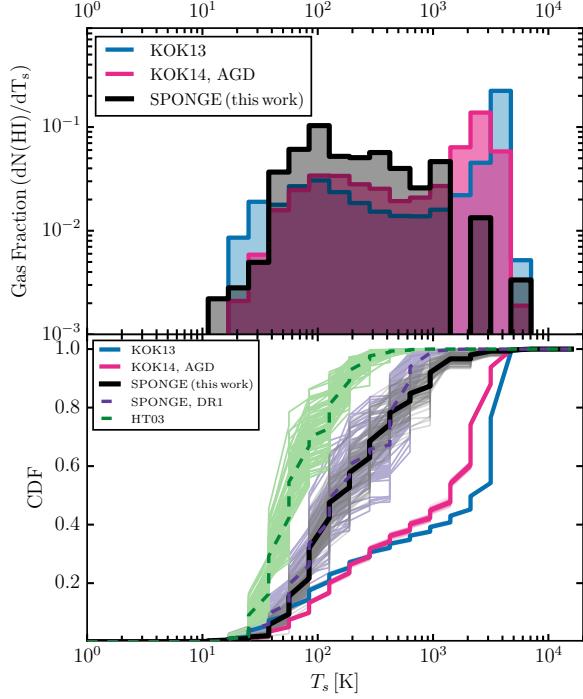


Figure 5.4 Fraction of total column density detected by 21-SPONGE as a function of spin temperature. Chapter 2 (purple dashed) and HT03 (green dashed) results are shown for comparison. We also include our Gaussian analysis of KOK14 synthetic spectral lines (pink) and the results of the KOK13 simulation (blue; i.e., Figure 8d of KOK13).

5.5 Results

5.5.1 Fitted Parameters

For the subsequent analysis, we focus solely on the absorption-detected components for which we have constraints on both τ and T_B for determining accurate spin temperatures and column densities. Furthermore, in the presence of stray radiation in the Arecibo H I emission profiles (see Section 2.2.5), the components fit to the H I emission profile may contain spurious emission. Therefore, we use these components solely as a means of quantifying the total fraction of emission that is not accounted for by absorption-detected components.

Figure 5.3 displays cumulative distribution functions (CDFs) of the fitted parameters for all absorption-detected components. In the top row, these parameters are from the best fit to

$\tau(v)$, and include: optical depth amplitude (τ_0), FWHM (Δv_0), absolute mean velocity ($|v_0|$). In the bottom row, we include derived physical properties. On the left, we compute the maximum kinetic temperature ($T_{k,\max}$) given by,

$$T_{k,\max} = \frac{m_H}{8k_B \ln 2} \Delta v_0^2 = 21.866 \times \Delta v_0^2, \quad (5.6)$$

for hydrogen mass m_H , and Boltzmann's constant k_B (e.g., HT03 Murray et al. 2015). Next, we compute the spin temperature (T_s) of each component, which we define as,

$$T_s = \frac{T_{B,n} \times \Delta v_{\text{em},n}}{(1 - e^{-\tau_{0,n}}) \times \Delta v_{0,n}}, \quad (5.7)$$

where $(T_{B,n}, \Delta v_{\text{em},n})$ are the amplitude in brightness temperature and FWHM of each absorption-detected component from the best fit to $T_B(v)$. In this panel (and all panels in Figure 5.3), we keep only those components with physically reasonable values of T_s , defined as $T_s > 3$ K. Components with $T_s \leq 3$ K are displayed in grey in Figure 5.2. A total of 225/289 (78%) components have physically reasonable T_s . For 22% of LOS, all absorption-fitted components are recovered in the fit to $T_B(v)$ with reasonable T_s , and the recovery fraction is ≥ 0.5 for 98% of cases. We find similar recovery results in the decomposition of KOK14 spectra.

Next, we compute the H I column density,

$$N(\text{HI}) = 1.064 \times C_0 \times \tau_0 \times \Delta v_0 \times T_s, \quad (5.8)$$

where 1.064 converts the product to the area under a Gaussian function with the given height and width (c.f., HT03 Murray et al. 2015).

In all panels of Figure 5.3, we bootstrap each sample 100 times and include the resampled CDFs to illustrate the effect of outliers on the distributions. For comparison, in Figure 5.3 we include the results of HT03, Chapter 2, and a re-processing of the synthetic H I spectral line pair database from KOK14 (Section 5.4.1).

The parameters plotted in Figure 5.3 for all sources are listed in Appendix B, Table B.1.

5.5.2 Mass distribution of H I as a function of T_s

In Figure 5.4, we display a histogram (top panel) and CDF (bottom panel) of the fraction of total column density found in each T_s bin ($dN(\text{HI})/dT_s$), a quantity that we denote the “gas fraction” as a function of spin temperature (black). We include the results of our reprocessing of KOK14 synthetic data (pink), as well as HT03 (green dashed) and Chapter 2 (purple dashed). We also include the distribution of gas fractions from the KOK13 simulation (i.e., not processed via spectral line analysis, but gathered from the simulation results, blue).

5.5.3 Correcting for observational bias

To correct the observed 21-SPONGE mass distribution as a function of spin temperature from observational biases introduced by our analysis method (i.e., bias towards certain spectral line shapes imposed by AGD), we compare the results of the KOK13 simulation with our Gaussian analysis of the KOK14 synthetic H I spectra (i.e., pink and blue distributions in Figure 5.4). In particular, we construct a “transfer function”, $\mathcal{T}(T_s)$, where,

$$\left(\frac{dN(\text{HI})}{dT_s} \right)_{\text{true}} = \mathcal{T}(T_s) \times \left(\frac{dN(\text{HI})}{dT_s} \right)_{\text{obs}} \quad (5.9)$$

If we assume that the “true” gas fraction is traced by KOK13, and the observed gas fraction is traced by KOK14, then we have,

$$\mathcal{T}(T_s) = \left(\frac{dN(\text{HI})}{dT_s} \right)_{\text{KOK13}} / \left(\frac{dN(\text{HI})}{dT_s} \right)_{\text{KOK14}}. \quad (5.10)$$

We display $\mathcal{T}(T_s)$ for all bins with $dN(\text{HI})/dT_s > 0.001$ in the top panel of Figure 5.5. In the bottom panel of Figure 5.5, we show the original 21-SPONGE gas fraction distribution (black, Figure 5.4), then multiply this distribution by $\mathcal{T}(T_s)$ (blue) to produce a bias corrected gas fraction as a function of T_s . We bootstrap the 21-SPONGE sample 100 times and include each resampled CDF to illustrate the effect of outliers on the distribution (red).

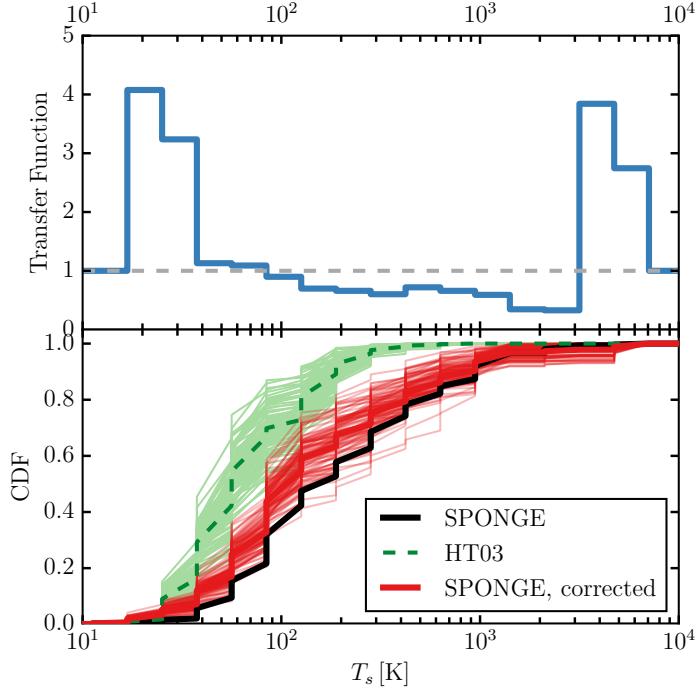


Figure 5.5 Top: Transfer function, $\mathcal{T}(T_s)$, derived by comparing the temperatures and densities of ISM simulated by KOK13 with those inferred by the synthetic observations of KOK14. Bottom: Fraction of total column density detected by 21-SPONGE as a function of spin temperature (black), corrected for observational bias using $\mathcal{T}(T_s)$ (red) and bootstrapped over 100 trials to illustrate sampling uncertainty. The HT03 results (green dashed), with 100 bootstrapped trials, are included for comparison.

In Table 5.2 we list the mass fractions of gas detected in absorption and emission from each ISM phase. We define each phase by a range of T_s : the CNM ($T_s < 200$ K), UNM ($200 < T_s < 2000$ K) and WNM ($T_s > 2000$ K). We estimate the fraction of total column density found in each T_s range by computing the mean and standard deviation over all bootstrapped trials with and without correction by $\mathcal{T}(T_s)$ (Figures 5.4 and 5.5). We emphasize that the mass fractions in the UNM and WNM are lower limits, given that we do not incorporate emission that is not detected in absorption (i.e., for which we have no constraint on τ or T_s). In addition, the CNM mass fraction estimate is an upper limit, as we do not incorporate the absorption components which were omitted during the fit to emission (i.e., those with estimated $T_s < 3$ K).

5.6 Discussion

21-SPONGE is the largest, highest-sensitivity survey for H I absorption to date. As seen in Figure 5.1, the majority of LOS have RMS uncertainty in optical depth of $\sigma_\tau < 1 \times 10^{-3}$ per channel. The outliers in the histogram are due to degraded sensitivity from lack of flux density in the cases of our 9 resolved sources. This exceptional sensitivity to optical depth makes 21-SPONGE sensitive to absorption by H I in all stable and thermally unstable ISM phases.

As seen in the right panel of Figure 5.1, we generally probe low- τ LOS. We observe $\tau_{\text{peak}} \geq 3$ in only two cases (3C410A and 3C410B). Our sources lie generally at high Galactic latitude by design, to avoid strong velocity crowding associated with the Galactic plane, and therefore the generally small τ_{peak} is consistent with expectations. In addition, although the sample size of 21-SPONGE is insufficient to probe small-scale optical depth variations across the whole sky, we note that we do not see evidence for the widespread, high-opacity ($\tau \geq 1$) H I claimed in recent analysis of *Planck* all-sky dust temperature maps and the contribution of optically-thick H I to the Galactic mass budget (Fukui et al. 2015). Future surveys of H I absorption towards fainter sources which sample the local ISM with higher density will address this issue further (e.g., GASKAP; Dickey et al. 2013).

The results of our decomposition and radiative transfer fitting process are shown in Figure 5.2. By inspection of these figures, there is generally good qualitative agreement between

Table 5.2. Mass Fractions as a function of T_s

Phase	T_s Range [K]	Fraction	Corrected Fraction ^a
CNM	< 200	0.58 ± 0.06	0.68 ± 0.06
UNM	[200, 2000)	0.37 ± 0.06	0.28 ± 0.06
WNM	≥ 2000	0.04 ± 0.02	0.04 ± 0.05

Note. — ^a: Following correction by $\mathcal{T}(T_s)$ (Equation 5.10, Figure 5.5).

the velocity structure of $\tau(v)$ and $T_B(v)$, although $T_B(v)$ is generally wider. In other words, the majority of components fitted to $\tau(v)$ are recovered by similar velocity structure observed in $T_B(v)$.

The components plotted in gray in the $\tau(v)$ are those with unphysical spin temperatures ($T_s \leq 3\text{ K}$; Equation 5.7). These components arise when a velocity structure observed in $\tau(v)$ does not appear in $T_B(v)$, caused by beam mismatch, or the fact that the emission profiles are measured on much larger angular scales than the absorption profiles, making them less sensitive to small-scale cold gas structures. In addition, the emission profile is necessarily measured using adjacent LOS from the absorption profile in order to avoid the background continuum source, which means that the two profiles are not sampling identical populations of H I structures. An example of this effect is shown in the case of 3C041A (Figure 5.2). The absorption line clearly detected at $v \sim -30\text{ km s}^{-1}$ is not recovered in $T_B(v)$, likely due to a beam mismatch or LOS effect. We also see examples of strongly blended components wherein one of two overlapping lines is lost in the fit to $T_B(v)$, for example as seen in the case of a component at $v \sim 10\text{ km s}^{-1}$ towards 3C410A (Figure 5.2).

Overall, 22% of the fitted absorption lines are not recovered in the subsequent fit to $T_B(v)$. We do not expect 100% recovery because the structure of H I emission is significantly affected not only by gas temperature, but also by Galactic rotation and velocity crowding, and is further modified by absorption along the LOS (e.g., Dickey & Lockman 1990). Furthermore, the emission spectra from Arecibo (i.e. A2770 and HT03) have not been corrected for stray radiation. We expect the contribution from stray radiation to be $\lesssim 1\text{ K}$ and spread over wide velocities (10s of km s^{-1} in width; Peek et al. 2011), which most strongly contaminates the wings of the $T_B(v)$ profiles. Therefore, our analysis method was designed to apply as few assumptions as possible in the fit to $T_B(v)$. For example, in the case of PKS2127, $T_B(v)$ exhibits a prominent non-Gaussian wing at high velocities, which GaussPy neglected to fit with additional broad components. This results in a degraded reduced χ^2 for the overall fit, but avoids over-constraining the Gaussian

shapes in emission.

In Figure 5.3 we summarize the results of the decomposition of all 57 $\tau(v)$ profiles. We find that varying the decomposition scheme did not have large effect on the results or derived physical parameters. Allowing the component parameters to vary between 1% and 20% has the effect of changing the properties of the fitted emission-only components. Following the cut for components with unphysical T_s (i.e., $T_s \leq 3$ K), we end up with roughly ~ 230 components, which are stable to all “knobs” in the Gaussian process described in Section 5.4 (e.g., parameter variation, whether $T_{B,n}$ is constrained). We conclude that the main effect in producing components with unphysical temperatures is beam mismatch between absorption and emission, as well as strong line blending at low Galactic latitudes ($|b| < 20^\circ$), where we recover only $\sim 50\%$ of the physical H I gas structures along the LOS (Chapter 4).

By comparison with the previous observations shown in Figure 5.3, we are encouraged that our decomposition results (black) are statistically indistinguishable from the by-hand analysis of the first half of the 21-SPONGE survey. We are also generally consistent with HT03, except in the case of τ_0 , wherein the superior sensitivity of 21-SPONGE allows us to probe smaller H I optical depths. We also detect higher T_s in the present analysis than found by HT03, which is also attributable to our improved observational sensitivity.

Furthermore, our analysis of KOK14 spectra shows that our method is able to recover the full range of temperatures present in the simulation; the transfer function in the top panel of Figure 5.5 is remarkably flat in the range of 40 – 3000 K, suggesting that the Gaussian method is accurately reproducing the mass distribution within this regime. At high and low T_s , AGD appears to underestimate the simulated gas fractions more strongly. It is likely that our analysis misses the lowest spin temperatures in the simulation, $T_s < 40$ K, as these components are predominately blended within H I structures along complex LOS and are difficult to recover in both $\tau(v)$ and $T_B(v)$. At the high end, $T_s > 3000$ K, we are limited by observational sensitivity. We added synthetic Gaussian noise with RMS = 1×10^{-3} to each synthetic absorption profile to mimic the

21-SPONGE results, and AGD is unable to recover lines with $T_s \geq 3000\text{ K}$ at $\text{S/N} \geq 5$ with this level of spectral noise. Furthermore, high temperatures in the simulation are strongly affected by the implementation of the WF effect, which is assumed to be constant throughout the disk. In addition, as discussed in Chapter 4, the KOK13 simulations are missing a hot ionized medium phase populated by thermal feedback from supernovae, which produces unrealistic H I absorption lines with temperatures in the $1000 - 3000\text{ K}$ range. It will be important to test future simulations with a more detailed implementation of the WF effect and supernova feedback.

5.6.1 Significant fraction of thermally unstable H I

Following correction by $\mathcal{T}(T_s)$, we detect a substantial fraction of thermally unstable H I in absorption: $\sim 28\%$ by mass. This is generally consistent with previous observational results. For example, a smaller high-sensitivity H I absorption line study from Roy et al. (2013a) argued that at most 28% of H I is thermally unstable. HT03 argued that 48% of the Galactic H I mass is thermally unstable. We emphasize that HT03’s conclusions for H I with $T_s \gtrsim 500\text{ K}$ were based on upper limits to the spin temperature derived using emission-detected line widths, and not detections in absorption, and therefore their estimate of $\sim 50\%$ instability is an upper limit. Although our estimate is lower than HT03, our superior sensitivity allows us to place stronger constraints on the T_s values of non-CNM gas. Furthermore, our estimate is a lower limit to the total unstable gas fraction, given that we only account for emission that is also detected in absorption. The emission for which we do not have constraints on τ or T_s is likely due to the UNM or WNM. We note that KOK13 also find a substantial fraction of gas out of thermal equilibrium, due to strong turbulence, expanding shocks from supernovae, and a time-dependent heating rate (KOK13). Below $T_s \sim 3000\text{ K}$, our observed distribution qualitatively agrees with their results (c.f., Figure 8d of KOK13).

Objections to previous observational estimates of a high UNM fraction in the Galactic ISM (e.g., HT03) suggested that single-dish measurements of 21 cm absorption are unreliable (e.g., Roy et al. 2013b), and that Gaussian decomposition is inherently non-unique and not appropriate

for characterizing H I structures (e.g., Braun & Kanekar 2005). We showed that Arecibo H I absorption measurements are entirely consistent with interferometric VLA observations of bright sources (> 3 Jy), thereby proving that spectra from single-dish studies such as HT03 are not affected by resolved H I emission in the majority of cases. Furthermore, via comparison with synthetic observations of numerical simulations, we demonstrated that Gaussian analysis of 21 cm spectral lines can successfully recover significant populations of real H I structures. Although the simulations have known limitations, and the uncertainty in observational estimates of T_s and $N(\text{H}\text{I})$ is large, we used the comparison between inferred and “true” conditions in the simulation to correct for the associated biases of our analysis method. Overall, the detection of a significant fraction of H I in the thermally unstable regime by 21-SPONGE supports the previous observational and theoretical argument that time-dependent radiative and dynamical processes on both macro and micro-physical scales are crucial for determining the thermodynamic state of the ISM.

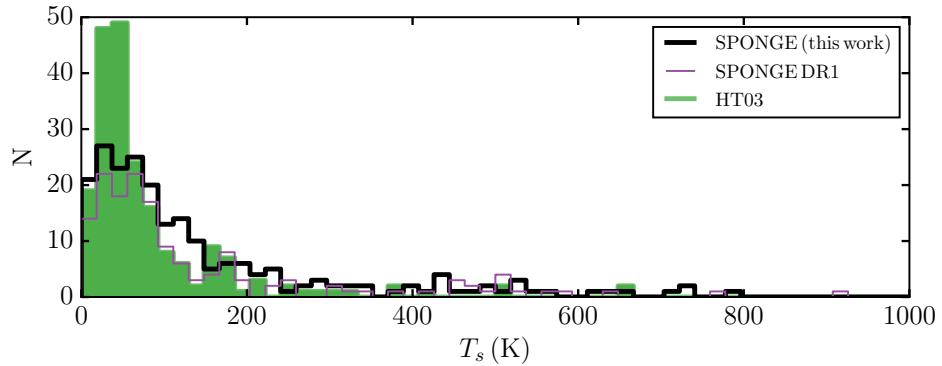


Figure 5.6 Histogram of spin temperatures (T_s) for components with $T_s \leq 1000$ K from 21-SPONGE (57 LOS; thick black), the first data release (“DR1”, i.e., Chapter 2; thin purple) and HT03 (79 LOS, filled green). The higher sensitivity of 21-SPONGE has the effect of broadening the T_s distribution for the CNM ($T_s \lesssim 200$ K) in both by-hand analysis following the HT03 method (DR1) and the decomposition method presented in this chapter.

5.6.2 Broad CNM distribution

In addition to a significant fraction of UNM, we detect a high mass fraction of the CNM in absorption, $\sim 68\%$. An interesting result from the 21-SPONGE survey is that with improved sensitivity in optical depth, signatures of the CNM in the form of weak, narrow absorption lines can be detected. This agrees with the analysis of Stanimirović & Heiles (2005), who showed that increased integration time on non-detections from HT03 revealed CNM absorption. Out of 57 total LOS, we have only 5 nondetections ($\sim 91\%$ detection rate). Our CNM temperature histogram is also broader than what has been found by HT03. In Figure 5.6, we display histograms of T_s from 21-SPONGE (thick black) and HT03 (filled green). The 21-SPONGE results from this work agree superbly with the results of Chapter 2 (thin purple; Murray et al. 2015), which followed an identical methodology as presented in HT03. This agreement suggests that the broadening of the CNM distribution is not related to the method of deriving T_s , but is likely an effect of the improved sensitivity of 21-SPONGE over HT03.

We note that correcting the observed gas mass distribution with $\mathcal{T}(T_s)$ recovers low- T_s components “missed” by our decomposition and radiative transfer method. Furthermore, the HT03 and corrected 21-SPONGE distributions are consistent within sampling uncertainties below $T_s \sim 100$ K (illustrated by the bootstrapped trials in the bottom panel of Figure 5.5), indicating that the two results are statistically indistinguishable at low temperatures.

5.6.3 Lack of absorption by WNM with $T_s > 2000$ K

The lack of WNM ($\sim 3\%$) detected by 21-SPONGE in absorption is puzzling. As previously discussed in Chapter 4, the isolated, strong, broad synthetic 21 cm absorption lines associated with the significant peak of gas at $2000 < T_s < 4000$ K are simply not observed in 21-SPONGE or other absorption line studies. Our analysis method is able to detect gas at these temperatures (Figure 5.4), and the correction by $\mathcal{T}(T_s)$ accounts for the T_s -dependent biases discussed in detail in Chapter 4. We note that our estimate for the WNM fraction is a lower limit, given we do not account for the emission not detected in absorption for which we do not have constraints for τ or

T_s . If the WNM is strongly affected by supplemental excitation beyond collisions such as the WF effect, individual LOS from 21-SPONGE are likely still not sensitive enough to detect absorption by gas with the resulting temperature of $T_s \gtrsim 7000$ K (Chapter 3; Murray et al. 2014). Larger samples of 21 cm absorption lines at high sensitivity are required to improve these uncertainties.

5.7 Summary

One of the main goals of 21-SPONGE is to conduct a statistically significant survey of the temperature and density of Galactic H I. In this chapter, we have presented the final data release of the 21-SPONGE survey. To this end, we model each H I spectral line pair using Gaussian functions, and compute properties of individual H I structures along each LOS using a new decomposition and radiative transfer method. The efficient, objective nature of the analysis method presented here enabled us to compare our results with synthetic observations of numerical simulations, which is crucial for understanding the population of H I structures probed by the 21-SPONGE observations. Our main results are summarized below:

1. 21-SPONGE reaches excellent sensitivity in H I optical depth, $\sigma_\tau < 10^{-3}$, thanks to careful calibration considerations (c.f., Appendix A) and the upgraded capabilities of the WIDAR correlator at the VLA for producing wide spectral bandwidths and narrow velocity resolution.
2. The 21 cm absorption detection rate of 21-SPONGE is high: $\sim 91\%$. We demonstrate that with improved sensitivity to optical depth, narrow absorption lines associated with the CNM are detected ubiquitously.
3. The results of the AGD-based decomposition and radiative transfer method described in this work (Section 5.4), which builds from the methods of HT03 and Chapter 4, agree well with previous by-hand analysis of 21-SPONGE spectra, and with HT03 (Figure 5.3). The superior sensitivity of 21-SPONGE relative to HT03 guarantees that we are sensitive to lower optical depths and higher spin temperatures.

4. We analyzed a sample of 9355 synthetic H I spectral line pairs from Kim et al. (2014) with the same methodology, and constructed a correction function for the fraction of total column density as a function of spin temperature (i.e., gas mass fraction) by comparing the inferred spin temperatures and column densities from the synthetic spectra with the simulation results. We find that the correction function is generally flat in the range of $40 < T_s < 3000$ K, indicating that our analysis method is successfully recovering H I properties within this regime.
5. After “correcting” the observed 21-SPONGE mass fraction as a function of temperature, we compute the fractions of H I detected in emission *and* absorption by 21-SPONGE in each canonical H I phase, including the CNM (68%, $3 < T_s < 200$ K), WNM ($\sim 3\%$; $T_s > 2000$ K) and UNM ($\sim 28\%$, $200 < T_s < 2000$ K). The large fraction of mass in the thermally unstable regime is generally consistent with previous observational estimates (e.g. Heiles & Troland 2003b; Roy et al. 2013b). We emphasize that the fractions of mass in the UNM and WNM are lower limits, given that we do not incorporate emission along each LOS that was not detected in absorption.

Overall, these results call into question the validity of steady-state models of heating and cooling in the ISM. As argued by analytical and numerical simulations, time-dependent or dynamical processes such as supernovae and turbulence are important for generating pressures required to push gas into the thermally unstable regime. Future comparisons between 21-SPONGE and next-generation simulations which more accurately account for Galactic dynamics and star formation feedback at high enough spatial resolution for resolving CNM structures are necessary for identifying the mechanism(s) responsible for the significant population of UNM confirmed in this work.

This work was supported by the NSF Early Career Development (CAREER) Award AST-1056780. C. E. M. acknowledges support by the National Science Foundation Graduate Research

Fellowship and the Wisconsin Space Grant Institution. S. S. thanks the Research Corporation for Science Advancement for their support. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. The Arecibo Observatory is operated by SRI International under a cooperative agreement with the National Science Foundation (AST-1100968), and in alliance with Ana G. Méndez-Universidad Metropolitana, and the Universities Space Research Association.

References

- Audit, E., & Hennebelle, P. 2005, *A&A*, 433, 1
- Braun, R., & Kanekar, N. 2005, *A&A*, 436, L53
- Carilli, C. L., Dwarakanath, K. S., & Goss, W. M. 1998, *ApJ*, 502, L79
- Chengalur, J. N., Kanekar, N., & Roy, N. 2013, *MNRAS*, 432, 3074
- Clark, B. G. 1965, *ApJ*, 142, 1398
- Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, *AJ*, 115, 1693
- Dalgarno, A., & McCray, R. A. 1972, *ARA&A*, 10, 375
- Dickey, J. M., & Benson, J. M. 1982, *AJ*, 87, 278
- Dickey, J. M., & Lockman, F. J. 1990, *ARA&A*, 28, 215
- Dickey, J. M., Mebold, U., Stanimirovic, S., & Staveley-Smith, L. 2000, *ApJ*, 536, 756
- Dickey, J. M., Terzian, Y., & Salpeter, E. E. 1978, *ApJS*, 36, 77
- Dickey, J. M., McClure-Griffiths, N., Gibson, S. J., et al. 2013, *PASA*, 30, e003
- Dwarakanath, K. S., Carilli, C. L., & Goss, W. M. 2002, *ApJ*, 567, 940
- Ewen, H. I., & Purcell, E. M. 1951, *Nature*, 168, 356
- Ferrière, K. M., Zweibel, E. G., & Shull, J. M. 1988, *ApJ*, 332, 984
- Fukui, Y., Torii, K., Onishi, T., et al. 2015, *ApJ*, 798, 6
- Hagen, J. P., Lilley, A. E., & McClain, E. F. 1955, *ApJ*, 122, 361
- Haud, U., & Kalberla, P. M. W. 2007, *A&A*, 466, 555
- Heiles, C., & Troland, T. H. 2003a, *ApJS*, 145, 329
- . 2003b, *ApJ*, 586, 1067
- Kim, C.-G., Ostriker, E. C., & Kim, W.-T. 2013, *ApJ*, 776, 1

- . 2014, ApJ, 786, 64
- Lindner, R. R., Vera-Ciro, C., Murray, C. E., et al. 2015, AJ, 149, 138
- Mac Low, M.-M., Balsara, D. S., Kim, J., & de Avillez, M. A. 2005, ApJ, 626, 864
- McKee, C. F., & Ostriker, J. P. 1977, ApJ, 218, 148
- Mebold, U., Düsterberg, C., Dickey, J. M., Staveley-Smith, L., & Kalberla, P. 1997, ApJ, 490, L65
- Mohan, R., Dwarakanath, K. S., & Srinivasan, G. 2004, Journal of Astrophysics and Astronomy, 25, 185
- Muller, C. A., & Oort, J. H. 1951, Nature, 168, 357
- Murray, C. E., Stanimirović, S., Kim, C.-G., et al. 2017, ApJ, 837, 55
- Murray, C. E., Lindner, R. R., Stanimirović, S., et al. 2014, ApJ, 781, L41
- Murray, C. E., Stanimirović, S., Goss, W. M., et al. 2015, ApJ, 804, 89
- Peek, J. E. G., Heiles, C., Douglas, K. A., et al. 2011, ApJS, 194, 20
- Rohlfs, K., & Wilson, T. L. 2004, Tools of radio astronomy
- Roy, N., Kanekar, N., & Chengalur, J. N. 2013a, MNRAS, 436, 2366
- . 2013b, MNRAS, 436, 2366
- Stanimirović, S., & Heiles, C. 2005, ApJ, 631, 371
- Stanimirović, S., Murray, C. E., Lee, M.-Y., Heiles, C., & Miller, J. 2014, ApJ, 793, 132
- Vázquez-Semadeni, E., Gazol, A., & Scalo, J. 2000, ApJ, 540, 271
- Verschuur, G. L., & Magnani, L. 1994, AJ, 107, 287
- Vishniac, E. T., & Lazarian, A. 1999, ApJ, 511, 193
- Winkel, B., Kerp, J., Flöer, L., et al. 2016, A&A, 585, A41
- Wolfire, M. G., McKee, C. F., Hollenbach, D., & Tielens, A. G. G. M. 2003, ApJ, 587, 278

Chapter 6

Summary and Future Work

As stars form and evolve within galaxies, they generate a rich, multiphase structure of gas in the ISM via radiative and dynamical feedback. A theoretical understanding of these feedback mechanisms in the local Universe, and throughout cosmic time, requires following the flow between gas reservoirs and star formation in detail. Specifically, physical properties of the diffuse, neutral ISM bear crucial clues to the nature of turbulence, supernovae feedback and radiative excitation mechanisms. However, previous observational capabilities and systematic uncertainties in data analysis have prevented statistically significant measurements of the temperature and density of the warm neutral medium (WNM) and the thermally unstable medium (UNM). These properties govern the formation efficiency of the clouds which host the births and deaths of stars, and therefore they are essential for models of the ISM and galaxy evolution.

In this thesis, I have presented an observational study of the physical properties of multiphase H I in the Galactic ISM, including the CNM, WNM and UNM. This effort was enabled by 21-SPONGE, the largest survey for Galactic H I absorption undertaken to date with the VLA. Exceptional sensitivity to 21 cm absorption uniquely enables 21-SPONGE to constrain the mass distribution of H I as a function of temperature — specifically to measure the properties of the WNM and UNM via direct detections in absorption — for the first time. The VLA absorption observations were complemented by emission observations from the Arecibo Observatory, whose $\sim 3.5'$ beam at 21 cm provides the best available single-dish complement to the $\sim 1''$ VLA beam. As a result, 21-SPONGE is orders of magnitude closer than previous studies to sampling the same gas populations in absorption and emission.

In Chapter 2, I described the 21-SPONGE survey design, observing strategy and initial results (Murray et al. 2015). As a result of careful calibration considerations, namely maximizing signal-to-noise in our bandpass calibration solutions, characterizing periodic structures and combining calibration observations over time (c.f., Appendix A), I achieve median RMS noise in optical depth of $\sigma_\tau < 10^{-3}$ per channel for the first half of the 21-SPONGE sample. Following the Gaussian decomposition and radiative transfer method of Heiles & Troland (2003a), I estimate

the spin temperature (T_s) and column density ($N(\text{H}\text{I})$) for 157 individual spectral components. The maximum measured T_s is a factor of two higher than the maximum T_s for absorption lines in the Millennium Arecibo HI Survey (Heiles & Troland 2003b), demonstrating the power of optical depth sensitivity for probing gas at higher temperatures. Furthermore, the 21-SPONGE line of sight column densities and CNM fractions are consistent with previous observations and with three-dimensional numerical simulations, indicating the success of our Gaussian decomposition method.

Despite the sensitivity of 21-SPONGE, I detected very little WNM (< 10%) following the first half of the survey. To improve sensitivity to diffuse, warm H I even further, I developed a spectral stacking method for Galactic H I spectra and detected a pervasive population of WNM gas at high temperature presented in Chapter 3 (Murray et al. 2014). Following bootstrapped Monte Carlo analysis, the absorption feature has a positive equivalent width of $\langle \text{EW} \rangle = 0.0182^{+0.0044}_{-0.0036} \text{ km s}^{-1}$, which corresponds to a 1 in 2×10^6 chance of being spurious. The superb baseline stability of the VLA, analyzed in Appendix A, enabled the high confidence in this detection. We estimate the spin temperature of the residual feature to be $\langle T_s \rangle = 7200^{+1800}_{-1200} \text{ K}$, which is significantly (98% confidence) higher than theoretical predictions based on collisional excitation alone, likely due to the thermalization of HI by resonant Ly α scattering. This analysis provides the first observational evidence that the Ly α excitation mechanism is acting throughout the bulk of the Galactic WNM, and demonstrates that the Ly α radiation field, a quantity that is difficult to measure yet vitally important for interpreting HI signals from early epochs of cosmic structure formation, can be probed using measurements of the WNM.

In Chapter 4, I addressed the question of how well 21 cm spectral lines reproduce real H I properties by analyzing synthetic spectra from numerical simulations by Kim et al. (2014), and comparing the results with the simulation outputs (Murray et al. 2017). To perform the analysis of 10^4 synthetic spectral line pairs, I developed a novel decomposition and radiative transfer method based on the Autonomous Gaussian Decomposition algorithm (AGD Lindner

et al. 2015). I quantified the completeness of H I structure recovery by Gaussian-fitted 21 cm spectral lines for the first time and found that completeness declines with decreasing Galactic latitude due to strong line blending and turbulent cloud motions. Furthermore, the observed spin temperatures agree within a factor of two with the simulated value for the majority of H I structures; the decomposition method tends to overestimate H I temperature above ~ 1000 K due to velocity offsets between H I absorption and emission lines caused by turbulent motions. These comparisons will allow for important bias corrections in future observational results.

After quantifying the performance of the analysis method, I compared the synthetic spectra with 21-SPONGE. I found more fitted absorption and emission lines per line of sight and with a wider range of velocities in 21-SPONGE than the synthetic spectra which reflects the limitations of local box simulations and simplified treatments of supernova feedback and Galactic rotation. I further found that synthetic spectra include wide, strong absorption lines that are not seen in 21-SPONGE despite being well above our sensitivity and resolution limits. These components are likely the result of the absence of a hot, large filling-factor gas phase in the simulations, or possibly the simple treatment of the simulated WF effect. In particular, I find that excluding the WF effect enhances the population of these discrepant H I absorption components, and suggests that the WF effect is important for producing realistic 21 cm spectral line properties. Furthermore, these components have $1000 < T_s < 4000$ K, which are temperatures not detected in 21-SPONGE despite sufficient sensitivity. I suggest that this gas may have even higher temperatures ($T_s \sim 7000$ K) than what 21-SPONGE is sensitive to in individual spectra, and than what the simulated implementation of the WF effect allows. Future work is required to understand the importance of the WF effect in more detail. It is also important to test whether more realistic treatment of star formation feedback including a hot ISM reduces the incidence of simulated gas with $T_s \sim 1000 - 4000$ K.

In Chapter 5, I presented the final data release of the 21-SPONGE survey (Murray et al. 2017b, *in prep*). We have a very high detection rate, as 91% of spectra exhibit significant

21 cm absorption. Building on the analysis of synthetic H I spectral lines from Chapter 4, I developed an objective, efficient method for estimating spin temperatures and column densities of individual clouds along all 57 non-saturated lines of sight. I find excellent agreement between the decomposed properties of the full 21-SPONGE sample and the analysis presented in Chapter 2. By comparing the results of an identical analysis of synthetic spectral lines from Kim et al. (2014) with the “true” simulated gas distribution, I constructed a correction function for observational bias. I find that the correction function is remarkably flat between $40 < T_s < 3000$ K, indicating that the decomposition and radiative transfer analysis is performing well in this range. After correcting the observed 21-SPONGE distribution, I find a significant fraction of UNM ($\sim 28\%$) which agrees with previous indirect estimates based on H I emission alone. This confirms that the steady-state models of the ISM are invalid, and that time-dependent, turbulent radiative and dynamical ISM processes are important and responsible for generating thermally unstable gas.

6.1 Future Work

Despite the contributions of 21-SPONGE presented here, the model of the multiphase neutral ISM is not complete. We still lack constraints for the dominant physical mechanisms responsible for the phase distribution of H I as a function of temperature, and an understanding for how H I properties vary with environment within the Milky Way. Several future directions are described below:

(1) *Variation in WNM with Galactic environment:* To detect WNM-like absorption, we needed to stack the 21-SPONGE spectra to maximize sensitivity to broad, shallow absorption lines. The complete source sample is large enough to break into latitude bins to investigate the variations in WNM properties with environment. This allows us to test the WF effect hypothesis: we expect stronger signal in low latitude spectra, however these are also the most difficult to model and residual absorption profiles could be contaminated by poor model subtraction. Figure 6.1 shows a preliminary stacking analysis of the full 21-SPONGE absorption line sample, with 12 sources excluded due to non-Gaussian properties. The equivalent width (EW) of the stacked

signal (45 sources) is smaller than the detection from Chapter 3 (Murray et al. 2014). This is likely due to the difference between the by-hand analysis used in Murray et al. (2014, 2015), and the automatic decomposition and radiative transfer method used in Chapter 5 to model the full data sample. For example, the T_s distribution in Chapter 5 extends to ~ 6000 K which is a factor of three higher than the maximum value detected in Murray et al. (2015). However, in Figure 6.1, we see preliminary evidence that the properties of the stacked absorption feature vary with latitude. For example, the EW of the stacked residual absorption from the high latitude sample (green) is significantly smaller than the low latitude sample (blue), which is consistent with the hypothesis that sources of Ly α radiation from the Galactic plane such as massive stars are important for exciting the 21 cm transition in the WNM, as discussed in Chapter 3. In future work, I plan to investigate these differences so that this method can be used on larger samples of lower-sensitivity absorption lines (e.g., GASKAP; Dickey et al. 2013).

(2) *Expand the sample of Galactic 21 cm absorption sources:* To measure variations in H I properties with Galactic environment, we need larger samples of absorption sources. As the

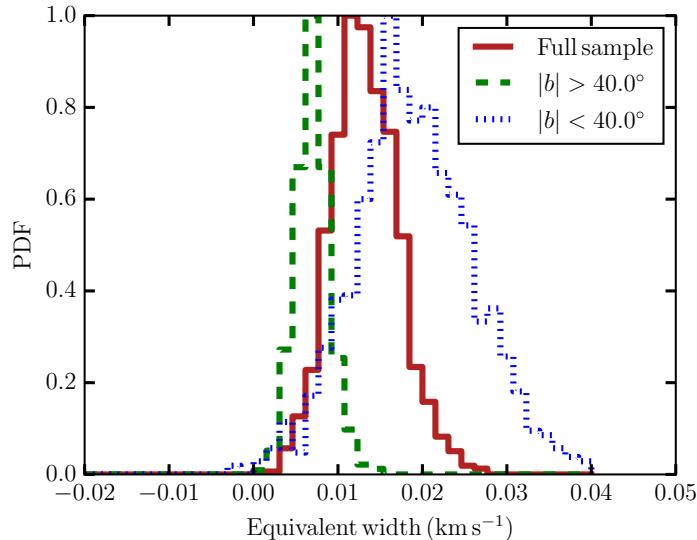


Figure 6.1 Equivalent width of stacked 21-SPONGE absorption following the methods of Chapter 3. The full sample (red) is divided into two latitude bins, high (dashed green $|b| > 40^\circ$) and low (dotted blue $|b| < 40^\circ$).

principal investigator, I am conducting a 21 cm absorption line survey of 99 compact continuum sources at high latitude using the VLA titled “Measuring Absorption by Cold Hydrogen” (MACH). Although MACH will target slightly worse sensitivity than 21-SPONGE, we will double the census of high latitude Milky Way H I absorption lines. This will improve our observational statistics for detailed comparisons with simulations: there are only ~ 100 H I absorption measurements beyond the Galactic plane (e.g., Heiles & Troland 2003; Roy et al. 2013; Murray et al. 2015), in contrast with orders of magnitude more synthetic H I spectra (e.g., 10^4 spectral pairs; Kim et al. 2014). Furthermore, MACH will be an excellent pilot study for upcoming H I surveys with next-generation interferometers (e.g., ASKAP, SKA). However, MACH will provide a unique view of H I properties in the Milky Way disk-halo interface, as future Galactic H I absorption surveys will prioritize the Galactic plane and the Magellanic System (Dickey et al. 2013).

(3) Apply analysis tools to next-generation datasets: The decomposition and radiative transfer method developed in this thesis will be extremely useful in the era of next-generation simulations and observations of 21 cm spectral line pairs. By excising the need for human input, and ensuring that the results are reproducible, I plan to use this approach to characterize structures within the Milky Way traced by multi-wavelength observations. I will begin by decomposing high-resolution H I emission data (e.g., GALFA-HI; Peek et al. 2011). Then, I will identify coherent structures using statistical clustering algorithms, and kinematically associate these structures with available H I absorption measurements (e.g., 21-SPONGE, HT03) and molecule-rich gas tracers (e.g., CO, OH, [CII]). This project will probe the volumetric extents of the underlying structures which give rise to strong absorption, and estimate their contribution to the total H I mass. I will use the same methodology to analyze synthetic data from numerical simulations to quantify the systematic limitations in tracing coherent physical structures within observed data cubes.

References

- Dickey, J. M., McClure-Griffiths, N., Gibson, S. J., et al. 2013, PASA, 30, e003
- Heiles, C., & Troland, T. H. 2003a, ApJS, 145, 329
- . 2003b, ApJ, 586, 1067
- Lindner, R. R., Vera-Ciro, C., Murray, C. E., et al. 2015, AJ, 149, 138
- Murray, C. E., Stanimirović, S., Kim, C.-G., et al. 2017, ApJ, 837, 55
- Murray, C. E., Lindner, R. R., Stanimirović, S., et al. 2014, ApJ, 781, L41
- Murray, C. E., Stanimirović, S., Goss, W. M., et al. 2015, ApJ, 804, 89
- Peek, J. E. G., Heiles, C., Douglas, K. A., et al. 2011, ApJS, 194, 20

Appendix A

21-SPONGE Calibration Techniques

*This Appendix has previously appeared
in the EVLA memo series: 171, 176, 197*

In this appendix, I describe bandpass calibration techniques developed for 21-SPONGE. These include instrument setup for calibration via frequency switching, filtering H I emission by excluding short baselines, and combining bandpass calibration observations in time. In addition, I characterize a 59 kHz “ripple” in the bandpass caused by finite impulse response filters applied prior to correlation.

A.1 Calibration Issues

Achieving high sensitivity in optical depth ($\sigma_{\tau} < 10^{-3}$ per channel) in H I absorption requires careful attention to calibration. In this section, we describe several important calibration issues.

A.1.1 Contamination by HI at $v \sim 0 \text{ km s}^{-1}$

In contrast to spectroscopy at most other radio wavelengths, it is difficult to calibrate observations at the Galactic H I target frequency, namely, at LSR velocity $v_{\text{LSR}} = 0 \text{ km s}^{-1}$ (hereafter v). Absorption and emission by ambient H I in the direction of any calibrator source will contaminate the calibration solutions because Galactic H I is present at $v = 0 \text{ km s}^{-1}$ nearly everywhere in the sky.



Figure A.1 Instrument configuration display from the OPT. We perform bandpass calibration by frequency switching. The two 0.5 MHz bandpass calibration subbands, “high” and “low”, are offset from the central HI observation subband by 1.5 MHz ($\sim 316 \text{ km s}^{-1}$) in order to avoid HI absorption lines from the strong bandpass calibrators, resolved emission in the close array configurations, and increased system temperature noise from strong HI emission.

To resolve this issue, it is possible to calibrate at off-line frequencies. By observing bandpass calibrators at frequencies offset from $v = 0 \text{ km s}^{-1}$, we can characterize the bandpass without incorporating the absorption or emission of H I located in front of the calibrators into the solutions. We describe the instrument setup and calibration strategy below.

A.1.1.1 Frequency-Switched Calibration Setup

Our observations utilize three standard L-band configurations, each with one dual polarization IF of 0.5 MHz bandwidth with 256 channels (1.95 kHz channel spacing). We use two configurations, one centered at 1.421908 GHz (1.5 MHz, or about 316 km/s, higher than the H I rest frequency, called “high”) and one at 1.418908 GHz (1.5 MHz lower, called “low”) in order to perform bandpass calibration for H I line observations.

Following the recent upgrades to the WIDAR correlator, we were able to observe with all three instrument configurations (i.e. high, low and target H I line) simultaneously. Figure A.1 displays the instrument configuration from the OPT.

For bandpass calibration, we favor the strong calibrators, 3C286 ($S_{1.4 \text{ GHz}} = 15 \text{ Jy}$), 3C147 ($S_{1.4 \text{ GHz}} = 23 \text{ Jy}$) and 3C48 ($S_{1.4 \text{ GHz}} = 16.5 \text{ Jy}$). As a C-priority project, we operate with short ($< 1 \text{ hour}$) SBs, and therefore we must maximize efficiency (and minimize slew time). If a target is bright enough ($S_{1.4 \text{ GHz}} \geq 6 \text{ Jy}$), we use it as its own bandpass calibrator. This is possible given

our small bandwidth (0.5 MHz). For a 6 Jy source, in order to reach rms noise levels of $\sim 5 \times 10^{-4}$, we need at least 3 hours of total integration time per bandpass calibrator.

A.1.2 Example: Observing Galactic H I Absorption towards 4C12.50

As an example of how we perform bandpass calibration via frequency switching to avoid contamination by H I structure at $v = 0 \text{ km s}^{-1}$, we consider an observation of the compact radio source 4C12.50 (RA=13:47:33, DEC=12:17:24; $S_{1.4 \text{ GHz}} = 5.2 \text{ Jy}$).

To measure H I absorption in the direction of 4C12.50 with an RMS noise in optical depth of $\sigma_\tau = 5 \times 10^{-4}$, we consult the exposure calculator to find that we need an on-source integration time of ~ 4.5 hours. Fortunately, 4C12.50 is a well-known calibrator source, and therefore can be used as its own complex gain calibrator throughout the observations. For bandpass calibration, we select the nearby well-known calibrator source 3C286 ($S_{1.4 \text{ GHz}} = 15 \text{ Jy}$). To achieve the required sensitivity, the exposure calculator indicates that we need at least 30 minutes of integration time on the bandpass calibrator.

With the ability to observe multiple subbands simultaneously, we observe all target and calibrator sources with the “high”, “low” and target H I subbands simultaneously as described in Section A.1.1.1.

A.1.2.1 Bandpass Calibration of 3C286 via Frequency-Switching

Following observations and initial flagging, we calibrate the bandpass using the AIPS task BPASS with observations of 3C286. Given that our bandpass calibrator is a well-modeled source, we incorporate a model of the source during calibration. We compute the bandpass solution for the “low”, target H I line, and “high” subbands and display the results in Figure A.2. These plots were constructed using AIPS task POSSM with “APARM(8) 2” and are zoomed-in to the central channels (selected channels 30 through 200 out of 256 total channels with “BCHAN 30; ECHAN 200”) to display the amplitude structure of the bandpass in more detail. We note that there is a striking “ripple” structure with period $\sim 60 \text{ kHz}$ period present in all bandpass solutions

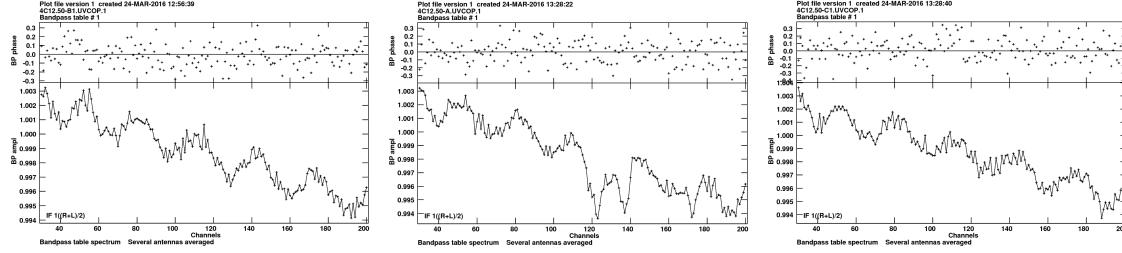


Figure A.2 Left: BP solution constructed with observations of 3C286 in the “low” subband (see Figure A.1 for the instrument setup); Center: BP solution constructed with observations of 3C286 at the target H I frequency; Right: BP solution constructed with observations of 3C286 in the “high” subband (see Figure A.1 for the instrument setup). Extraneous bandpass structure at $v = 0 \text{ km s}^{-1}$ from absorption by H I towards 3C286 is present at the H I frequency (central panel). The “ripple” present in all bandpass solutions is discussed in Section A.2.

throughout this memo. We will address this structure and its cause in Section A.2.

As shown in Figure A.2, at the H I frequency (central panel), there is extraneous structure in the bandpass solution imposed by intervening H I absorption at $v = 0 \text{ km s}^{-1}$ in comparison with the off-line frequencies (left and right panels). To avoid incorporating this structure into our target observation of source 4C12.50, we use the bandpass information from the “high” and “low” subbands (see Figure A.1 for the instrument setups), rather than data from the target H I subband.

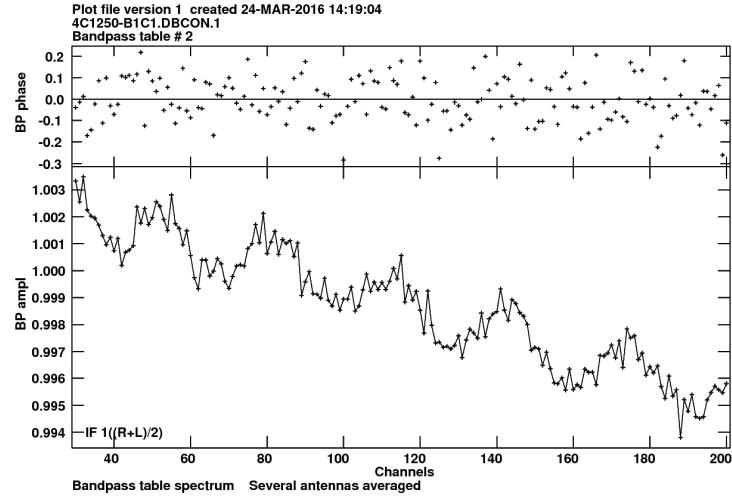


Figure A.3 BP solution constructed with observations of 3C286 after combining data from the “low” and “high” subbands (see Figure A.1 for the instrument setup).



Figure A.4 Instrument configuration display from the OPT. Testing six 0.5 MHz subbands offset from the central HI observation subband by 1.5 MHz ($\sim 316 \text{ km s}^{-1}$) each in order to avoid HI absorption lines from the strong bandpass calibrators, resolved emission in the close array configurations, and increased system temperature noise from strong HI emission.

After constructing BP solutions for the high and low subbands individually, we use the task BPLOT to examine the solution from each antenna separately and remove any noisy individual antennas. Finally, we combine the high and low subbands using the AIPS task DBCON and re-compute the BP solution with BPASS. This averaging process allows us to interpolate over

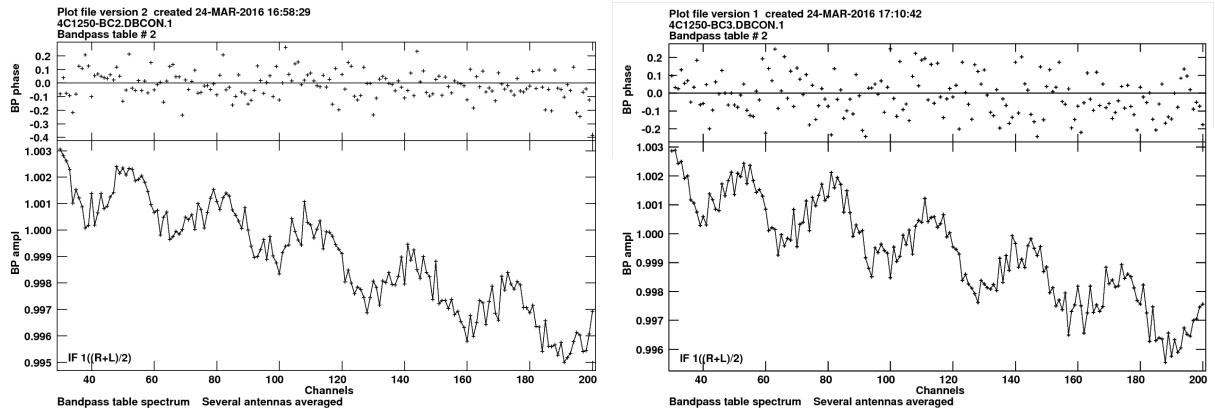


Figure A.5 Left: BP solution constructed with observations of 3C286 after combining data from subbands “low 2” and “high 2” (see Figure A.4 for the instrument setup), spaced by 3 MHz from the target H I line; Right: BP solution constructed with observations of 3C286 after combining data subbands “low 3” and “high 3” (see Figure A.4 for the instrument setup), spaced by 4.5 MHz from the target H I line. The “ripple” present in all bandpass solutions is discussed in Section A.2.

the intervening target H I frequencies. The resulting BP solution, which contains accurate phase information for the target observation, is shown in Figure A.3. This is the BP solution that we will apply to the target H I frequency observations of 4C12.50 later in the reduction process.

A.1.3 Spacing of Frequency-Switched Subbands

Given the WIDAR correlator capability of including many subbands simultaneously, we experimented with different subband spacing above and below the H I rest frequency. Figure A.4 displays a screenshot of the OPT from an example test setup, including six subbands spaced by 1.5 MHz each above and below the target H I subband.

Using the same observation of 4C12.50 with calibrator 3C286 described in Section A.1.2, we show example BP solutions constructed using various combinations of additional subbands spaced above and below the target H I line as shown in the setup from Figure A.4.

Figure A.5 shows the BP solutions from subbands spaced by 3 MHz (left) and 4.5 MHz (right) from the target H I line. By inspection, the solution does not change significantly with increased spacing from the target H I line.

Figure A.6 shows BP solutions after combining two “low” and two “high” subbands (left; 4 total subbands combined) and after combining three “low” and three “high” subbands (right; 6 total subbands combined) on either side of the target H I line. By inspection the noise in the BP solution does not improve significantly between the two examples, and also does not improve significantly from the solution presented in Figure A.3, which contains data from only two subbands.

Overall, increased subband spacing did not change the quality or stability of the calibration solutions. Furthermore, we found that combining data from multiple subbands does not significantly improve the noise in the bandpass solution.

A.1.4 Increased System Temperature

In addition to bandpass structure contamination at $v = 0 \text{ km s}^{-1}$, strong emission by H I causes the system temperature (T_{sys}) to vary with frequency according to the strength of the emission. To display this effect, in the left panel of Figure A.7 we show the scalar-averaged amplitude spectrum (with no calibration applied yet, generated with AIPS task POSSM) towards source J2136+0041 ($S_{1.4 \text{ GHz}} = 3.3 \text{ Jy}$). The structure reflects the variation of T_{sys} across the bandwidth, and ressembles the Galactic H I emission profile in this direction. For comparison, the brightness temperature spectrum in the direction of J2136+0041, from the Leiden-Argentine-Bonn (LAB) all-sky H I survey (angular resolution $\sim 36'$; Kalberla et al. 2005), is shown in the right panel of Figure A.7.

A.1.5 Resolved HI Emission

In observations of radio continuum point sources designed to detect absorption by H I, it is clear that the structure of H I emission along the line of sight is easily resolved on baselines of order ~ 100 's of meters. At a wavelength of 21 cm, this baseline length corresponds to $\sim 100 \text{ m}/0.21 \text{ m} = 0.5 \text{ k}\lambda$, or kilo-wavelengths. We show an example of this effect in Figure A.8.

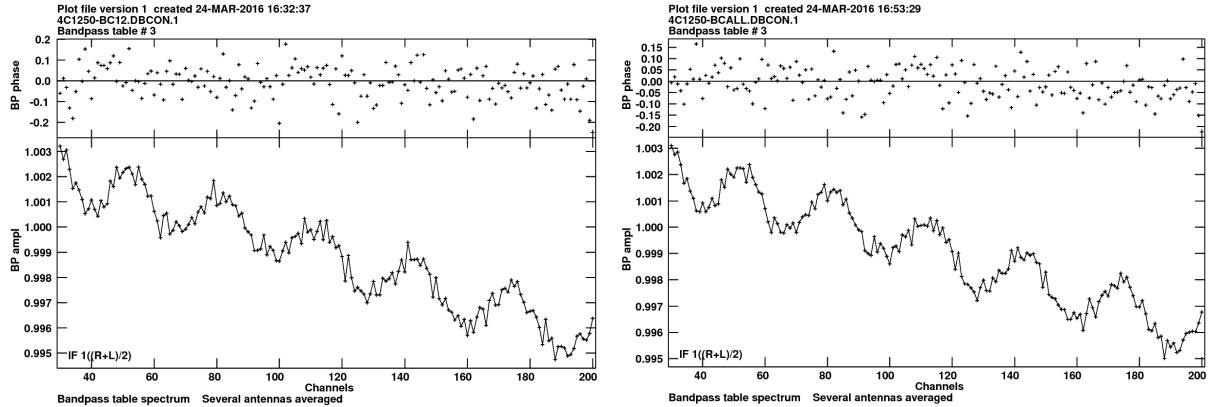


Figure A.6 Left: BP solution constructed with observations of 3C286 after combining data from two “low” and two “high” subbands (see Figure A.4 for the instrument setup); Right: BP solution constructed with observations of 3C286 after combining data from three “low” and three “high” subbands (see Figure A.4 for the instrument setup). The “ripple” present in all bandpass solutions is discussed in Section A.2.

The left panel displays a vector-averaged spectrum towards the continuum source J2136+0041, observed during a hybrid array configuration between A and D arrays (and therefore includes several baselines > 100 m in length), with no calibration applied. The right panel displays the same plot with all baselines shorter than 300 m removed, which removes the resolved emission structure from the data.

After testing a range of minimum baselines, we have adopted 300 m as a baseline cutoff to ensure that most resolved emission structure is removed from the data without compromising sensitivity from the loss of too many baselines.

The effect of resolved emission on the fully-calibrated data is not as dramatic, but is still apparent. In Figure A.9, we display two vector-averaged cross-power spectra of the same source, J2136+0041, following calibration with (left) and without (right) all baselines shorter than 300 m included. The depth of the absorption line at channel 100 in particular is significantly affected by resolved emission structure from the short baselines.

To remove data from baselines shorter than 300 m, we used the AIPS task UVCOP with setting “UVRAN 1.4 0” to select all baselines longer than $300 \text{ m} = 1.4 \text{ k}\lambda$.

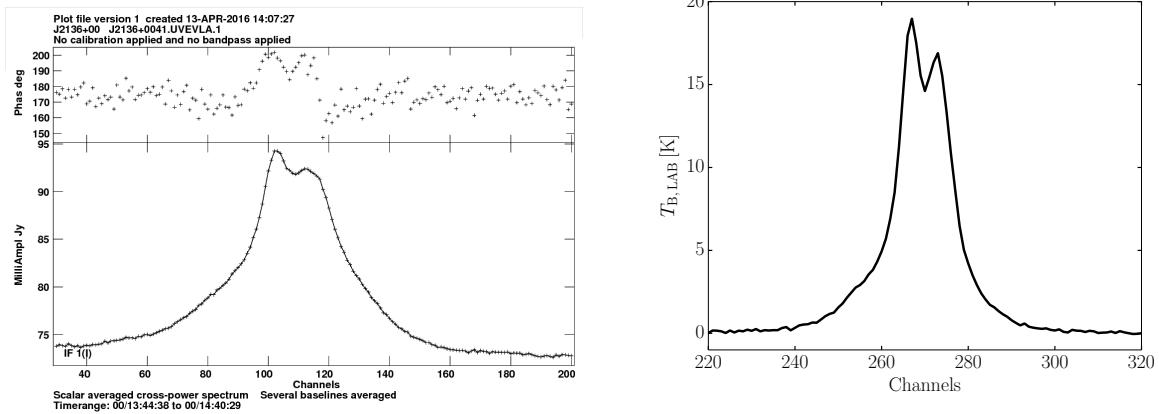


Figure A.7 Left: Scalar averaged cross-power spectrum at the H I frequency in the direction of source J2136+0041 with no calibration applied, constructed with the AIPS task POSSM and exported with LWPLA. Right: Brightness temperature spectrum towards source J2136+0041 from the LAB survey [2].

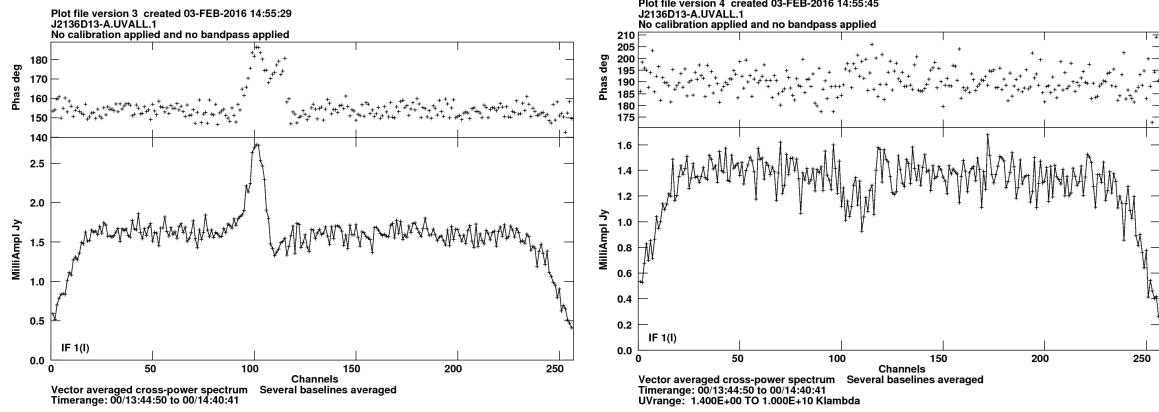


Figure A.8 Left: Vector averaged cross-power spectrum at the H I frequency in the direction of source J2136+0041 ($S_{1.4\text{GHz}} = 3.3\text{Jy}$) with no calibration applied. Right: same as left, with baselines shorter than 300 m (1.4 kilo-wavelengths) removed. Both plots constructed with the AIPS task POSSM and exported with LWPLA.

A.2 59 kHz Ripple

We find that all bandpass solutions obtained by 21-SPONGE contain a sinusoidal “ripple” with a period of 59 kHz (or FWHM $\Delta v \sim 12\text{ km/s}$ at our velocity resolution) and amplitude of ~ 0.0015 . The period, amplitude and phase of this ripple are constant between array configurations and time of year. After investigating the stability of the ripple, we determined that it is caused by

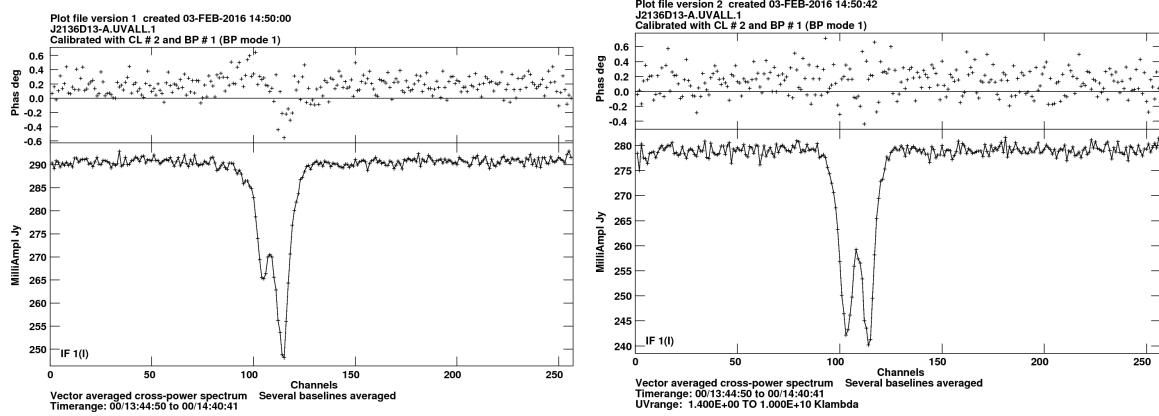


Figure A.9 Left: Vector averaged cross-power spectrum at the H I frequency in the direction of source J2136+0041 ($S_{1.4\text{GHz}} = 3.3\text{Jy}$) with bandpass and gain calibration applied. Right: same as left, with baselines shorter than 300 m (1.4 kilo-wavelengths) removed. Both constructed with the AIPS task POSSM and exported with LWPLA.

finite impulse response (FIR) filters in the WIDAR correlator applied prior to correlation which are used to shape the bandpass. The constant relative amplitude, phase and period of the ripple can all be explained in this context.

To estimate the quantitative properties of the ripple, we first fit and remove the linear bandpass slope and zoom in on the central channels of the bandpass solution (channels 50-200 out of a total 256, shown in Figure A.10b). The linear slope does not affect our ability to detect wide absorption lines, and we are only interested in determining the parameters of the ripple and correcting for it.

We then compute the Fourier transform (FT) spectrum of this flattened, zoomed solution (see Figure A.10a). Next, we multiply this FT spectrum with a simple step function filter to isolate the dominant periodic components. The boundaries of the filter are shown in Figure A.10a as vertical dashed lines. We then compute the inverse FT of the filter-multiplied FT spectrum (i.e., the thick black line in Figure A.10a) to model the periodic component. The modeled signal is displayed by the solid black line in Figure A.10b, along with the original slope-subtracted bandpass

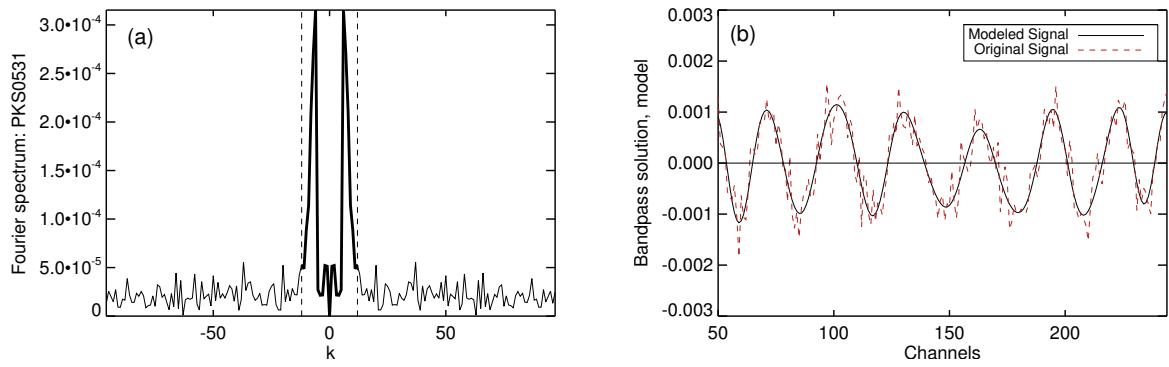


Figure A.10 (a): FT spectrum of a bandpass solution, with the boundaries of the applied step-function filter overlaid as dashed lines. (b): Model of the bandpass solution (solid black), computed as the inverse FT of the filter-convolved FT spectrum of the bandpass solution. The bandpass solution is overlaid as the dashed red line.

solution (dashed red line).

As a test of the success of the model, we compute the residuals by subtracting the model from the bandpass solution (see Figure A.11a). We then compute a histogram of these residuals to observe if they are well-represented by a Gaussian distribution, which we expect if the model is reasonable (see Figure A.11b). In this example, a K-S test between the residuals and a Gaussian distribution returns $p=0.55$, indicating that the noise is consistent with being drawn from a Gaussian distribution. From this, we can conclude that the model does a reasonably good job of isolating the dominant periodic components of the bandpass solution. We compute the RMS noise in these residuals as an indicator of the noise level in the bandpass solution that is not due to the periodic components. We record various parameters associated with the ripple fit in Table A.1 for all bandpass solutions that we tested, including RMS noise in the model residuals, ripple amplitude, period and phase. The amplitude, period and phase of the ripple for each observation were calculated using a periodogram, following the methods of Horne & Baliunas (1986).

To test if the presence of the ripple is a function of particular antennas, we separated the

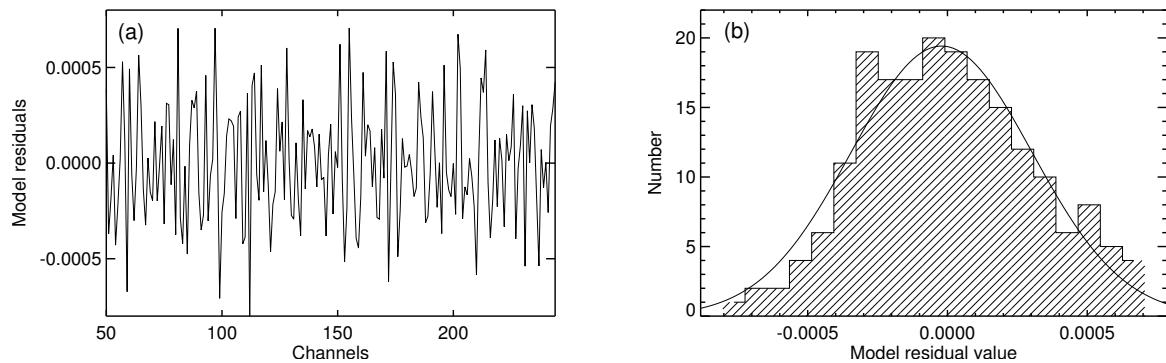


Figure A.11 (a): Residuals of the FT model to the bandpass solution displayed in Figure A.10 (i.e., bandpass solution minus model). (b): Histogram of the residuals to check if they are well-represented by a Gaussian distribution, as would be expected in the case of a successful model. A Gaussian model is overlaid (solid black line) to illustrate this.

bandpass solutions into three groups of about 9 antennas (i.e antennas 1-9, 10-18, 19-28). After fitting the solutions from these antenna subsets, we find that the ripple exists at a similar power as in the full solution in all cases. Therefore, we conclude that the effect is not antenna-based.

The ripple is caused by finite impulse response filters applied prior to correlation which shape the bandpass, and is stable in time (the period, amplitude and phase are constant between all examined solutions acquired all 4 years of observations) so that we are able to model and

Table A.1. Bandpass Solution Parameters

Information:		Fit Parameters:				
Target Name	Obs. Date	Array Config.	RMS noise ^a (per channel)	Amplitude ^b	Period ^b (channels)	Phase ^c (channel)
3C120D1	03/11	B	0.00049	0.0014	30.1	120
3C120D2	03/11	B	0.00047	0.0014	29.9	120
3C286	05/11	B	0.00064	0.0016	29.9	122
3C225BD1	05/11	BnA	0.00049	0.0015	30.0	121
3C225BD2	05/11	BnA	0.00049	0.0015	30.3	121
3C225BD3	05/11	BnA	0.00042	0.0013	30.1	120
4C12.50	05/11	BnA	0.00048	0.0014	30.0	120
3C345	06/11	BnA-A	0.00074	0.0017	30.1	120
3C298	06/11	BnA-A	0.00057	0.0015	29.8	121
4C32.44D1	05/11	BnA-A	0.00040	0.0014	29.9	122
4C32.44D2	06/11	BnA-A	0.00063	0.0015	29.9	120
4C32.44D3	06/11	BnA-A	0.00057	0.0015	29.8	122
3C48	08/11	A	0.00048	0.0014	30.0	120
4C16.09	09/11	A	0.00037	0.0014	30.0	120
4C16.09.2	09/11	A	0.00031	0.0013	29.9	120
3C133	09/11	A-D	0.00035	0.0014	30.1	120
PKS0531	09/11	A-D	0.00031	0.0013	29.9	120
3C111	04/12	C	0.00042	0.0014	29.9	120
3C154D1	04/12	C	0.00057	0.0017	30.2	122
3C154D2	04/12	C	0.00033	0.0013	29.9	120
3C123	05/12	CnB	0.00041	0.0015	30.1	120
3C138	05/12	CnB	0.00041	0.0014	29.9	122
3C410	07/12	B	0.00057	0.0016	29.6	123

Note. — ^a: the RMS noise calculated in the residuals of the bandpass solution after removing the ripple model. ^b: calculated using the periodogram analysis of the solution (i.e. these are the parameters of the dominant component) ^c: starting channel of a new period nearest to the center of the solution. .

remove it after combining the observations.

Figure A.12 displays the filter plus mixer response required to achieve the target H I sky frequency (courtesy K. Sowinski). The aforementioned ripple, caused by the 0.5 MHz filter, along with a small slope from the 8 MHz filter, is clearly evident.

A.3 Time Averaging

To minimize noise levels in bandpass (BP) solutions, we combine bandpass observations acquired at different times. To compare noise levels in the BP solutions from the original

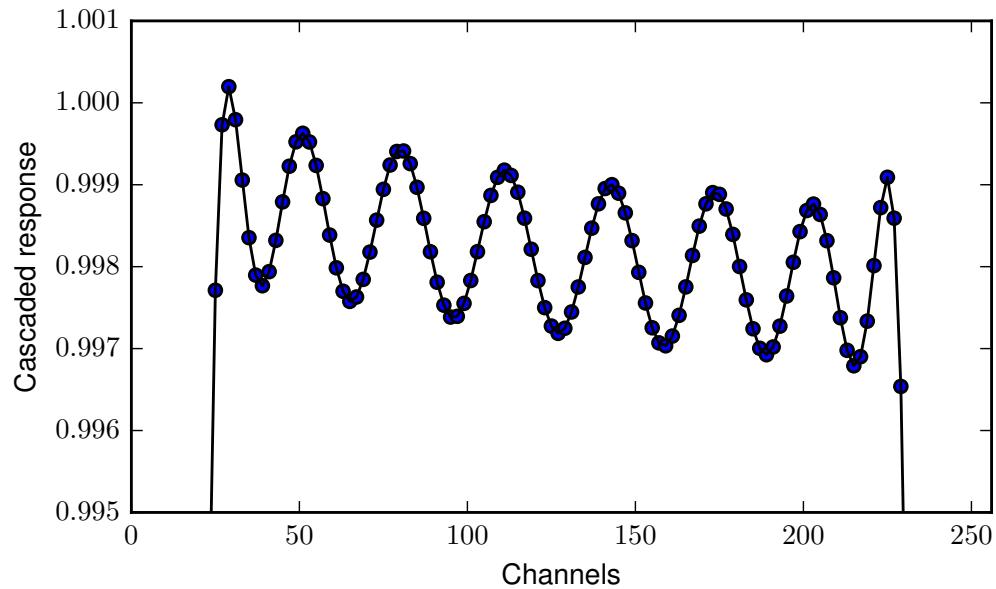


Figure A.12 Cascaded response of all the filters and mixer needed to achieve the desired sky frequency and bandwidth for the subband containing the H I line (i.e. the “target” subband in Figure A.1) vs. channels for a 21-SPONGE observation conducted on 2013/07/27. The 59 kHz ripple of the 0.5 MHz filter is evident, with a small slope from the 8 MHz filter (data courtesy K. Sowinski).

observation and the combined observations, we first model and remove the 59 kHz ripple present in all solutions. Figure A.13 shows an example of the modeling process. A zoom-in onto a bandpass solution from one observation of 3C147 (33 min integration) is shown in the top left panel, with the ripple model overlaid in red. The top right panel displays the residuals following the removal of the ripple model, from which the noise in the bandpass solution is computed. The bottom panels are the same, after adding an additional observation of 3C147 from 4 days later (for a total of $33+30=63$ min integration).

We expect the noise in the bandpass solution to decline according to the factor of increase in integration time given by the ideal radiometer equation, or by $1/\sqrt{t_2/t_1}$. For example, from Figure 15, for an increase in integration time of $t_2/t_1 = 63/33 = 1.9$, we expect an improvement in noise by a factor of $1/\sqrt{t_2/t_1} = 0.72$. The improvement in noise is $0.00055/0.00073 = 0.75$.

In Figure A.14 we show the RMS noise in the bandpass solution (computed after removing the ripple) as a function of t_2/t_1 for several different observations (symbols). The solid line displays a fit to the data points, which is linear in log space with a slope of -0.44 ± 0.02 . This is nearly equal to the theoretically expected fit index of -0.5 (e.g. a noise decrease by factor of $1/\sqrt{t_2/t_1}$). This result is highly encouraging, as it demonstrates that VLA bandpass solutions at this resolution from different days are stable enough to be combined to increase the sensitivity of observations while significantly saving observing time, which is especially useful for very deep observations.

In this process, we are able to combine observations of different sources. In Figure A.14, the green triangles are from a trial combining observations of the sources 3C410 ($S_{1.4\text{GHz}} = 10\text{ Jy}$) and 3C454.3 ($S_{1.4\text{GHz}} = 11\text{ Jy}$) as bandpass calibrators. The other three cases are combinations of observations of the same bandpass calibrators, 3C147 ($S_{1.4\text{GHz}} = 23\text{ Jy}$) and 3C286 ($S_{1.4\text{GHz}} = 15\text{ Jy}$) between several days. These examples represent a common trend in our results.

In later observations, we experimented with including additional 0.5 MHz subbands spaced ± 1.5 , 3 and 4.5 MHz from the HI line. We found that combining these subbands does not

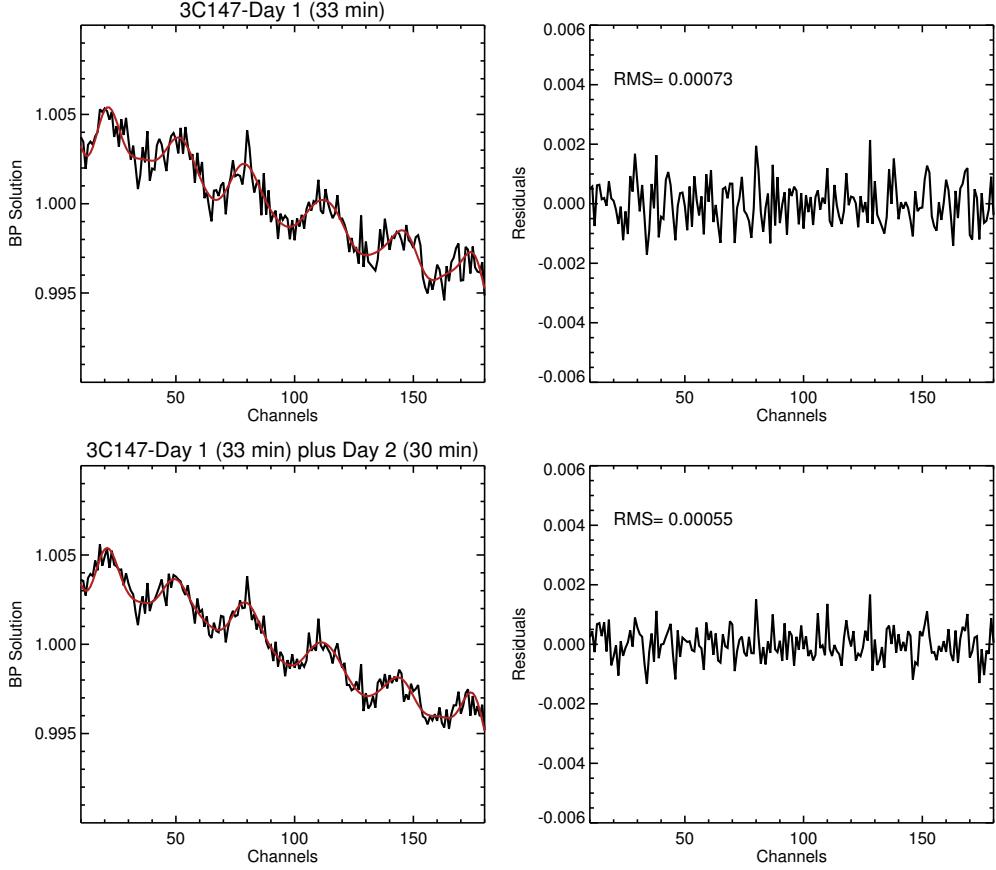


Figure A.13 An example of combining bandpass solutions. We fit the bandpass solutions (left) for the presence of a 59 kHz ripple, remove this model (shown in red) to produce the panels on the right, from which we measure the RMS noise in the solution. The top panels are from a single observation, and the bottom panels are the results of adding an observation 4 days later.

significantly improve the noise in the bandpass solution.

References

Horne, J. H., & Baliunas, S. L. 1986, ApJ, 302, 757

Kalberla, P., Burton, W., Hartmann, D., et al. 2005, A&A, 440, 775

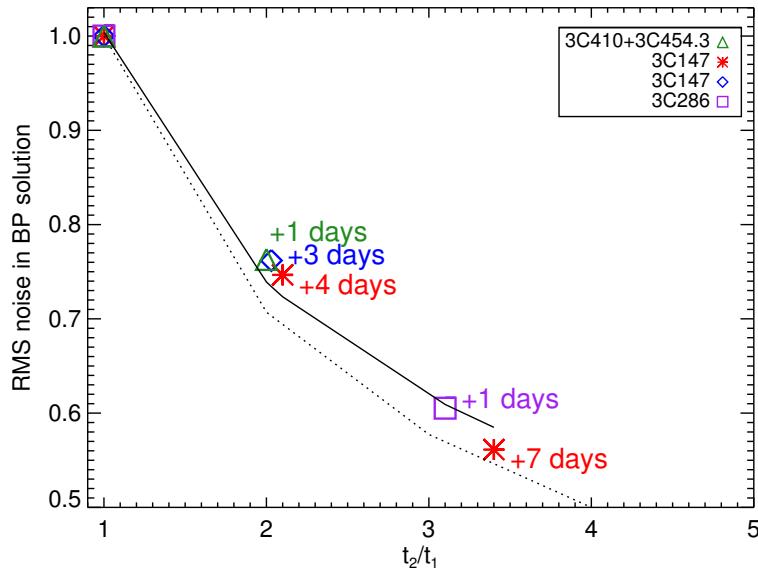


Figure A.14 The improvement in RMS noise in a bandpass solution (y-axis) by combining observations in time, thereby increasing the total integration time (x-axis). Symbols denote measurements, and the solid line is a linear fit to the data. The index of the fit is -0.44 ± 0.02 , compared to the theoretical expectation of an index equal to -0.5 (e.g. noise decrease by factor of $1/\sqrt{t_2/t_1}$, shown by the dotted line).

Appendix B

21-SPONGE Continuum Images and Fitted Spectral Line Properties

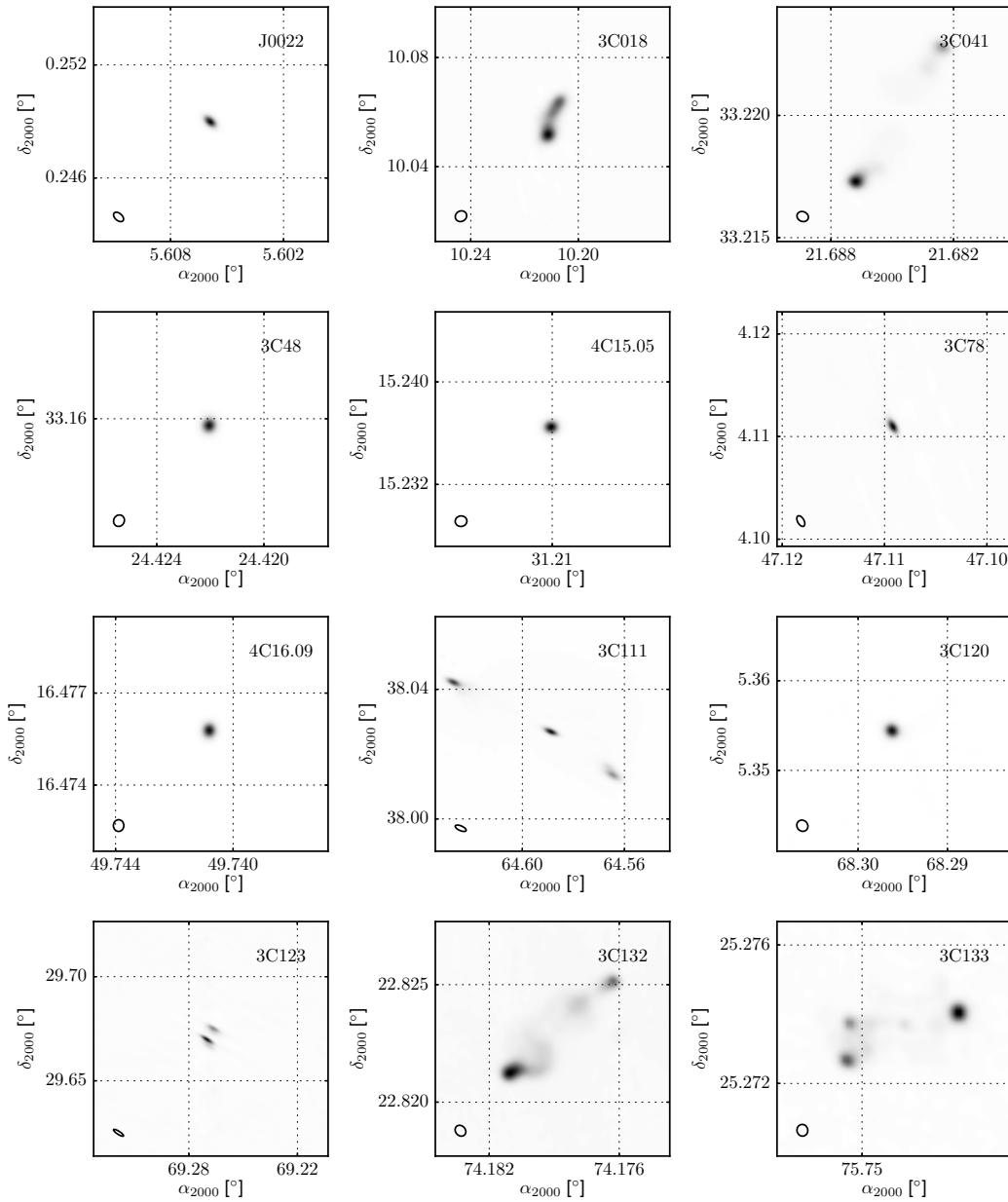
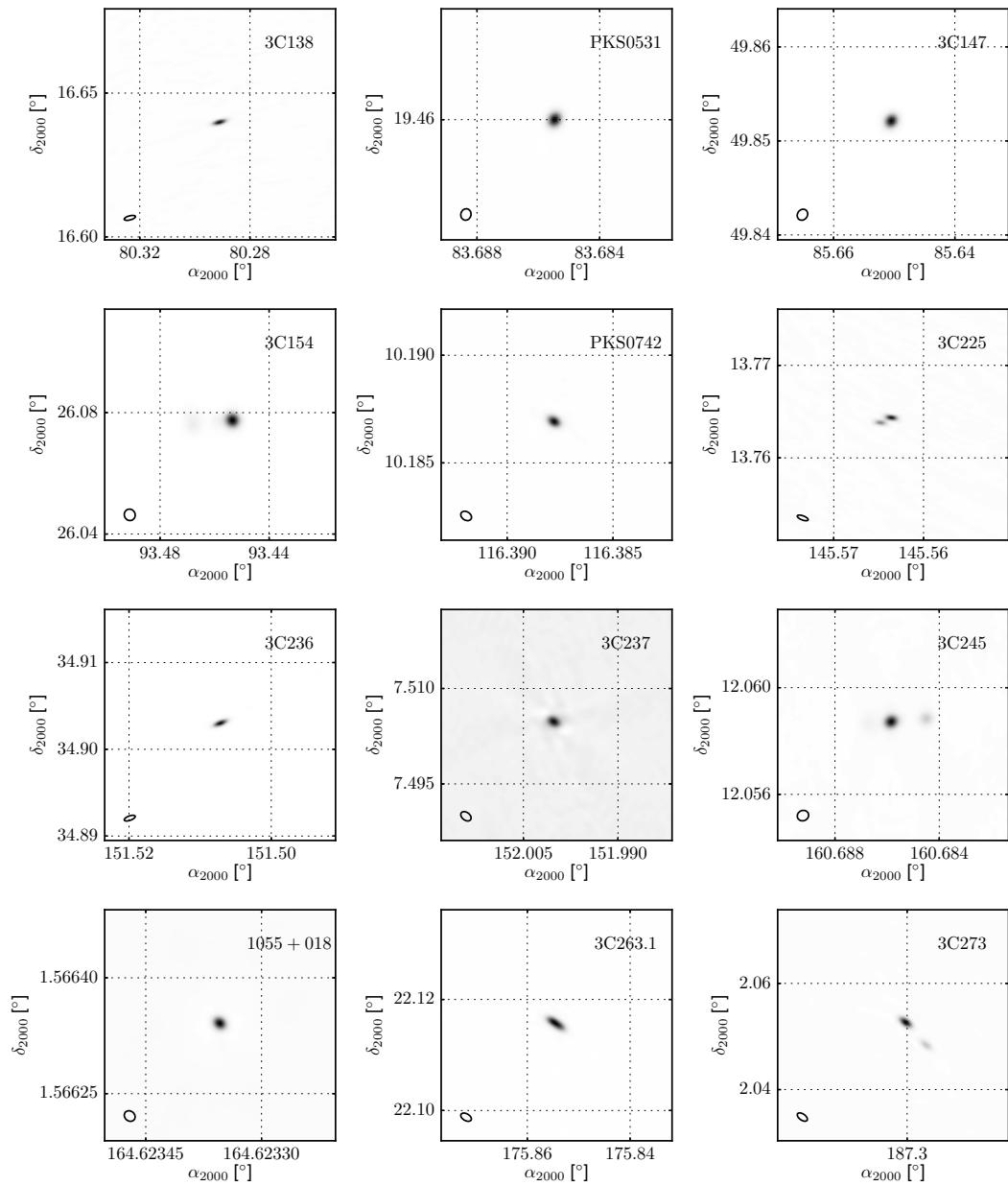
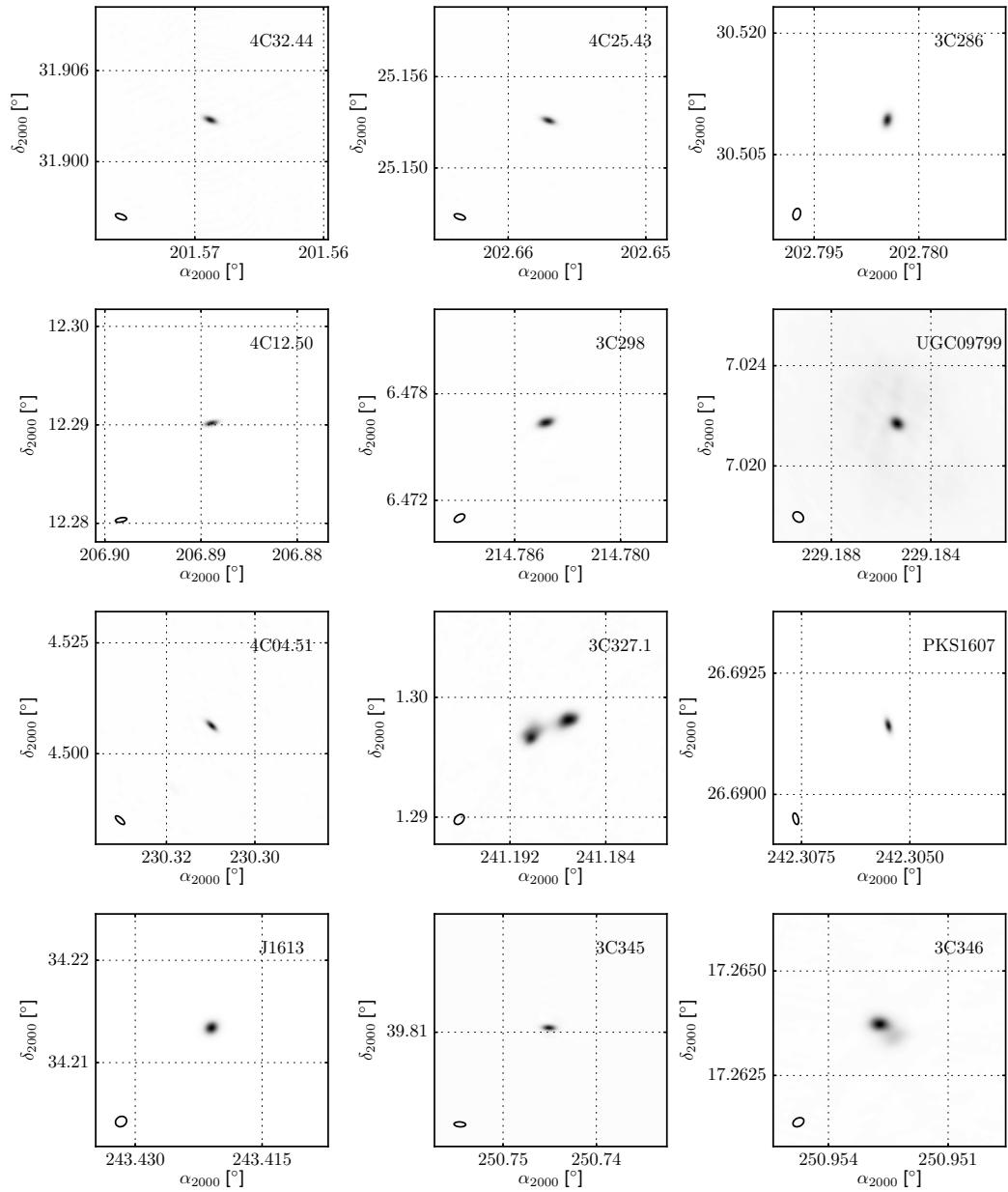


Figure B.1 21SPONGE 1.42040575 GHz continuum images. The source name is printed within each panel, and the synthesized beam used to construct each image is included in the bottom left corner. Each image is scaled so that the peak flux density is unity.

**Figure B.0** (contd.)

**Figure B.0** (contd.)

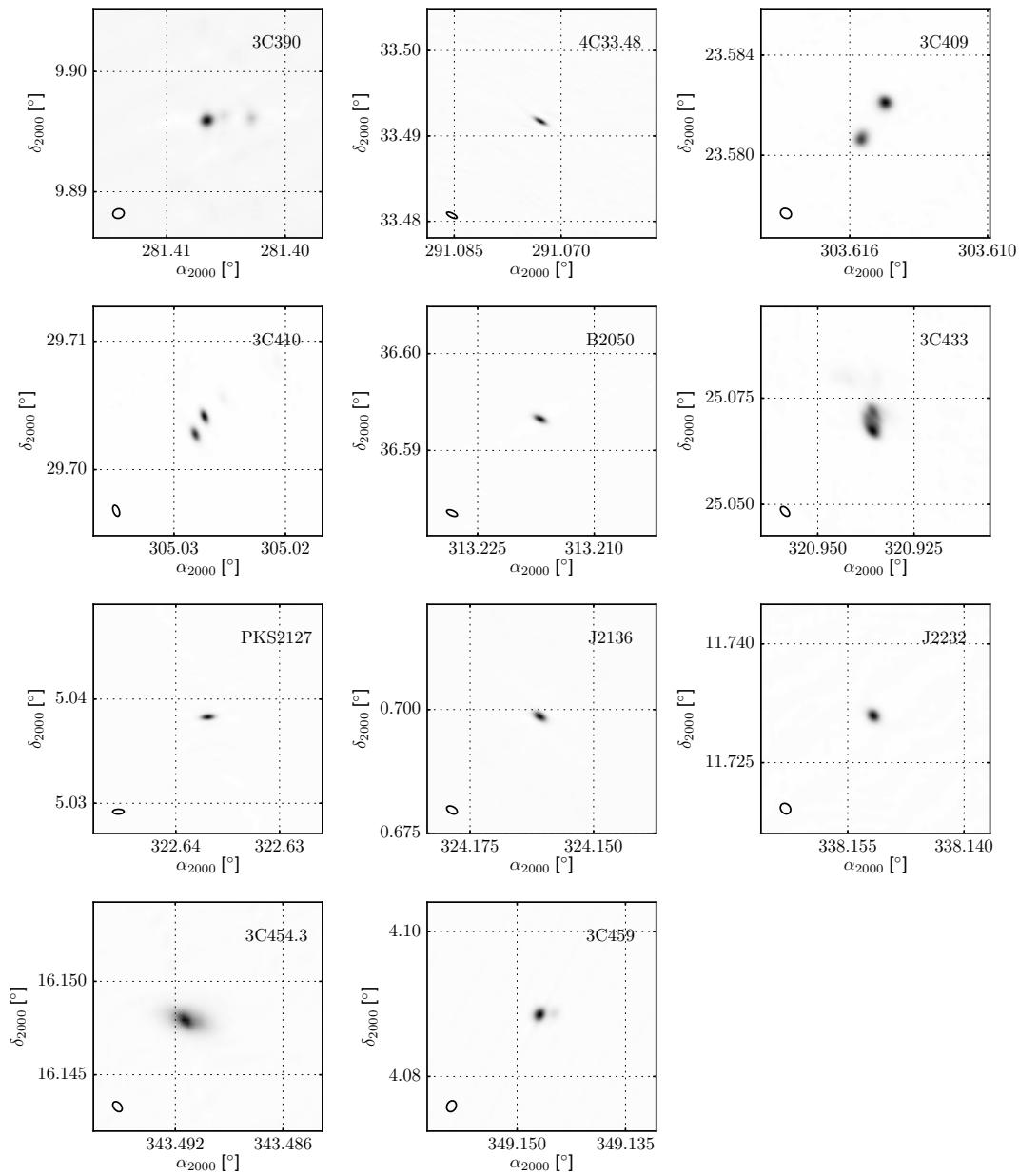


Figure B.0 (contd.)

Table B.1. Fitted Parameters

Source (name)	τ_0	Δv_0 (km s $^{-1}$)	v_0 (km s $^{-1}$)	$T_{B,n}$ (K)	$\Delta v_{0,n}$ (km s $^{-1}$)	$v_{0,n}$ (km s $^{-1}$)	T_s (K)	N(HI) (10 20 cm $^{-2}$)	\mathcal{O}	\mathcal{F}
J0022	0.018±0.000	2.8±0.1	-7.8±0.0	0.9±0.0	3.4±0.1	-9.5±0.0	49±1	0.05±0.00	3	1.0
	0.008±0.000	10.4±0.3	-4.3±0.2	5.0±0.1	12.6±0.3	-4.1±0.2	643±21	1.02±3.64	3	
	0.009±0.000	1.7±0.1	0.3±0.0	0.2±0.0	1.8±0.1	0.4±0.0	24±1	0.01±0.00	3	
	0.003±0.000	1.0±0.2	-14.6±0.1	
3C018A	0.565±0.007	2.5±0.0	-9.1±0.0	5.5±0.1	3.0±0.0	-10.3±0.0	14±2	0.39±0.00	2	1.0
	0.134±0.003	5.5±0.2	-6.2±0.1	23.1±0.6	6.0±0.2	-5.5±0.1	193±9	2.83±1.30	2	
	0.084±0.003	1.5±0.1	-5.0±0.0	
	0.007±0.003	0.7±0.3	24.4±0.1	
3C018B	0.524±0.003	2.4±0.0	-9.0±0.0	3.1±0.9	2.7±0.3	-10.3±0.3	5±2	0.14±0.00	2	1.0
	0.149±0.002	6.2±0.1	-6.8±0.0	22.0±0.5	7.0±0.1	-6.7±0.0	157±5	2.86±0.74	1	
3C041A	0.033±0.000	8.9±0.1	-1.4±0.1	12.8±0.2	9.5±0.1	-1.4±0.1	393±6	2.28±2.30	2	1.0
	0.010±0.001	1.2±0.2	1.8±0.1	0.9±0.1	1.5±0.2	2.0±0.1	92±13	0.02±0.01	2	
	0.029±0.001	1.7±0.1	-30.5±0.0	
3C041B	0.022±0.002	1.7±0.2	-10.5±0.1	0.4±0.0	2.0±0.2	-10.2±0.1	19±3	0.01±0.00	2	1.0
	0.042±0.001	7.1±0.2	-1.4±0.1	4.0±0.1	8.6±0.2	-1.3±0.1	99±3	0.59±0.17	2	
	0.023±0.003	0.8±0.1	-8.0±0.0	
3C48	0.004±0.000	0.7±0.1	-55.1±0.0	0.1±0.0	0.8±0.1	-57.7±0.0	29±4	0.00±0.00	6	1.0
	0.008±0.000	5.5±0.3	-15.6±0.1	3.7±0.1	6.6±0.3	-15.0±0.1	493±13	0.40±0.48	6	
	0.016±0.000	2.8±0.1	-11.0±0.0	2.2±0.0	3.4±0.1	-11.1±0.0	137±3	0.12±0.02	6	
	0.004±0.000	1.3±0.2	-5.0±0.1	0.9±0.1	1.5±0.2	-4.9±0.1	230±27	0.02±0.01	6	
	0.020±0.000	9.2±0.1	-2.0±0.1	8.2±0.1	10.1±0.1	-2.0±0.1	426±4	1.51±0.88	1	
	0.040±0.000	2.3±0.0	2.6±0.0	2.1±0.0	2.8±0.0	2.8±0.0	55±2	0.10±0.00	6	
4C15.05	0.051±0.001	2.4±0.0	-10.5±0.0	2.1±0.0	2.9±0.0	-11.3±0.0	43±2	0.11±0.00	2	0.0
	0.034±0.002	7.8±0.4	-9.4±0.2	10.5±0.6	9.4±0.4	-8.9±0.2	314±21	1.65±4.48	1	
	0.060±0.002	3.1±0.0	-5.3±0.0	3.4±0.1	3.1±0.0	-5.3±0.0	59±3	0.22±0.02	6	
	0.004±0.002	13.9±2.0	-3.5±2.6	8.7±3.9	13.9±2.0	-3.8±2.6	2236±990	2.41±99.	6	
	0.010±0.001	1.8±0.1	-1.9±0.1	1.9±0.1	2.2±0.1	-1.9±0.1	192±18	0.07±0.04	6	
	0.018±0.001	2.4±0.1	0.2±0.1	2.2±0.1	2.9±0.1	0.2±0.1	124±5	0.11±0.02	6	
3C78	0.148±0.002	3.1±0.1	-7.8±0.0	10.3±0.2	3.1±0.1	-7.8±0.0	75±4	0.70±0.04	2	1.0
	0.027±0.001	16.3±0.9	-7.2±0.3	11.7±0.7	16.6±0.9	-7.1±0.3	448±25	3.84±30.50	1	
	0.195±0.002	1.8±0.0	4.2±0.0	11.0±0.1	2.1±0.0	4.4±0.0	63±2	0.43±0.01	6	
	1.370±0.002	2.2±0.0	6.9±0.0	38.3±0.1	2.7±0.0	7.2±0.0	56±4	3.41±0.03	6	
	0.118±0.011	2.0±0.1	10.4±0.1	4.0±0.5	2.0±0.1	10.7±0.1	42±6	0.20±0.03	6	
	0.058±0.006	2.7±0.4	12.4±0.3	10.2±1.0	3.3±0.4	12.1±0.3	184±21	0.58±0.55	6	
4C16.09	0.014±0.001	1.3±0.1	-9.7±0.0	0.4±0.0	1.6±0.1	-11.7±0.0	28±5	0.01±0.00	5	1.0
	0.436±0.002	4.2±0.0	-1.0±0.0	39.3±0.2	4.1±0.0	-1.0±0.0	117±4	4.23±0.15	5	
	0.227±0.004	1.8±0.0	0.8±0.0	7.2±0.2	1.8±0.0	0.9±0.0	61±11	0.50±0.02	5	
	0.086±0.001	5.4±0.1	5.9±0.0	23.6±0.3	5.9±0.1	5.7±0.0	288±3	2.67±0.89	5	
	0.042±0.001	1.8±0.1	7.6±0.0	4.9±0.2	1.9±0.1	7.5±0.0	133±9	0.19±0.03	5	
	0.013±0.001	4.1±0.2	-8.7±0.1	
3C111A	0.040±0.001	2.7±0.1	-22.0±0.0	2.8±0.1	3.3±0.1	-21.1±0.0	70±2	0.15±0.01	8	1.0
	0.144±0.008	3.4±0.3	-10.8±0.1	27.6±1.6	3.6±0.3	-11.4±0.1	208±11	2.01±1.46	8	
	0.251±0.013	2.8±0.1	-8.1±0.1	26.6±1.6	2.8±0.1	-8.1±0.1	122±7	1.70±0.60	2	
	0.464±0.003	2.4±0.0	-4.3±0.0	28.0±0.2	2.9±0.0	-4.3±0.0	78±4	1.76±0.05	8	
	0.539±0.024	1.6±0.0	-2.6±0.0	15.9±1.3	1.7±0.0	-2.7±0.0	53±10	0.88±0.06	8	
	0.820±0.008	2.0±0.0	-0.9±0.0	37.8±0.6	2.4±0.0	-0.9±0.0	75±5	2.38±0.10	8	
	0.398±0.029	1.5±0.0	1.0±0.0	13.4±1.8	1.6±0.0	1.0±0.0	63±13	0.74±0.10	8	
	0.066±0.002	1.2±0.0	8.4±0.0	0.5±0.0	1.4±0.0	10.2±0.0	5±1	0.01±0.00	8	
	0.119±0.008	4.3±0.2	-15.6±0.2	
	0.038±0.001	2.4±0.1	-32.2±0.0	
	0.252±0.011	2.2±0.0	-17.0±0.0	
	0.233±0.001	5.1±0.1	6.2±0.1	

Table B.1 (cont'd)

Source (name)	τ_0	Δv_0 (km s $^{-1}$)	v_0 (km s $^{-1}$)	$T_{B,n}$ (K)	$\Delta v_{0,n}$ (km s $^{-1}$)	$v_{0,n}$ (km s $^{-1}$)	T_s (K)	N(HI) (10 20 cm $^{-2}$)	\mathcal{O}	\mathcal{F}
	0.571±0.013	2.9±0.1	2.2±0.0
3C111B	0.019±0.002	1.7±0.2	-54.8±0.1	0.5±0.0	2.1±0.2	-56.3±0.1	25±3	0.02±0.00	1	1.0
	0.090±0.002	11.8±0.4	-11.2±0.4	29.8±0.8	11.7±0.4	-12.1±0.4	350±9	7.38±14.48	7	
	0.405±0.006	2.8±0.0	-8.3±0.0	6.1±0.1	3.4±0.0	-8.3±0.0	24±7	0.56±0.01	7	
	0.981±0.021	3.8±0.1	-1.7±0.1	27.1±1.2	3.7±0.1	-1.5±0.1	56±9	4.12±0.38	7	
	0.294±0.073	2.4±0.2	2.0±0.0	12.7±5.9	2.9±0.2	2.0±0.0	84±22	1.16±1.12	7	
	0.409±0.037	4.8±0.5	3.5±0.4	33.1±3.7	5.7±0.5	4.2±0.4	102±11	3.94±3.40	2	
	0.242±0.007	2.3±0.0	7.9±0.0	1.5±0.1	2.8±0.0	7.9±0.0	9±3	0.11±0.00	7	
	0.188±0.005	0.8±0.0	-2.3±0.0	
	0.015±0.002	2.2±0.3	-51.3±0.1	
	0.110±0.002	2.6±0.1	-16.8±0.0	
3C111C	0.153±0.003	9.8±0.5	-10.2±0.3	31.0±0.8	9.7±0.5	-10.1±0.3	235±16	7.00±7.03	6	1.0
	0.307±0.012	1.8±0.0	-5.5±0.0	3.5±0.2	1.8±0.0	-5.5±0.0	18±10	0.21±0.01	2	
	1.082±0.008	2.7±0.0	-2.1±0.0	23.1±0.3	3.3±0.0	-2.1±0.0	42±9	2.49±0.06	6	
	1.015±0.008	2.6±0.0	1.1±0.0	25.1±0.4	2.9±0.0	0.9±0.0	47±9	2.47±0.06	6	
	0.182±0.010	1.2±0.1	3.2±0.0	1.9±0.2	1.2±0.1	2.8±0.0	17±10	0.08±0.00	6	
	0.273±0.003	4.5±0.1	4.8±0.1	28.6±0.3	5.4±0.1	4.7±0.1	110±12	2.69±0.30	6	
	0.029±0.002	2.6±0.3	-27.8±0.1	
	0.456±0.008	1.7±0.0	-8.4±0.0	
	0.027±0.002	2.4±0.3	-31.4±0.1	
	0.298±0.003	2.5±0.0	-16.7±0.0	
3C120	0.052±99.	0.1±99.	-4.7±99.	
	0.188±0.007	1.8±0.1	3.8±0.0	0.0±1.2	2.1±2.8	3.7±0.2	4±11	0.03±0.00	2	1.0
	1.762±0.011	2.4±0.0	6.3±0.0	28.8±4.2	2.9±0.2	5.3±0.0	45±12	3.88±0.12	1	
	1.634±0.006	3.3±0.0	9.1±0.0	48.9±0.4	3.8±0.0	9.9±0.0	62±2	6.67±0.14	3	
3C123A	0.064±0.001	3.1±0.0	-19.9±0.0	1.7±0.0	3.8±0.0	-21.1±0.0	30±4	0.12±0.00	6	1.0
	0.379±0.035	2.2±0.1	1.6±0.0	3.8±0.8	2.2±0.1	1.9±0.0	21±19	0.36±0.03	6	
	0.044±0.004	10.1±0.3	3.7±0.1	31.4±3.2	12.2±0.3	3.9±0.1	705±100	6.24±88.47	6	
	0.810±0.071	4.4±0.1	5.3±0.2	26.8±8.0	5.3±0.1	5.2±0.2	117±52	8.33±7.35	6	
	0.628±0.037	1.8±0.0	5.5±0.0	5.0±0.5	1.7±0.0	6.3±0.0	12±16	0.28±0.01	6	
	0.008±0.000	4.0±0.3	20.2±0.1	1.1±0.1	4.0±0.3	17.9±0.1	144±16	0.09±0.06	6	
	1.020±0.096	2.3±0.0	3.7±0.0	
3C123B	0.064±0.001	3.2±0.1	-19.3±0.0	1.2±0.0	3.4±0.1	-20.6±0.0	23±3	0.10±0.00	6	1.0
	0.018±0.002	14.1±0.8	3.3±0.5	20.5±2.6	15.5±0.8	3.7±0.5	1161±159	5.84±99.	6	
	1.676±0.003	4.4±0.0	4.4±0.0	33.5±0.2	5.3±0.0	4.3±0.0	52±10	7.66±0.23	6	
	0.242±0.004	1.4±0.0	5.5±0.0	7.1±0.4	1.7±0.0	5.8±0.0	109±49	0.75±0.04	6	
	0.068±0.006	1.5±0.1	8.5±0.0	2.5±0.3	1.5±0.1	6.9±0.0	43±68	0.09±0.01	6	
	0.017±0.002	3.0±0.9	10.4±0.5	5.5±0.7	3.6±0.9	9.2±0.5	274±69	0.27±0.64	6	
3C132	0.018±0.007	0.8±2.3	0.8±0.0	0.1±0.1	0.8±2.3	0.9±0.0	20±19	0.01±0.00	5	1.0
	0.156±0.025	6.2±0.0	1.4±99.	5.6±1.8	6.2±0.0	1.4±99.	71±25	1.38±1.96	5	
	0.070±0.021	11.6±0.0	4.7±0.1	45.4±13.8	11.5±0.0	4.2±0.1	665±198	10.72±99.	5	
	1.446±0.011	1.9±0.0	7.8±0.1	33.2±0.8	2.2±0.0	7.9±0.1	72±18	3.96±0.10	5	
	0.357±0.011	1.8±0.3	9.3±0.0	21.6±0.9	2.1±0.3	9.6±0.0	80±9	1.01±0.13	5	
	0.231±0.015	5.7±0.2	13.2±1.9	
	0.151±0.004	1.9±0.2	1.9±0.1	
3C133	0.035±0.002	9.2±0.4	-29.7±0.2	16.0±1.1	9.1±0.4	-29.8±0.2	471±32	3.04±17.09	5	1.0
	0.082±0.003	3.1±0.1	-27.8±0.0	2.0±0.1	3.0±0.1	-26.7±0.0	29±4	0.15±0.01	5	
	0.206±0.011	6.5±0.6	-0.1±0.5	27.7±1.9	6.4±0.6	-0.1±0.5	166±23	4.39±5.07	5	
	0.864±0.031	3.0±0.0	3.8±0.0	21.6±2.0	3.6±0.0	4.1±0.0	64±28	3.27±0.44	5	
	0.887±0.012	2.9±0.0	8.2±0.0	29.8±1.0	3.5±0.0	8.5±0.0	78±30	3.98±0.26	5	
	0.805±0.014	1.5±0.0	7.6±0.0	
	0.082±0.023	1.9±0.3	-1.4±0.0	

Table B.1 (cont'd)

Source (name)	τ_0	Δv_0 (km s $^{-1}$)	v_0 (km s $^{-1}$)	$T_{B,n}$ (K)	$\Delta v_{0,n}$ (km s $^{-1}$)	$v_{0,n}$ (km s $^{-1}$)	T_s (K)	N(HI) (10 20 cm $^{-2}$)	\mathcal{O}	\mathcal{F}
	0.061±0.009	1.5±0.2	-4.2±0.1
3C138	0.033±0.000	2.4±0.0	-22.0±0.0	1.5±0.0	2.9±0.0	-20.8±0.0	47±2	0.07±0.00	7	1.0
	0.015±0.000	4.4±0.0	-6.6±0.0	1.9±0.0	4.3±0.0	-7.0±0.0	170±58	0.22±0.00	7	
	0.275±0.000	3.3±0.0	-0.5±0.0	12.8±0.0	3.3±0.0	-0.4±0.0	74±18	1.34±0.01	3	
	0.171±0.000	2.0±0.0	1.6±0.0	16.3±0.0	2.2±0.0	1.6±0.0	93±21	0.62±0.00	1	
	0.035±0.000	15.8±0.0	2.6±0.0	40.0±0.0	17.4±0.0	3.1±0.0	1126±46	12.34±0.01	2	
	1.026±0.000	2.3±0.0	6.3±0.0	14.6±0.0	2.8±0.0	5.2±0.0	30±9	1.47±0.02	7	
	0.417±0.000	3.0±0.0	9.0±0.0	19.1±0.0	3.5±0.0	8.8±0.0	60±14	1.51±0.01	7	
	0.000±0.000	8.7±0.0	10.1±0.0	
PKS0531	0.001±0.001	10.9±3.9	-25.0±2.0	4.3±2.4	13.2±3.9	-24.7±2.0	3101±2085	0.94±99.	9	1.0
	0.007±0.001	3.5±0.5	-22.9±0.2	0.5±0.1	3.8±0.5	-20.0±0.2	75±10	0.04±0.02	9	
	0.008±0.013	10.2±4.2	-5.7±5.4	19.1±29.2	12.3±4.2	-5.6±5.4	2305±4233	3.88±99.	9	
	0.018±0.005	2.7±0.4	-2.0±0.2	7.6±2.4	2.7±0.4	-2.0±0.2	483±125	0.46±3.19	9	
	0.405±0.006	2.2±0.0	1.7±0.0	3.7±0.3	2.2±0.0	1.4±0.0	47±24	0.82±0.02	9	
	0.208±0.019	7.6±0.4	5.2±0.2	74.0±7.3	8.7±0.4	5.5±0.2	391±40	12.34±64.87	9	
	0.157±0.011	2.1±0.1	9.5±0.0	9.6±1.0	2.1±0.1	9.5±0.0	98±20	0.64±0.18	9	
	0.033±0.012	4.0±0.9	11.2±0.6	16.2±6.1	4.4±0.9	11.4±0.6	521±199	1.36±20.09	9	
	0.005±0.000	4.9±0.6	20.7±0.3	4.0±0.4	5.9±0.6	20.5±0.3	789±85	0.39±2.61	9	
	0.000±0.029	11.3±99.	1.6±99.	
3C147	0.013±0.001	2.0±0.1	-19.4±0.1	0.6±0.0	2.0±0.1	-20.3±0.1	49±4	0.02±0.00	6	1.0
	0.057±0.009	1.9±0.1	-13.7±0.0	1.0±0.2	2.2±0.1	-15.3±0.0	17±5	0.04±0.00	6	
	0.112±0.021	6.5±0.3	-11.1±0.0	24.9±5.0	6.4±0.3	-11.0±0.0	243±44	3.51±20.04	2	
	0.173±0.020	2.3±0.1	-10.7±0.0	4.5±0.5	2.0±0.1	-9.9±0.0	25±6	0.20±0.03	6	
	0.727±0.006	1.7±0.0	-8.0±0.0	4.7±0.1	1.9±0.0	-6.8±0.0	12±2	0.30±0.00	6	
	0.281±0.004	5.9±0.1	0.5±0.0	27.1±0.5	7.1±0.1	0.5±0.0	122±9	4.01±0.89	6	
	0.205±0.004	2.0±0.0	-0.1±0.0	
	0.042±0.003	1.5±0.1	1.8±0.0	
	0.036±0.002	1.7±0.1	5.1±0.0	
3C154	0.070±0.003	3.5±0.2	-23.2±0.1	5.8±0.3	3.5±0.2	-24.8±0.1	85±3	0.42±0.11	6	1.0
	0.040±0.023	6.5±3.1	-6.7±2.7	22.6±12.6	6.4±3.1	-6.9±2.7	562±322	2.90±99.	6	
	1.076±0.033	4.7±0.1	-2.5±0.0	55.8±3.7	4.6±0.1	-2.5±0.0	111±15	11.14±3.55	6	
	1.280±0.020	2.3±0.0	2.0±0.0	31.9±1.7	2.7±0.0	1.6±0.0	89±22	5.25±0.35	6	
	0.832±0.006	4.3±0.1	4.7±0.0	42.1±0.6	5.2±0.1	4.7±0.0	95±20	6.80±0.43	6	
	0.472±0.021	2.5±0.0	10.8±0.0	12.2±1.0	3.0±0.0	10.9±0.0	46±16	1.07±0.11	6	
	0.006±0.004	34.1±11.4	-3.0±3.9	
	0.283±0.009	1.1±0.0	-2.2±0.0	
	0.093±0.021	1.1±0.2	11.1±0.0	
PKS0742	0.009±0.000	3.2±0.1	1.5±0.0	1.8±0.1	3.8±0.2	1.7±0.0	190±11	0.11±0.03	1	0.0
3C225A	0.043±0.002	1.8±0.1	-40.2±0.0	1.2±0.1	1.8±0.1	-39.6±0.0	28±1	0.04±0.00	5	1.0
	0.020±0.001	4.8±0.3	-37.4±0.2	1.0±0.0	5.8±0.3	-38.2±0.2	51±1	0.10±0.01	5	
	0.048±0.001	2.5±0.0	-27.2±0.0	0.7±0.0	3.1±0.0	-27.8±0.0	16±1	0.04±0.00	5	
	0.013±0.000	7.7±0.3	-5.2±0.1	7.1±0.2	8.5±0.3	-4.2±0.1	543±19	1.09±2.89	5	
	0.806±0.001	1.3±0.0	4.0±0.0	9.3±0.0	1.5±0.0	3.6±0.0	18±1	0.38±0.00	5	
3C225B	0.044±0.002	2.0±0.1	-40.3±0.0	0.9±0.0	1.9±0.1	-39.8±0.0	21±0	0.04±0.00	5	1.0
	0.023±0.001	4.0±0.3	-37.2±0.1	3.9±0.1	4.0±0.3	-37.4±0.1	168±5	0.31±0.14	5	
	0.053±0.001	2.4±0.1	-27.3±0.0	1.2±0.0	2.9±0.1	-27.8±0.0	24±1	0.06±0.00	5	
	0.013±0.001	8.3±0.4	-5.6±0.2	7.1±0.3	8.3±0.4	-4.5±0.2	530±20	1.18±3.91	5	
	0.774±0.001	1.3±0.0	4.0±0.0	9.0±0.0	1.5±0.0	3.5±0.0	17±0	0.34±0.00	5	
3C236	
3C237	0.006±0.000	15.0±0.6	-4.0±0.3	2.5±0.1	14.8±0.6	-4.0±0.3	443±20	0.74±4.37	2	1.0
	0.415±0.001	1.2±0.0	2.3±0.0	5.0±0.0	1.2±0.0	2.1±0.0	15±0	0.15±0.00	2	
	0.005±0.001	1.9±0.2	-2.4±0.1	

Table B.1 (cont'd)

Source (name)	τ_0	Δv_0 (km s $^{-1}$)	v_0 (km s $^{-1}$)	$T_{B,n}$ (K)	$\Delta v_{0,n}$ (km s $^{-1}$)	$v_{0,n}$ (km s $^{-1}$)	T_s (K)	N(HI) (10 20 cm $^{-2}$)	σ	\mathcal{F}
3C245A	0.006±0.001 0.010±0.001	1.6±0.2 5.3±0.2	-9.8±0.1 -9.1±0.1	0.4±0.1 3.2±0.2	1.5±0.3 5.3±0.2	-9.1±0.1 -8.9±0.1	66±14 324±29	0.01±0.00 0.35±0.81	2 2	1.0
3C245B
1055+018	0.006±0.000	7.0±0.2	-7.3±0.1	6.8±0.2	8.5±0.2	-7.3±0.1	1047±35	0.94±3.74	1	0.0
3C263.1	0.020±0.000 0.007±0.001 0.007±0.001	2.0±0.1 0.6±0.1 1.3±0.1	-52.8±0.0 15.6±0.0 -68.9±0.1	1.0
3C273	0.019±0.000 0.005±0.000 0.003±0.000 0.002±0.000	2.3±0.0 6.3±0.3 2.0±0.2 1.3±0.2	-6.3±0.0 -5.8±0.1 31.6±0.1 24.3±0.1	1.5±0.0 3.3±0.3 0.0±0.0 ...	2.8±0.0 7.7±0.3 2.0±0.2 ...	-5.7±0.0 -5.6±0.1 31.6±0.1 ...	82±2 623±63 3±0 ...	0.07±0.01 0.41±2.65 0.00±0.00 ...	3 3 3 ...	1.0
4C32.44	0.018±0.000 0.004±0.000	2.8±0.0 3.7±0.3	-16.2±0.0 -4.2±0.1	1.8±0.0 1.0±0.1	2.8±0.0 3.7±0.3	-16.2±0.0 -4.4±0.1	95±1 292±17	0.10±0.01 0.08±0.10	2 2	0.0
4C25.43	0.004±0.001	0.8±0.1	5.3±0.1	1.0
3C286	0.006±0.000 0.005±0.000 0.007±0.000	2.4±0.1 3.2±0.1 4.3±0.1	-28.5±0.0 -14.2±0.1 -7.3±0.0	0.7±0.0 0.4±0.0 0.9±0.0	2.9±0.1 3.9±0.1 5.2±0.1	-28.9±0.0 -13.7±0.1 -5.9±0.0	107±4 91±4 130±3	0.03±0.01 0.03±0.01 0.08±0.02	3 3 3	0.0
4C12.50	0.016±0.001 0.077±0.001	6.6±0.1 2.2±0.0	-2.5±0.1 -1.0±0.0	4.6±0.1 7.1±0.1	6.8±0.1 2.7±0.0	-2.5±0.1 -1.0±0.0	284±9 97±1	0.60±0.69 0.33±0.01	2 1	1.0
3C298	0.019±0.000	3.6±0.1	-1.1±0.0	2.4±0.2	4.3±0.4	-1.1±0.2	124±10	0.17±0.02	1	0.0
UGC09799	0.056±0.009 0.013±0.003	2.3±0.2 3.6±1.8	-4.4±0.1 -1.8±1.2	4.1±0.7 1.5±0.4	2.2±0.2 3.6±1.8	-4.5±0.1 -1.8±1.2	78±12 118±27	0.20±0.11 0.11±0.23	2 2	1.0
4C04.51	0.002±0.000 0.025±0.000 0.067±0.000	14.0±1.5 2.4±0.1 3.4±0.0	-9.1±0.8 -4.0±0.0 -6.0±0.0	1.2±0.1 5.1±0.1 13.2±0.1	17.0±1.5 2.4±0.1 4.2±0.0	-8.2±0.8 -4.0±0.0 -0.6±0.0	529±38 204±4 201±2	0.34±2.96 0.24±0.04 0.91±0.07	3 3 3	0.0
3C327.1A	0.126±0.001 0.434±0.006 0.397±0.003	3.3±0.1 1.9±0.0 2.1±0.0	-2.7±0.0 -0.0±0.0 2.0±0.0	19.0±0.2 0.0±0.1 23.7±0.3	4.0±0.1 1.9±0.0 2.4±0.0	-2.9±0.0 -0.0±0.0 2.0±0.0	163±3 13±13 75±3	1.35±0.14 0.23±0.00 1.27±0.04	2 3 1	1.0
3C327.1B	0.117±0.001 0.375±0.006 0.409±0.005 0.013±0.002 0.011±0.001	3.1±0.1 2.0±0.0 2.1±0.0 0.7±0.1 4.3±0.4	-2.8±0.0 0.0±0.0 1.9±0.0 46.1±0.1 20.1±0.2	18.2±0.2 17.4±0.3 23.5±0.3 ...	3.8±0.1 2.0±0.0 2.4±0.0 ...	-2.9±0.0 0.0±0.0 2.0±0.0 ...	166±3 62±6 73±4 ...	1.22±0.13 0.93±0.04 1.24±0.04 ...	3 3 3 ...	1.0
PKS1607	0.129±0.001 0.013±0.000 0.066±0.004 0.123±0.005	2.0±0.0 6.8±0.3 3.8±0.1 2.2±0.0	-10.3±0.0 -7.6±0.2 -1.2±0.1 -2.3±0.0	2.9±0.0 2.0±0.1 8.2±0.5	2.0±0.0 6.7±0.3 4.6±0.1	-11.0±0.0 -6.2±0.2 -1.1±0.1 ...	26±1 155±10 131±12	0.13±0.00 0.28±0.14 0.65±0.36	1 2 3	1.0
J1613	0.004±0.001	0.6±0.1	-44.2±0.1	1.0
3C345	0.009±0.001	0.6±0.1	-5.2±0.0	0.0
3C346	0.034±0.001 0.280±0.001 0.198±0.001	4.7±0.1 1.9±0.0 1.9±0.0	-6.3±0.0 -0.6±0.0 2.1±0.0	5.8±0.1 18.8±0.1 3.8±0.0	4.9±0.1 2.4±0.0 2.3±0.0	-6.3±0.0 -0.6±0.0 2.1±0.0	174±4 80±3 32±11	0.55±0.19 0.87±0.01 0.24±0.00	3 3 3	1.0
3C390	0.101±0.004 0.148±0.002 0.082±0.001 0.044±0.002 0.028±0.006 0.093±0.001 0.069±0.001 0.021±0.001 0.007±0.000 0.024±0.002	2.2±0.0 3.7±0.1 5.5±0.1 2.0±0.1 1.1±0.1 10.0±0.1 1.4±0.0 1.6±0.1 9.8±1.2 1.8±0.3	1.0±0.0 3.7±0.0 8.8±0.1 11.2±0.0 23.5±0.1 0.3±0.1 1.3±0.1 4.8±0.2 42.1±0.3 24.7±0.2	11.8±0.5 37.6±0.5 34.3±0.4 1.6±0.1 0.3±0.1 39.4±0.4 0.0±0.0 4.8±0.2 18.3±1.1 ...	2.4±0.0 4.1±0.1 6.2±0.1 2.5±0.1 1.3±0.1 12.1±0.1 1.7±0.0 2.0±0.1 11.9±1.2 ...	0.8±0.0 3.2±0.0 8.8±0.1 12.3±0.0 22.4±0.1 28.0±0.0 31.3±0.0 37.3±0.0 41.9±0.3 ...	127±8 277±4 437±5 42±6 17±4 436±6 10±4 235±13 2722±202 ...	0.56±0.13 3.03±0.72 3.94±2.08 0.08±0.01 0.01±0.00 8.02±6.69 0.02±0.00 0.16±0.06 3.58±99. ...	9 9 9 9 9 2 1 9 9 ...	1.0

Table B.1 (cont'd)

Source (name)	τ_0	Δv_0 (km s $^{-1}$)	v_0 (km s $^{-1}$)	$T_{B,n}$ (K)	$\Delta v_{0,n}$ (km s $^{-1}$)	$v_{0,n}$ (km s $^{-1}$)	T_s (K)	N(HI) (10 20 cm $^{-2}$)	\mathcal{O}	\mathcal{F}
4C33.48	0.172±0.002	2.2±0.0	3.5±0.0	3.7±0.1	2.2±0.0	3.5±0.0	32±9	0.24±0.00	3	1.0
	0.139±0.003	6.8±0.1	25.5±0.1	25.6±0.7	7.4±0.1	23.4±0.1	212±14	3.95±2.53	3	
	0.127±0.003	1.7±0.1	28.8±0.0	1.2±0.0	2.0±0.1	29.5±0.0	12±6	0.05±0.00	2	
	0.278±0.004	2.3±0.0	23.8±0.0	
3C409A	0.431±0.004	3.3±0.0	4.2±0.0	19.7±0.2	4.0±0.0	3.4±0.0	59±8	1.66±0.05	5	1.0
	0.269±0.007	2.5±0.0	7.9±0.0	18.6±0.5	3.0±0.0	7.8±0.0	70±23	0.93±0.09	2	
	0.242±0.010	9.0±0.3	13.2±0.2	51.5±3.0	10.8±0.3	12.5±0.2	306±58	13.19±28.60	3	
	0.892±0.006	1.8±0.0	15.9±0.0	3.2±0.2	2.0±0.0	15.8±0.0	30±23	0.99±0.03	5	
	0.017±0.001	4.2±0.3	24.6±0.1	7.2±0.3	4.2±0.3	23.0±0.1	382±73	0.56±0.82	1	
	0.902±0.012	2.3±0.0	13.9±0.0	
	0.080±0.006	2.1±0.1	18.2±0.1	
3C409B	0.444±0.001	2.9±0.0	4.0±0.0	18.3±0.0	3.6±0.0	4.0±0.0	64±11	1.67±0.01	2	1.0
	0.289±0.003	3.0±0.0	7.9±0.0	22.3±0.3	2.9±0.0	7.6±0.0	116±18	1.97±0.13	5	
	0.367±0.008	7.1±0.1	14.1±0.0	53.6±1.9	7.8±0.1	14.2±0.0	224±43	11.64±8.19	5	
	0.593±0.007	2.2±0.0	15.7±0.0	0.6±0.1	2.0±0.0	15.7±0.0	12±36	0.33±0.02	5	
	0.032±0.001	4.6±0.2	23.3±0.1	5.8±0.2	5.5±0.2	24.8±0.1	237±69	0.69±0.30	1	
	0.844±0.008	2.1±0.0	13.8±0.0	
	0.444±0.001	2.9±0.0	4.0±0.0	18.3±0.0	3.6±0.0	4.0±0.0	64±11	1.67±0.01	2	1.0
3C410A	0.014±0.003	1.5±0.3	-30.2±0.1	0.0±0.0	1.8±0.3	-29.9±0.1	7±18	0.00±0.00	10	1.0
	0.020±0.002	3.9±0.4	-22.7±0.2	14.3±1.2	3.9±0.4	-23.4±0.2	725±65	1.14±5.26	10	
	0.618±0.016	3.6±0.1	-0.2±0.1	15.1±0.7	4.0±0.1	-0.2±0.1	45±32	2.04±0.19	10	
	0.635±0.022	3.2±0.1	2.7±0.1	14.3±1.2	3.9±0.1	2.7±0.1	53±41	2.15±0.26	10	
	1.805±0.099	3.2±0.1	7.5±0.1	51.6±9.8	3.9±0.1	7.5±0.1	103±53	12.01±4.59	10	
	0.214±0.112	1.8±0.3	11.0±0.1	13.4±14.1	1.9±0.3	9.7±0.1	132±239	1.01±2.46	10	
	0.509±0.096	4.0±0.4	11.6±0.5	55.5±14.2	4.0±0.4	11.6±0.5	156±50	6.35±14.69	10	
	0.185±0.002	3.4±0.1	17.9±0.0	24.2±0.3	3.9±0.1	16.8±0.0	144±31	1.81±0.17	10	
	0.075±0.005	1.8±0.1	24.6±0.0	4.5±0.3	2.0±0.1	20.9±0.0	62±15	0.17±0.02	10	
	0.061±0.005	5.1±0.2	25.3±0.1	11.0±0.9	5.1±0.2	22.9±0.1	198±41	1.23±1.88	10	
	1.772±0.081	1.5±0.0	8.1±0.0	
	0.056±0.002	1.9±0.1	-46.5±0.0	
	0.048±0.002	2.8±0.2	-4.8±0.1	
	0.476±0.005	2.7±0.0	-0.2±0.0	29.1±0.4	2.6±0.0	-0.2±0.0	76±0	1.92±0.08	7	1.0
	0.426±0.010	2.3±0.1	2.5±0.0	27.5±0.8	2.3±0.1	2.6±0.0	79±1	1.53±0.13	7	
	0.344±0.023	11.4±0.5	5.1±0.3	54.4±4.2	13.8±0.5	4.3±0.3	187±15	14.55±40.28	7	
	0.292±0.011	1.5±0.0	5.9±0.0	17.1±0.7	1.4±0.0	5.6±0.0	67±2	0.57±0.04	7	
	2.798±0.014	2.2±0.0	8.0±0.0	108.0±1.5	2.7±0.0	7.6±0.0	114±0	13.97±0.45	7	
	0.431±0.137	1.9±0.1	10.9±0.1	72.3±26.4	2.1±0.1	9.9±0.1	206±72	3.39±8.33	7	
	0.041±0.017	19.2±2.0	12.7±3.1	47.8±19.5	23.2±2.0	10.3±3.1	1175±574	18.51±99.	7	
	0.021±0.001	5.3±0.4	-47.6±0.2	
	0.019±0.002	1.8±0.3	-30.2±0.1	
	0.119±0.003	2.5±0.1	18.4±0.0	
	0.347±0.087	2.6±0.3	12.2±0.4	
	0.112±0.002	2.5±0.1	25.2±0.0	
B2050	0.007±0.000	3.4±0.2	-21.3±0.1	0.8±0.1	3.7±0.2	-21.1±0.1	128±10	0.06±0.03	1	1.0
	0.011±0.001	2.6±0.2	-11.5±0.1	0.6±0.0	3.1±0.2	-12.0±0.1	66±11	0.04±0.01	2	
	0.014±0.001	1.2±0.1	-6.2±0.0	0.1±0.0	1.5±0.1	-6.9±0.0	16±5	0.01±0.00	6	
	0.045±0.002	9.5±0.3	-0.2±0.1	13.4±0.6	9.4±0.3	-0.2±0.1	339±35	2.90±7.13	4	
	0.191±0.001	6.1±0.1	10.2±0.0	23.5±0.2	7.3±0.1	10.3±0.0	150±14	3.44±0.34	6	
	0.074±0.001	3.7±0.0	16.0±0.0	2.1±0.0	3.7±0.0	17.4±0.0	34±6	0.19±0.01	3	
	0.150±0.001	1.7±0.0	9.0±0.0	
	0.084±0.002	2.7±0.0	-0.8±0.0	
	0.039±0.001	2.0±0.1	2.5±0.0	
3C433	0.302±0.005	3.9±0.0	3.5±0.0	23.6±0.5	3.9±0.0	2.9±0.0	104±12	2.48±0.34	2	1.0

Table B.1 (cont'd)

Source (name)	τ_0	Δv_0 (km s $^{-1}$)	v_0 (km s $^{-1}$)	$T_{B,n}$ (K)	$\Delta v_{0,n}$ (km s $^{-1}$)	$v_{0,n}$ (km s $^{-1}$)	T_s (K)	N(HI) (10 20 cm $^{-2}$)	\mathcal{O}	\mathcal{F}
	0.080±0.002	2.0±0.1	7.3±0.0	3.4±0.1	2.0±0.1	7.3±0.0	51±7	0.17±0.01	1	
	0.059±0.001	2.5±0.1	16.5±0.0	2.1±0.0	3.0±0.1	15.2±0.0	35±1	0.10±0.00	3	
	0.183±0.006	1.5±0.0	2.6±0.0	
PKS2127	0.071±0.001	2.7±0.0	-0.9±0.0	6.0±0.1	2.7±0.0	-1.0±0.0	82±3	0.32±0.02	2	0.0
	0.015±0.001	6.8±0.4	1.0±0.1	10.9±0.9	6.0±0.4	1.0±0.1	735±56	1.46±12.15	3	
	0.103±0.001	2.1±0.0	1.5±0.0	4.6±0.1	1.9±0.0	1.5±0.0	39±4	0.17±0.00	1	
	0.002±0.001	17.2±2.7	2.6±0.9	7.9±3.0	20.8±2.7	3.1±0.9	5298±2406	2.72±99.	4	
J2136	0.096±0.001	2.2±0.0	1.1±0.0	2.3±0.0	2.2±0.0	1.0±0.0	20±6	0.08±0.00	2	0.0
	0.062±0.001	7.0±0.0	3.5±0.0	4.7±0.1	7.7±0.0	3.4±0.0	72±7	0.62±0.12	1	
	0.079±0.001	2.5±0.0	5.8±0.0	5.5±0.1	3.0±0.0	6.0±0.0	67±8	0.26±0.01	3	
J2232	0.057±0.002	2.0±0.1	-14.4±0.0	2.8±0.1	2.0±0.1	-15.0±0.0	50±2	0.11±0.01	1	1.0
	0.011±0.002	4.3±0.4	-13.7±0.2	3.7±0.8	4.3±0.4	-13.6±0.2	328±71	0.32±1.89	5	
	0.085±0.002	2.5±0.0	-7.5±0.0	9.7±0.2	3.0±0.0	-7.4±0.0	120±3	0.50±0.07	5	
	0.094±0.008	3.6±0.2	-3.9±0.1	9.1±0.8	4.0±0.2	-3.9±0.1	106±10	0.73±0.45	5	
	0.047±0.000	3.4±0.1	2.0±0.0	3.4±0.0	4.1±0.1	2.4±0.0	78±4	0.25±0.01	5	
	0.063±0.008	1.7±0.1	-3.6±0.0	
3C454.3	0.010±0.000	3.5±0.2	-35.1±0.1	1.2±0.1	4.3±0.2	-35.1±0.1	126±6	0.09±0.03	8	1.0
	0.047±0.001	2.1±0.0	-30.5±0.0	0.4±0.0	2.1±0.0	-30.5±0.0	9±1	0.02±0.00	8	
	0.016±0.000	5.7±0.2	-16.5±0.1	2.1±0.0	5.7±0.2	-16.5±0.1	137±3	0.25±0.08	8	
	0.315±0.017	2.6±0.0	-10.3±0.0	12.8±0.9	2.5±0.0	-9.7±0.0	50±3	0.81±0.11	1	
	0.021±0.008	3.8±1.2	-8.4±1.1	0.7±0.3	3.8±1.2	-8.4±1.1	30±14	0.05±0.04	8	
	0.097±0.000	3.5±0.1	-2.1±0.0	12.5±0.1	3.5±0.1	-2.1±0.0	139±3	0.94±0.05	8	
	0.085±0.002	1.8±0.0	0.7±0.0	5.2±0.1	2.0±0.0	0.8±0.0	67±4	0.21±0.01	8	
	0.025±0.000	4.2±0.2	3.2±0.1	3.6±0.1	5.1±0.2	3.8±0.1	147±4	0.31±0.07	8	
3C459	0.040±0.001	5.2±0.2	-6.2±0.2	15.8±0.6	5.2±0.2	-6.6±0.2	407±16	1.70±2.52	2	1.0
	0.055±0.003	1.9±0.0	0.4±0.0	2.3±0.1	1.9±0.0	0.4±0.0	28±9	0.06±0.00	1	
	0.042±0.004	7.3±0.6	1.0±0.3	20.8±1.8	7.4±0.6	0.8±0.3	507±45	3.07±19.18	5	
	0.086±0.003	2.2±0.0	2.9±0.0	4.4±0.1	2.2±0.0	2.9±0.0	42±7	0.16±0.01	5	
	0.017±0.001	1.4±0.1	7.7±0.0	1.1±0.0	1.7±0.1	9.3±0.0	67±3	0.03±0.00	3	
	0.102±0.002	2.2±0.0	-7.3±0.0	
	0.010±0.000	2.8±0.1	-13.3±0.0	

Note. — From the analysis presented in Chapter 5. Cols. (2-4): Gaussian parameters fit to H I absorption (Equation 5.2). Cols. (5-7): Gaussian parameters fit to H I emission (Equations 5.3). Col. (8): Average spin temperature from all permutations of components with overlap along the line of sight (Equations 5.4, 5.5). Col. (9): Column density computed from fitted parameters (Equation 5.8). Col. (10): Order of components along the line of sight corresponding to the smallest model residuals. Components whose position along the line of sight is extremely uncertain or unaffected by order permutations are assumed to lie behind all others (i.e., $\mathcal{O} = N$, for N total components). Col (11): Fraction of WNM (emission-detected only) components which lie in front of all absorption-detected components, allowed to be 1.0 or 0.0 for all emission-detected components. Fit parameters for components with $T_s \leq 3$ K are omitted, as these are either spurious AGD detections or were not recovered in the fit to $T_B(v)$ due to strong line blending.