

COLD AND WARM ATOMIC GAS AROUND THE PERSEUS MOLECULAR CLOUD. II. THE IMPACT OF HIGH OPTICAL DEPTH ON THE HI COLUMN DENSITY DISTRIBUTION AND ITS IMPLICATION FOR THE HI-TO-H₂ TRANSITION

MIN-YOUNG LEE¹, SNEŽANA STANIMIROVIĆ², CLAIRE E. MURRAY², CARL HEILES³, AND JESSE MILLER⁴

¹ Laboratoire AIM, CEA/IRFU/Service d'Astrophysique, Bat 709, F-91191 Gif-sur-Yvette, France; min-young.lee@cea.fr

² Department of Astronomy, University of Wisconsin, Madison, WI 53706, USA

³ Department of Astronomy, University of California, Berkeley, CA 94720, USA

⁴ Department of Physics and Astronomy, Washington State University, P.O. Box 642814, Pullman, WA 99164-2814, USA

Received 2014 October 18; accepted 2015 May 10; published 2015 August 10

ABSTRACT

We investigate the impact of high optical depth on the HI saturation observed in the Perseus molecular cloud by using Arecibo HI emission and absorption measurements toward 26 radio continuum sources. The spin temperature and optical depth of individual HI components are derived along each line of sight, enabling us to estimate the correction for high optical depth. We examine two different methods for the correction, Gaussian decomposition and isothermal methods, and find that they are consistent (maximum correction factor ~ 1.2), likely due to the relatively low optical depth and insignificant contribution from the diffuse radio continuum emission for Perseus. We apply the correction to the optically thin HI column density on a pixel-by-pixel basis and find that the total HI mass increases by $\sim 10\%$. Using the corrected HI column density image and far-infrared data from the IRIS Survey, we then derive the H₂ column density on ~ 0.4 pc scales. For five dark and star-forming sub-regions, the HI surface density is uniform with $\Sigma_{\text{HI}} \sim 7\text{--}9 M_{\odot} \text{ pc}^{-2}$, in agreement with the minimum HI surface density required for shielding H₂ against photodissociation. As a result, $\Sigma_{\text{H}_2}/\Sigma_{\text{HI}}$ and $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$ show a tight relation. Our results are consistent with predictions for H₂ formation in steady state and chemical equilibrium and suggest that H₂ formation is mainly responsible for the Σ_{HI} saturation in Perseus. We also compare the optically thick HI with the observed “CO-dark” gas and find that the optically thick HI only accounts for $\sim 20\%$ of the “CO-dark” gas in Perseus.

Key words: ISM: individual objects (Perseus) – ISM: molecules – radio lines: ISM

1. INTRODUCTION

As the most abundant molecular species in the universe, molecular hydrogen (H₂) is the main constituent of giant molecular clouds, the exclusive birthplaces of stars (e.g., Kennicutt & Evans 2012). Recent observations of galaxies at both low and high redshifts have shown that star formation rates are strongly correlated with H₂ surface densities (Σ_{H_2}), the relation generally known as the “Kennicutt–Schmidt law” (e.g., Schmidt 1959; Kennicutt 1989; Bigiel et al. 2008; Wilson et al. 2009; Tacconi et al. 2010; Schruba et al. 2011; Genzel et al. 2013). This suggests that physical processes responsible for the atomic-to-molecular hydrogen (HI-to-H₂) transition play a key role in the evolution of galaxies.

Observationally, the HI-to-H₂ transition has been studied via ultraviolet (UV) absorption measurements along many random lines of sight through the Galaxy (e.g., Savage et al. 1977; Rachford et al. 2002; Gillmon et al. 2006). For these measurements, either early-type stars or active galactic nuclei were used as background sources, and HI and H₂ column densities, $N(\text{HI})$ and $N(\text{H}_2)$, were estimated from Ly α and Lyman–Werner (LW) band absorption. The UV studies probed the H₂ mass fraction $f_{\text{H}_2} = 2N(\text{H}_2)/[N(\text{HI}) + 2N(\text{H}_2)]$ ranging from $\sim 10^{-6}$ to $\sim 10^{-1}$ and found that f_{H_2} sharply increases at the total gas column density $N(\text{H}) = N(\text{HI}) + 2N(\text{H}_2)$ of $\sim (3\text{--}5) \times 10^{20} \text{ cm}^{-2}$. Additionally, the HI-to-H₂ transition has been indirectly inferred from the flattening of the relation between the HI column density and a tracer of total gas column density (e.g., far-infrared [FIR] or hydroxide [OH] emission; Reach et al. 1994; Meyerdericks & Heithausen 1996; Douglas &

Taylor 2007; Barriault et al. 2010; Liszt 2014). These studies found that the HI column density saturates to $\sim (5\text{--}10) \times 10^{20} \text{ cm}^{-2}$, suggesting the presence of H₂. The HI saturation has also been found in extragalactic observations on $\sim \text{kpc}$ scales (e.g., Wong & Blitz 2002; Blitz & Rosolowsky 2006; Leroy et al. 2008; Wong et al. 2009).

Theoretically, the HI-to-H₂ transition has been investigated as a central process in photodissociation regions (PDRs). In PDRs, the interstellar medium (ISM) is predominately atomic, and the molecular gas is only found in well-shielded regions where dissociating UV photons are sufficiently attenuated. Many studies have been presented with different treatments of chemistry, geometry, and radiative transfer (e.g., Spitzer 1948; Gould & Salpeter 1963; Glassgold & Langer 1974; van Dishoeck & Black 1986; Sternberg 1988; Elmegreen 1993; Draine & Bertoldi 1996; Spaans & Neufeld 1997; Browning et al. 2003; Goldsmith et al. 2007; Liszt 2007; Krumholz et al. 2009; Glover et al. 2010; Offner et al. 2013; Sternberg et al. 2014), and an excellent summary of these studies was recently provided by Sternberg et al. (2014).

Among the many studies, the Krumholz et al. (2009, hereafter KMT09) model has recently been tested with a variety of Galactic and extragalactic observations (e.g., Bolatto et al. 2011; Lee et al. 2012; Welty et al. 2012; Wong et al. 2013; Motte et al. 2014), thanks to its simple analytic predictions that allow a comparison with direct observables and an extrapolation of the model over a wide range of ISM environments. In the KMT09 model, a spherical cloud is illuminated by a uniform and isotropic radiation field, and the H₂ abundance is computed based on the balance between the

rate of formation on dust grains and the rate of dissociation by UV photons (chemical equilibrium). The authors derived two dimensionless parameters that determine the location of the HI-to-H₂ transition in the cloud, resulting in the following important predictions. First, they found that H₂ formation requires a certain amount of HI surface density (Σ_{HI}) for shielding against the dissociating radiation field. Interestingly, this shielding surface density primarily depends on metallicity and is expected to be $\sim 10 M_{\odot} \text{pc}^{-2}$ (corresponding to $N(\text{HI}) \sim 1.3 \times 10^{21} \text{cm}^{-2}$) for solar metallicity. Second, the H₂-to-HI ratio, $R_{\text{H}_2} = \Sigma_{\text{H}_2} / \Sigma_{\text{HI}}$, was predicted to linearly increase with the total gas surface density. This is because once the minimum HI surface density is obtained for shielding H₂ against photodissociation, all additional hydrogen is fully converted into H₂ and the HI surface density remains constant. As a result, R_{H_2} is simply a function of metallicity and total gas surface density. Another interesting feature of the KMT09 model is that the ISM is “self-regulated” in that pressure balance between the cold neutral medium (CNM) and the warm neutral medium (WNM) determines the ratio of the UV intensity to the HI density.

Aiming at testing the KMT09 model on sub-pc scales, we have recently focused on the Perseus molecular cloud (Lee et al. 2012). Perseus is one of the nearby molecular clouds in the Gould Belt and is located at a distance of $\sim 300 \text{ pc}$ (Herbig & Jones 1983; Černis 1990). It has a projected angular size of $\sim 6^\circ \times 3^\circ$ on the sky (based on the CO emission)⁵ and lies at high Galactic latitude $b \sim -20^\circ$, resulting in relatively simple HI spectra compared to other molecular clouds in the Galactic plane. With a total mass of $\sim 2 \times 10^4 M_{\odot}$ (Sancisi et al. 1974; Lada et al. 2010), Perseus is considered as a low-mass molecular cloud with an intermediate level of star formation (Bally et al. 2008). To test the KMT09 model, we derived Σ_{HI} and Σ_{H_2} images using HI data from the Galactic Arecibo L-band Feed Array HI Survey (GALFA-HI; Stanimirović et al. 2006; Peek et al. 2011) and FIR data from the Improved Reprocessing of the IRAS Survey (IRIS; Miville-Deschênes & Lagache 2005). The final images were at $\sim 0.4 \text{ pc}$ resolution and covered the far outskirts of the cloud, as well as the main body. We found that the HI surface density is relatively uniform with $\Sigma_{\text{HI}} \sim 6\text{--}8 M_{\odot} \text{pc}^{-2}$ for five dark and star-forming regions in Perseus (B5, B1E, B1, IC 348, and NGC 1333). In addition, the relation between R_{H_2} and $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$ on a log-linear scale was remarkably consistent for all individual regions, having a steep rise of R_{H_2} at small $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$, a turnover at $R_{\text{H}_2} \sim 1$, and a slow increase toward larger R_{H_2} . All these results were in excellent agreement with the KMT09 predictions for solar metallicity,⁶ suggesting that the KMT09 model captures well the fundamental physics of H₂ formation on sub-pc scales.

The observed HI saturation in Perseus, however, could alternatively result from the high optical depth HI. When the HI emission is optically thick, the brightness temperature (T_{B}) becomes comparable to the kinetic temperature (T_{k}). As a result, the HI surface density saturates in the optically thin approximation (which we used in Lee et al. 2012) since $\Sigma_{\text{HI}} \propto T_{\text{B}} \sim T_{\text{k}}$ and is underestimated. As the constant HI surface density is the key prediction from KMT09, it is critical to

evaluate how much of the HI column density distribution is affected by the optically thick HI.

In this paper, we assess the impact of high optical depth on the observed HI saturation in Perseus by using HI emission and absorption measurements obtained toward 26 background radio continuum sources (Stanimirović et al. 2014, hereafter Paper I). These observations provide the most direct way to measure the high optical depth HI, allowing us to derive the “true” total HI column density distribution. Specifically, we use the HI emission and absorption spectra to estimate the correction factor for high optical depth and apply the correction to the HI column density image computed by Lee et al. (2012) in the optically thin approximation. We take this localized approach rather than using data from existing all-sky surveys (e.g., Heiles & Troland 2003a, hereafter HT03a) in order to treat all spectra uniformly for the velocity range of Perseus and consider the possibility that CNM/WNM properties may vary with ISM environments (e.g., metallicity, star formation rate, etc.), as expected from theoretical models (e.g., McKee & Ostriker 1977; Koyama & Inutsuka 2002; Wolfire et al. 2003; Audit & Hennebelle 2005; Mac Low et al. 2005; Kim et al. 2013).

This paper is organized in the following way. We start with a summary of previous studies where various methods have been employed to derive the correction for high optical depth (Section 2). We then provide a description of the data used in this study (Section 3). In Section 4, we estimate the correction factor for high optical depth using two different methods and compare our results with previous studies. In Sections 5 and 6, we apply the correction to the HI column density image from Lee et al. (2012) on a pixel-by-pixel basis and revisit the HI saturation issue by rederiving the H₂ column density image and comparing our results with the KMT09 predictions. We then investigate whether or not the optically thick HI can explain the observed “CO-dark” gas in Perseus (Section 7), and finally we summarize our conclusions (Section 8).

2. BACKGROUND: METHODS TO ESTIMATE THE CORRECTION FOR HIGH OPTICAL DEPTH

In most radio observations, HI is detected in emission, and the intensity of radiation is measured as the brightness temperature as a function of radial velocity, i.e., $T_{\text{B}}(v)$. Since the HI optical depth (τ) can be measured only via absorption-line measurements in the direction of background radio continuum sources, the optically thin approximation of $\tau \ll 1$ is frequently employed to estimate the HI column density:

$$N(\text{HI}) \text{ (cm}^{-2}\text{)} = 1.823 \times 10^{18} \int T_{\text{B}}(v) dv \text{ (K km s}^{-1}\text{)}. \quad (1)$$

Over the past three decades, several approaches have been employed to estimate how much of the true total HI column density is underestimated in the optically thin approximation. Most of these approaches can be classified as “isothermal,” and the only multiphase approaches are by Dickey et al. (2000) and HT03a. Here we summarize the main results from some of the most important studies.

Dickey & Benson (1982) used 47 emission/absorption spectral line pairs in the direction of background sources and estimated the ratio of the HI column density from the absorption spectra to the HI column density in the optically thin approximation. Although ~ 1 at high and intermediate Galactic latitudes, the ratio reached ~ 1.8 at low latitudes. There

⁵ In this paper, $^{12}\text{CO}(J=1 \rightarrow 0)$ is quoted as CO.

⁶ Perseus has solar metallicity (González Hernández et al. 2009). See Section 7.2.1 of Lee et al. (2012) for a detailed discussion.

was considerable scatter in the ratio, however: several lines of sight at low latitudes showed small ratios, suggesting that the low-latitude directions with large ratios likely intersect dense molecular clouds. In order to compute the ratios, HI in each velocity channel was assumed to have a single temperature (“isothermal” approximation).

In the Dickey et al. (2000) study of the Small Magellanic Cloud (SMC), HI absorption observations were obtained in the direction of 13 background radio continuum sources. The corresponding emission spectra were derived by averaging HI profiles from Stanimirović et al. (1999) over a 3×3 pixel region (pixel size = $30''$) centered on the position of each source. The correction factor for high optical depth was calculated for each velocity channel in the isothermal approximation, and the line-of-sight integrated value f was expressed as a function of the uncorrected $N(\text{HI})$: $f = 1 + 0.667(\log_{10} N(\text{HI}) - 21.4)$ for $N(\text{HI}) > 10^{21.4} \text{ cm}^{-2}$. This relation was then applied to the $N(\text{HI})$ image of the SMC on a pixel-by-pixel basis, resulting in a $\sim 10\%$ increase of the total HI mass from $\sim 3.8 \times 10^8 M_{\odot}$ to $\sim 4.2 \times 10^8 M_{\odot}$. Although negligible at $N(\text{HI}) < 3 \times 10^{21} \text{ cm}^{-2}$, the correction factor increased with the uncorrected $N(\text{HI})$ up to ~ 1.4 at $N(\text{HI}) \sim 10^{22} \text{ cm}^{-2}$. In some cases, the correction factors for individual channels were larger than the integrated value, reaching up to ~ 2 . However, such values covered only a narrow range of channels, and their effect on $N(\text{HI})$ was relatively small. Finally, the authors rederived the correction factor in the two-phase approximation and found that the difference between the one- and two-phase cases depends on the relative location of cold and warm HI components along a line of sight (Section 4.3 for details).

In the Millennium Arecibo 21 cm Absorption Line Survey, HT03a obtained HI emission and absorption spectra toward 79 randomly positioned radio continuum sources and performed Gaussian decomposition to estimate the physical properties of individual CNM and WNM components (column density, optical depth, spin temperature T_s , etc.). These multiphase analyses showed that two or more components with very different spin temperatures can contribute to a single velocity channel, implying that the isothermal treatment may not be satisfactory. Heiles & Troland (2003b, hereafter HT03b) then calculated the correction factor using the Gaussian decomposition results, which they called $R_{\text{raw}} = 1/f$. There were interesting variations in R_{raw} , ranging from ~ 0.3 to ~ 1.0 ($f = \sim 1.0$ – 3.0 ; Appendix B for details). Specifically, $f \sim 1.3$ was found for the Taurus/Perseus region.

A very different approach was adopted in Braun et al. (2009) to calculate the correction for high optical depth in M31. They only used high-resolution HI emission observations for their modeling and assumed that a single cold component determines the brightness temperature along a line of sight. While previous similar studies have applied a single temperature to the images of entire galaxies (e.g., Henderson et al. 1982; Braun & Walterbos 1992), Braun et al. (2009) estimated the spin temperature and non-thermal velocity dispersion for each pixel. After excluding HI spectra that likely suffer from a high blending of different components along a line of sight, they noticed that the opaque HI is organized into filamentary complexes and isolated clouds down to their resolution limit of ~ 100 pc. The spin temperature was found to increase from ~ 20 to ~ 60 K with radius to 12 kpc, and then to decline smoothly down to ~ 20 K beyond 25 kpc. The estimated correction

resulted in a $\sim 30\%$ increase of the global HI mass of M31. Using the same methodology, Braun (2012) found that the correction for high optical depth increases the HI masses of the Large Magellanic Cloud (LMC) and M33 by the same amount ($\sim 30\%$). While the main advantage of the Braun et al. (2009) approach is clearly that galactic-scale images of the opaque component can be produced solely from HI emission observations, the method has several weaknesses. For example, it does not consider multiple components along a line of sight and how they self-absorb each other. In addition, it does not take account of the possibility that some of the brightness temperature could come from unabsorbing warm HI components.

Chengalur et al. (2013) tested the optically thin and isothermal approximations with Monte Carlo simulations of the multiphase ISM. They varied the fraction of gas in three phases (CNM, WNM, and the thermally unstable neutral medium) and the location of each phase along a line of sight. A wide range of values were assumed for the HI column density (10^{20} – 10^{24} cm^{-2}) and the spin temperature (20–5000 K). They found that the optically thin approximation underestimates the true HI column density by a factor of ~ 1.6 when $\int \tau dv \sim 1 \text{ km s}^{-1}$, while the underestimate can be as high as a factor of ~ 20 when $\int \tau dv \sim 10 \text{ km s}^{-1}$. On the other hand, the simulations showed that the isothermal estimate tracks the true HI column density to better than 10% even when $\int \tau dv \sim 5 \text{ km s}^{-1}$. Their conclusion that the isothermal estimate provides a good measure of the true HI column density of up to $\sim 5 \times 10^{23} \text{ cm}^{-2}$ was insensitive to the assumed gas temperature distribution and the positions of the different phases along a line of sight. We note that Equation (1) of Chengalur et al. (2013) does not include the contribution from the cosmic microwave background (CMB) and the Galactic synchrotron emission, which can be significant in certain cases (e.g., low Galactic latitudes). In addition, the authors did not consider self-absorption of the WNM by the foreground CNM.

Liszt (2014) compared $N(\text{HI})$ from Galactic HI surveys with $E(B - V)$ derived by Schlegel et al. (1998) and found a strong linear relation between $N(\text{HI})$ and $E(B - V) \sim 0.02$ – 0.08 mag and a flattening of the relation at $E(B - V) \gtrsim 0.08 \text{ mag}$. While this flattening, likely due to H_2 formation, was essentially the same effect as what Lee et al. (2012) found for individual regions in Perseus, the relation derived by Liszt (2014) covers a large spatial area with randomly selected lines of sight predominantly at $|b| > 20^\circ$. By using HI absorption data compiled by Liszt et al. (2010), the author then derived the correction for high optical depth and applied it to the $N(\text{HI})$ data. The flattening at $E(B - V) \gtrsim 0.08 \text{ mag}$ persisted after the correction, confirming its origin in the onset of H_2 formation. The derived correction factor increased from ~ 1.0 at $E(B - V) \sim 0.01 \text{ mag}$ to ~ 1.4 at $E(B - V) \sim 1 \text{ mag}$, and it was $\lesssim 1.2$ at $E(B - V) \lesssim 0.5 \text{ mag}$.

Recently, Fukui et al. (2015) suggested a new approach to estimate the correction for high optical depth by using *Planck* dust continuum data. They noticed that the dust optical depth at 353 GHz (τ_{353}) correlates with the HI column density, and the dispersion in this relation becomes much smaller when the data points are segregated based on the dust temperature (T_{dust}). The highest dust temperature was assumed to be associated with the optically thin HI, and the saturation seen in the τ_{353} – $N(\text{HI})$ relation was then solely attributed to the optically thick HI. By coupling the τ_{353} – $N(\text{HI})$ relation with radiative transfer

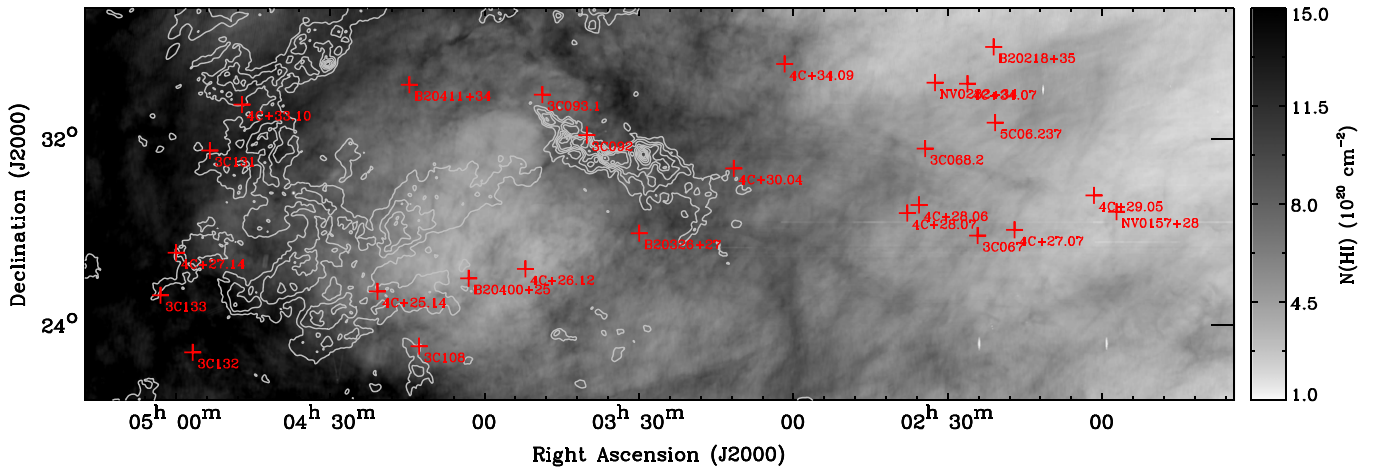


Figure 1. Twenty-six radio continuum sources overlaid on the HI column density image at 4' resolution (4C +32.14 excluded; Section 3.1 for details). The HI column density image is produced by integrating the GALFA-HI cube from $v_{\text{LSR}} = -5$ to $+15$ km s $^{-1}$, and the gray contours are from the CfA CO integrated intensity image at 8.4' resolution. The contour levels range from 10% to 90% of the peak value (69 K km s $^{-1}$) with 10% steps. In this figure, both the Perseus and Taurus molecular clouds are seen.

equations, Fukui et al. (2015) calculated T_s and τ for the Galactic sky at $|b| > 15^\circ$ on a pixel-by-pixel basis. They found that more than 70% of the data points have $T_s < 40$ K and $\tau > 0.5$, and similar results were obtained for the high-latitude molecular clouds MBM 53, MBM 54, MBM 55, and HLCG 92-35 (Fukui et al. 2014). The correction for high optical depth resulted in a factor of ~ 2 increase in the total HI mass in the solar neighborhood, implying that the optically thick HI may explain the “CO-dark” gas in the Galaxy.

3. DATA

3.1. HI Emission and Absorption Observations

We use the HI emission and absorption observations from Paper I. The observations were performed with the Arecibo telescope⁷ using the *L*-band wide receiver and were made toward 27 radio continuum sources located behind Perseus. The target sources were selected from the NRAO VLA Sky Survey (Condon et al. 1998) based on flux densities at 1.4 GHz greater than ~ 0.8 Jy, and they are distributed over a large area of ~ 500 deg 2 centered on the cloud (Figure 1).⁸ The angular resolution of the Arecibo telescope at 1.4 GHz is 3'.5. For the observations, a special procedure was adopted to make a “17-point pattern,” which includes 1 on-source measurement and 16 off-source measurements (HT03a; Stanimirović & Heiles 2005). This procedure was designed to consider HI intensity variations across the sky and instrumental effects involving telescope gains. The data were processed using the reduction software developed by HT03a, and the final products for each source include an HI absorption spectrum ($e^{-\tau(v)}$), an “expected” HI emission spectrum ($T_{\text{exp}}(v)$; HI profile that we would observe at the source position if the continuum source were not present), and their uncertainty profiles. Among the 27 sources, 4C +32.14 was excluded from further analyses

because of its saturated absorption spectrum. With an average integration time of 1 hr, the rms noise level in the optical depth profiles was $\sim 1 \times 10^{-3}$ per 1 km s $^{-1}$ velocity channel. Finally, the derived optical depth and “expected” emission spectra were decomposed into separate CNM and WNM components using the technique of HT03a, and physical properties (optical depth, spin temperature, column density, etc.) were computed for the individual components. We refer to Paper I for details on the observations, data reduction, line fitting,⁹ and CNM/WNM properties.

3.2. HI Emission Data from the GALFA-HI Survey

In order to evaluate different methods for deriving the correction for high optical depth, we also use the HI emission data from the GALFA-HI survey (Stanimirović et al. 2006; Peek et al. 2011). GALFA-HI uses ALFA, a seven-beam array of receivers at the focal plane of the Arecibo telescope, to map the HI emission in the Galaxy. Each of the seven dual-polarization beams has an effective beam size of $3'.9 \times 4'.1$.

For Perseus, Lee et al. (2012) produced an HI cube centered at (R.A., decl.) = (03^h29^m52^s, +30°34'1") in J2000¹⁰ with a size of $\sim 15^\circ \times 9^\circ$ by combining a number of individual GALFA-HI projects. We use the same data here, but extend the HI cube up to $\sim 60^\circ \times 18^\circ$ to include all radio continuum sources in Paper I. The HI column density image derived from the extended HI cube is shown in Figure 1 along with our continuum sources (4C +32.14 excluded).

3.3. HI and H₂ Distributions of Perseus

We use the $N(\text{HI})$ and $N(\text{H}_2)$ images from Lee et al. (2012). To derive the $N(\text{HI})$ image, we integrated the HI emission from $v_{\text{LSR}} = -5$ to $+15$ km s $^{-1}$ in the optically thin approximation. This velocity range was determined based on the maximum correlation between the $N(\text{HI})$ image and

⁷ 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.

⁸ In this paper, we quote all velocities in the local standard of rest (LSR) frame, which is defined based on the average velocity of stars in the solar neighborhood: 20 km s $^{-1}$ toward (R.A., decl.) = (18^h, 30°) in B1900.

⁹ Pros and cons of our Gaussian fitting method were discussed in HT03a in detail. In the future, we plan to compare the Gaussian fitting method with results from numerical simulations to investigate biases that could be introduced by Gaussian fitting (R. Lindner et al. 2015, in preparation).

¹⁰ In this paper, we quote all coordinates in J2000.

the Two Micron All Sky Survey (2MASS) based A_V data from the COMPLETE Survey (Ridge et al. 2006). In order to construct the $N(\text{H}_2)$ image of Perseus with a large sky coverage, we used the 60 and 100 μm data from the IRIS survey (Miville-Deschênes & Lagache 2005) and derived the dust optical depth at 100 μm (τ_{100}) by assuming that dust grains are in thermal equilibrium. For this purpose, the emissivity spectral index of $\beta = 2$ was adopted, and the contribution from very small grains (VSGs) to the intensity at 60 μm (I_{60}) was removed by calibrating the derived T_{dust} image with DIRBE-based T_{dust} data from Schlegel et al. (1998). We then converted the τ_{100} image into the A_V image by finding the conversion factor X for $A_V = X\tau_{100}$ that results in the minimum difference between the derived A_V and the COMPLETE A_V . This calibration of τ_{100} to the COMPLETE A_V was motivated by Goodman et al. (2009), who showed that dust extinction at near-infrared wavelengths is the best probe of total gas column density. Finally, we measured a local dust-to-gas ratio (D/G) by examining the A_V - $N(\text{HI})$ relation for diffuse regions and derived the $N(\text{H}_2)$ image by

$$N(\text{H}_2) = \frac{1}{2} \left[\frac{A_V}{\text{D/G}} - N(\text{HI}) \right]. \quad (2)$$

The derived $N(\text{HI})$ and $N(\text{H}_2)$ images are at $4'.3$ resolution (corresponding to ~ 0.4 pc at the distance of 300 pc), and their median 1σ uncertainties are $\sim 5.6 \times 10^{19} \text{ cm}^{-2}$ and $\sim 3.6 \times 10^{19} \text{ cm}^{-2}$. See Sections 3 and 4 of Lee et al. (2012) for details on the derivation of the $N(\text{HI})$ and $N(\text{H}_2)$ images and their uncertainties.

3.4. CO Data from the CfA Survey

We use CO integrated intensity (I_{CO}) data from Dame et al. (2001). Dame et al. (2001) produced a composite CO survey of the Galaxy at $8'.4$ resolution by combining individual observations of the Galactic plane and local molecular clouds. The observations were conducted with the Harvard-Smithsonian Center for Astrophysics (CfA) telescope. The final cube had a uniform rms noise of 0.25 K per 0.65 km s^{-1} velocity channel. To estimate I_{CO} for Perseus, Dame et al. (2001) integrated the CO emission from $v_{\text{LSR}} = -15$ to $+15 \text{ km s}^{-1}$. See Section 2 of Dame et al. (2001) for details on the observations, data reduction, and analyses.

4. CORRECTION FOR HIGH OPTICAL DEPTH

In the direction of 26 radio continuum sources, we measured the optical depth profiles, which we use to estimate the true total HI column density, N_{tot} . Along the same lines of sight, we also have the emission spectra that can be used to calculate the HI column density in the optically thin approximation, $N_{\text{low-}\tau}$. This column density would be the only available information if no HI absorption data were present. In this section, we examine how $f = N_{\text{tot}}/N_{\text{low-}\tau}$, which we call the correction factor for high optical depth, varies with $N_{\text{low-}\tau}$. Our aims are to offer an analytic estimate of $f(N_{\text{low-}\tau})$ for Perseus using our HI emission and absorption measurements, and then to apply this correction to the $N(\text{HI})$ image from Lee et al. (2012) on a pixel-by-pixel basis. In this way, we can account for the optically thick HI that was missed in the GALFA-HI emission observations. Since there are several approaches to derive

$f(N_{\text{low-}\tau})$, we first compare different methods using full line-of-sight information.

4.1. Calculating the HI Column Density in the Optically Thin Approximation ($N_{\text{low-}\tau}$)

As the HI emission spectrum in the direction of a radio continuum source is affected by absorption, it is not possible to obtain an emission profile along exactly the same line of sight as probed by the HI absorption observation. For this reason, we instead estimated the “expected” HI emission spectrum $T_{\text{exp}}(v)$, which is the profile we would observe if the continuum source turned off, by modeling the “17-point pattern” measurements. In this modeling, spatial derivatives of the HI emission (up to the second order) were carefully taken into account (see Section 2.1 of Paper I for details). Then the HI column density in the optically thin approximation, N_{exp} , can be calculated by

$$N_{\text{exp}} (\text{cm}^{-2}) = 1.823 \times 10^{18} \int T_{\text{exp}}(v) dv (\text{K km s}^{-1}). \quad (3)$$

We compute N_{exp} over the velocity range where $T_{\text{exp}}(v)$ is above its 3σ noise level. This velocity range covers all CNM and WNM components for each source, and the median velocity range for all 26 sources is $v_{\text{LSR}} = -39$ to $+20 \text{ km s}^{-1}$. We then estimate the uncertainty in N_{exp} by propagating the $T_{\text{exp}}(v)$ error spectrum through Equation (3), finding a median of $\sim 3.0 \times 10^{18} \text{ cm}^{-2}$.

Additionally, we use spectra from the GALFA-HI cube (pixel size = $1'$) in the direction of our radio continuum sources to derive the “expected” emission profiles. This simpler approach has been employed when HI absorption spectra were obtained without any special strategy such as the “17-point pattern” (e.g., Dickey et al. 2000; McClure-Griffiths et al. 2001; Dickey et al. 2003). We extract HI emission spectra from a 9×9 pixel region (roughly the “17-point pattern” grid size; Figure 1 of HT03a) centered on each continuum source. As the HI spectra right around the continuum source are likely affected by absorption of the background emission, we exclude the HI spectra from the central 3×3 pixel region (roughly one Arecibo beamwidth across). By averaging the remaining 72 spectra, we then compute an average emission spectrum $T_{\text{avg}}(v)$ and estimate the corresponding HI column density N_{cube} by

$$N_{\text{cube}} (\text{cm}^{-2}) = 1.823 \times 10^{18} \int T_{\text{avg}}(v) dv (\text{K km s}^{-1}). \quad (4)$$

The uncertainty in N_{cube} is estimated by calculating the standard deviation of the extracted 72 spectra, $\sigma_{T_{\text{avg}}(v)}$, and propagating it through Equation (4). The median value for all 26 sources is $\sim 1.4 \times 10^{19} \text{ cm}^{-2}$.

In Figure 2, we compare the HI column densities calculated using the two methods. Most sources probe the HI column density of $\sim (5-16) \times 10^{20} \text{ cm}^{-2}$, and the last five sources (3C 132, 3C 131, 3C 133, 4C +27.14, and 4C +33.10) extend this range by a factor of ~ 2 . We find an excellent agreement between the two methods up to $\sim 3 \times 10^{21} \text{ cm}^{-2}$. The ratio of N_{exp} to N_{cube} ranges from ~ 0.9 to ~ 1.1 with a median of ~ 1.0 , suggesting that the two HI column density estimates are consistent within 10%. Considering the more careful examination of spatial variations in the HI emission, we continue by using N_{exp} as $N_{\text{low-}\tau}$ in the following sections.

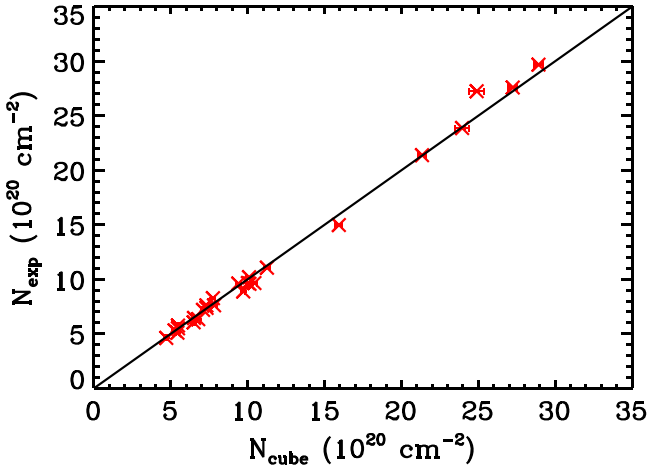


Figure 2. Comparison of the HI column densities calculated using the two different methods: N_{exp} from the spatial derivative method vs. N_{cube} from the simple averaging method. Both quantities are estimated in the optically thin approximation, and the black solid line shows a one-to-one relation.

4.2. METHOD 1—Gaussian Decomposition to Estimate N_{tot}

In Paper I, we performed Gaussian decomposition of the optical depth and “expected” emission spectra and calculated the properties of individual CNM and WNM components while considering self-absorption of both the CNM and the WNM by the CNM. All Gaussian decomposition results (peak brightness temperature, peak optical depth, spin temperature, etc.) for each component are presented in Table 2 of Paper I. These results enable us to derive the true total HI column density along a line of sight by

$$\begin{aligned}
 N_{\text{tot}} (\text{cm}^{-2}) &= N_{\text{CNM}} + N_{\text{WNM}} \\
 &= 1.823 \times 10^{18} \int \left(\sum_0^{N-1} T_{s,n} \tau_{0,n} e^{-[(v-v_{0,n})/\delta v_n]^2} \right. \\
 &\quad \left. + \sum_0^{K-1} T_{0,k} e^{-[(v-v_{0,k})/\delta v_k]^2} \right) dv \quad (\text{K km s}^{-1}), \quad (5)
 \end{aligned}$$

where the components with subscript n refer to the CNM, the components with subscript k refer to the WNM, τ_0 is the peak optical depth, v_0 is the central velocity, T_0 is the peak brightness temperature, and δv is the $1/e$ width of the component. Here N_{tot} is calculated over the velocity range determined as having $T_{\text{exp}}(v)$ higher than its 3σ noise. For the uncertainty in N_{tot} , we use errors of the fitted parameters provided by Gaussian decomposition to perform a Monte Carlo simulation where 1000 N_{CNM} and N_{WNM} values are computed from normally distributed parameters. The standard deviations of the N_{CNM} and N_{WNM} distributions are then added in quadrature to estimate the uncertainty in N_{tot} . This method of deriving N_{tot} was used by HT03b for their HI absorption measurements toward background sources randomly located over the whole Arecibo sky. In this study, we focus on a localized group of background sources in the direction of Perseus.

The (integrated) correction factor, $f = N_{\text{tot}}/N_{\text{low-}\tau} = N_{\text{tot}}/N_{\text{exp}}$, is shown in Figure 3 (left) as a function of N_{exp} . Clearly, the correction factor increases with N_{exp} . We then present the $(1/\sigma^2)$ -weighted mean values as the blue squares and

the linear fit to all 26 data points as the green solid line¹¹:

$$\begin{aligned}
 f &= \log_{10}(N_{\text{exp}}/10^{20}) \times a + b \\
 &= \log_{10}(N_{\text{exp}}/10^{20})(0.32 \pm 0.06) + (0.81 \pm 0.05). \quad (6)
 \end{aligned}$$

In general, the correction factor ranges from ~ 1.0 at $\sim 3.9 \times 10^{20} \text{ cm}^{-2}$ to ~ 1.2 at $\sim 1.3 \times 10^{21} \text{ cm}^{-2}$ (maximum uncorrected HI column density in Perseus). While f and N_{exp} show a good correlation (Spearman’s rank correlation coefficient of 0.80), there are two sources with relatively high correction factors at $\sim 10^{21} \text{ cm}^{-2}$, 3C 092 and 3C 093.1. Interestingly, the two are located behind the main body of Perseus (Figure 1). Their high f values of ~ 1.5 – 1.6 could result from an increased amount of the cold HI in the molecular cloud relative to the surrounding diffuse ISM. The CNM fraction is indeed ~ 0.4 for both sources, which is higher than the median value of ~ 0.3 for all 26 sources (Section 4.3 of Paper I). However, this is not the maximum CNM fraction in our measurements (~ 0.6). Observing a denser grid of radio continuum sources behind Perseus and repeating the calculations would be an interesting way to test the cold HI hypothesis. Finally, the correction factor is also presented as a function of the integrated optical depth in Figure 3 (right). As expected, there is a clear correlation (Spearman’s rank correlation coefficient of 0.94).

We note that our results are not sensitive to HI components at $v_{\text{LSR}} < -20 \text{ km s}^{-1}$, which are likely unassociated with Perseus (Section 4.2 of Paper I): limiting the calculation of N_{exp} and N_{tot} to $v_{\text{LSR}} > -20 \text{ km s}^{-1}$ or excluding the five sources showing such HI components at large negative velocities (corresponding to the sources with $\log_{10}(N_{\text{exp}}/10^{20}) > 1.2$ in Figure 3; 3C 132, 3C 131, 3C 133, 4C +27.14, and 4C +33.10) results in linear fit coefficients that are consistent with what we present here within uncertainties.¹²

4.3. METHOD 2—Isothermal Estimate of N_{tot}

By assuming that each velocity channel represents gas at a single temperature, Dickey et al. (2000) showed that the correction factor per velocity channel can be written as

$$f_{\text{chan}}(v) = \frac{C_0 T_s(v) \tau(v)}{C_0 T_{\text{exp}}(v)}, \quad (7)$$

where $C_0 = 1.823 \times 10^{18} \text{ cm}^{-2}/(\text{K km s}^{-1})$. In addition, $T_{\text{exp}}(v)$ was expressed as

$$T_{\text{exp}}(v) = T_s(v) (1 - e^{-\tau(v)}). \quad (8)$$

¹¹ In attempting to be consistent with our calculation of the uncertainty in f for the isothermal method (Section 4.3), we run a full Monte Carlo simulation based on errors of the fitted parameters from Gaussian decomposition. In this simulation, 1000 N_{exp} and N_{tot} values are calculated from normally distributed parameters, and the standard deviation of 1000 f values is used as the uncertainty in f . We find that linear fit results from using this error estimate, $f = \log_{10}(N_{\text{exp}}/10^{20})(0.25 \pm 0.03) + (0.87 \pm 0.02)$, are consistent with Equation (6) within uncertainties.

¹² To be specific, limiting the calculation of N_{exp} and N_{tot} to $v_{\text{LSR}} > -20 \text{ km s}^{-1}$ results in $f = \log_{10}(N_{\text{exp}}/10^{20})(0.44 \pm 0.07) + (0.73 \pm 0.06)$ and $f = (0.034 \pm 0.005) \int \tau(v) dv + (1.004 \pm 0.020)$. Similarly, excluding the five sources that show the HI components at $v_{\text{LSR}} < -20 \text{ km s}^{-1}$ leads to $f = \log_{10}(N_{\text{exp}}/10^{20})(0.31 \pm 0.11) + (0.82 \pm 0.09)$ and $f = (0.040 \pm 0.011) \int \tau(v) dv + (0.988 \pm 0.027)$.

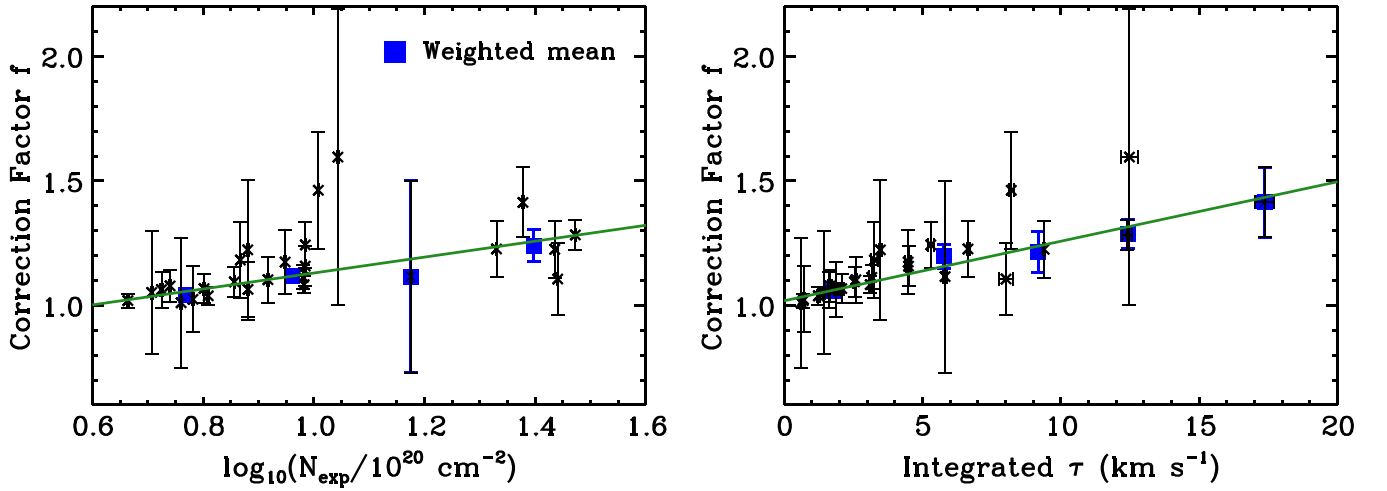


Figure 3. METHOD 1. Left: $f = N_{\text{tot}}/N_{\text{exp}}$ (N_{tot} : derived using the Gaussian decomposition results) as a function of $\log_{10}(N_{\text{exp}}/10^{20})$. The blue squares show the $(1/\sigma^2)$ -weighted mean values in 0.2-wide bins in $\log_{10}(N_{\text{exp}}/10^{20})$, and the linear fit determined for all 26 data points is indicated as the green solid line (Equation (6)). Right: f as a function of the integrated optical depth. The $(1/\sigma^2)$ -weighted mean values in 3.4 km s^{-1} -wide bins in the integrated optical depth are presented as the blue squares, and the green solid line shows the linear fit to all 26 data points: $f = (0.024 \pm 0.004) \int \tau(\nu) d\nu + (1.019 \pm 0.019)$.

This equation assumes the absence of any radio continuum source behind the absorbing HI cloud. As a result, $f_{\text{chan}}(\nu)$ simply becomes

$$f_{\text{chan}}(\nu) = \frac{\tau(\nu)}{1 - e^{-\tau(\nu)}}. \quad (9)$$

However, while the radio continuum source is absent in Equation (8), some diffuse radio continuum emission is always present and should not be ignored. This emission includes the CMB and the Galactic synchrotron emission that varies across the sky and becomes strong toward the Galactic plane. We refer to a combination of these contributions as T_{sky} , and Equation (8) then has to be rewritten as

$$T_{\text{exp}}^*(\nu) = T_s(\nu)(1 - e^{-\tau(\nu)}) + T_{\text{sky}}e^{-\tau(\nu)}. \quad (10)$$

Considering that HI emission spectra are generally baseline subtracted during the reduction process, T_{sky} can be removed from both sides of Equation (10). Then $T_{\text{exp}}(\nu) = T_{\text{exp}}^*(\nu) - T_{\text{sky}}$, the quantity we have been working with so far, can be expressed as

$$T_{\text{exp}}(\nu) = T_{\text{exp}}^*(\nu) - T_{\text{sky}} = (T_s(\nu) - T_{\text{sky}})(1 - e^{-\tau(\nu)}). \quad (11)$$

As a consequence, the correction factor becomes

$$f_{\text{chan}}(\nu) = \frac{T_s(\nu)}{T_s(\nu) - T_{\text{sky}}} \frac{\tau(\nu)}{1 - e^{-\tau(\nu)}}, \quad (12)$$

or with direct observables,

$$f_{\text{chan}}(\nu) = T_{\text{sky}} \frac{\tau(\nu)}{T_{\text{exp}}(\nu)} + \frac{\tau(\nu)}{1 - e^{-\tau(\nu)}}. \quad (13)$$

In order to estimate the contribution from the Galactic synchrotron emission, we use the Haslam et al. (1982) 408 MHz survey of the Galaxy. The brightness temperature at 408 MHz is converted to 1.4 GHz using the spectral index of -2.7 . As the absolute Galactic latitude of our continuum

sources is generally higher than 10° , the synchrotron contribution is small, with T_{sky} ranging from 2.78 to 2.80 K (Table 1 of Paper I). Based on the histogram of T_s for the individual CNM components (Figure 5(b) of Paper I), we can then provide a rough estimate of $T_s(\nu)/[T_s(\nu) - T_{\text{sky}}]$: the expected range is narrow, from ~ 1.0 to ~ 1.2 . Clearly, for molecular clouds located closer to the Galactic plane the contribution from the diffuse radio continuum emission will be more significant.

In Equation (7), $f_{\text{chan}}(\nu)$ essentially represents the correction that needs to be applied to $T_{\text{exp}}(\nu)$ to calculate the true brightness temperature profile. As a result, the true total HI column density can be obtained by

$$N_{\text{tot}} (\text{cm}^{-2}) = 1.823 \times 10^{18} \times \int f_{\text{chan}}(\nu) T_{\text{exp}}(\nu) d\nu (\text{K km s}^{-1}). \quad (14)$$

We derive $f_{\text{chan}}(\nu)$ for all 26 sources (Equation (13)) and present the results in Figure 4. About 91% of the $f_{\text{chan}}(\nu)$ values are between 1 and 2, and the fraction of velocity channels with $f_{\text{chan}}(\nu) > 2$ is very small. We then calculate N_{tot} using Equation (14) and show the (integrated) correction factor, $f = N_{\text{tot}}/N_{\text{exp}}$, as a function of N_{exp} in Figure 5 (left). Here the integration is done over the velocity range where $T_{\text{exp}}(\nu)$ is higher than its 3σ noise. The uncertainty in f is derived by running a Monte Carlo simulation where the optical depth and “expected” emission error spectra are propagated through Equations (13) and (14) to compute 1000 f values. The standard deviation of the f distribution is used as the final uncertainty in f .

Similar to the Gaussian decomposition method, we find a good correlation between f and N_{exp} (Spearman’s rank correlation coefficient of 0.84). The linear fit determined using all 26 data points is

$$f = \log_{10}(N_{\text{exp}}/10^{20})(0.25 \pm 0.02) + (0.87 \pm 0.02). \quad (15)$$

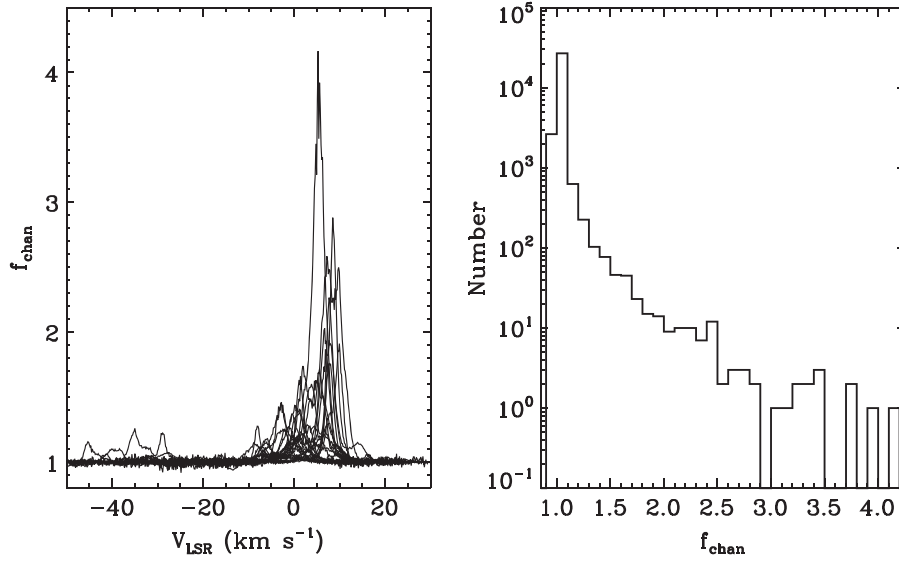


Figure 4. METHOD 2. Left: correction factor per velocity channel $f_{\text{chan}}(v)$ (estimated in the isothermal approximation) for all 26 sources. Right: histogram of $f_{\text{chan}}(v)$ values.

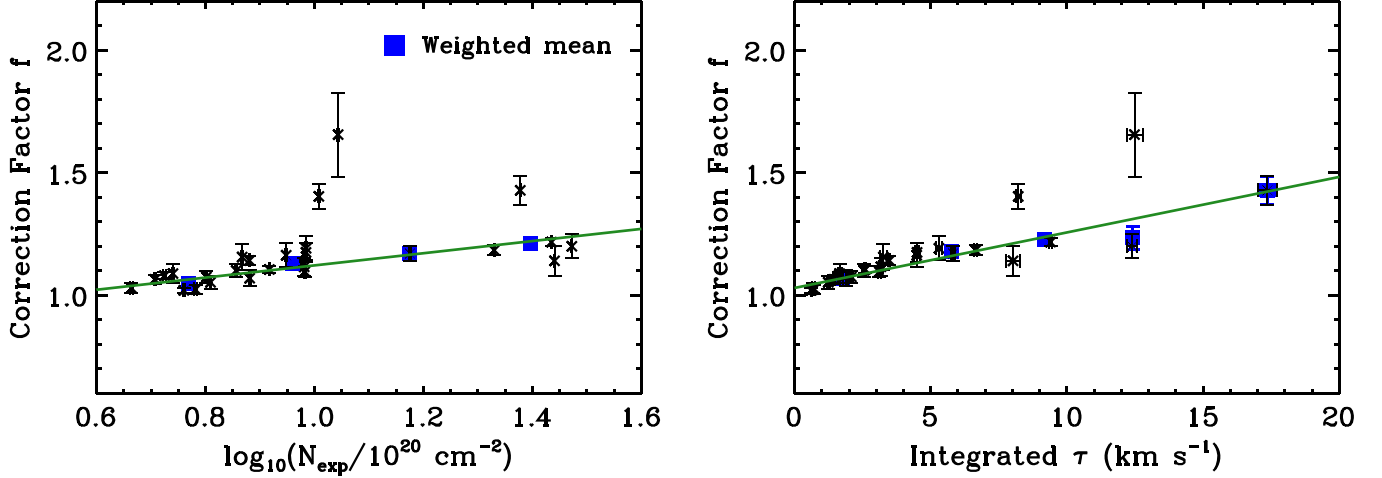


Figure 5. METHOD 2. Left: $f = N_{\text{tot}}/N_{\text{exp}}$ (N_{tot} : estimated in the isothermal approximation) as a function of $\log_{10}(N_{\text{exp}}/10^{20})$. The blue squares show the $(1/\sigma^2)$ -weighted mean values in 0.2-wide bins in $\log_{10}(N_{\text{exp}}/10^{20})$, and the linear fit determined for all 26 data points is indicated as the green solid line (Equation (15)). Right: f as a function of the integrated optical depth. The $(1/\sigma^2)$ -weighted mean values in 3.4 km s^{-1} -wide bins in the integrated optical depth are presented as the blue squares, and the green solid line shows the linear fit to all 26 data points: $f = (0.023 \pm 0.001) \int \tau(v) dv + (1.030 \pm 0.006)$.

Additionally, f is plotted as a function of the integrated optical depth in Figure 5 (right), again showing a clear correlation (Spearman’s rank correlation coefficient of 0.97).

Both graphs in Figure 5 are very similar to those in Figure 3 for the Gaussian decomposition method. Specifically, the linear fit coefficients are consistent within uncertainties. This is surprising considering that the two methods are very different. In particular, the isothermal method assigns a single spin temperature to each velocity channel, while the Gaussian decomposition method allows a single velocity channel to have contributions from several HI components with different spin temperatures.

In Dickey et al. (2000), the authors updated the isothermal method by incorporating the two-phase approximation. As input parameters, this method then required the spin temperature of the cold HI and the fraction of the warm HI that is in front of the cold HI, the quantity they referred to as q . Dickey et al. (2000) showed that for their SMC data there is no

difference between the one- and two-phase approximations regarding the correction factor if $q \gtrsim 0.5$, while the difference becomes more pronounced when $q \lesssim 0.25$. In the Gaussian decomposition method, this fraction (F_k in Paper I) is important as well, but it is difficult to constrain. Thus, the fitting process was repeated for $F_k = 0, 0.5$, and 1, and these results were used to estimate the final uncertainty in T_s (see Section 3 of Paper I and HT03a for details).

To understand why the two different methods result in similar correction factors, we compare Equations (5) and (14) (Appendix A) and find that they become comparable regardless of F_k when $\tau \ll 1$ and $T_{\text{sky}} \ll T_s$. In our observations of Perseus, the median peak optical depth for all CNM components is ~ 0.2 , and only a small number of the components has the peak optical depth higher than 1 (10 out of 107; Section 4.1 of Paper I), satisfying the condition. In addition, we already showed that T_{sky} is small, with $\sim 2.8 \text{ K}$ for Perseus. The difference between the Gaussian decomposition

and isothermal methods, however, will be more significant for molecular clouds that have a large amount of the cold, optically thick HI and/or a substantial contribution from the diffuse radio continuum emission. Due to the more self-consistent way to derive $T_{\text{exp}}(\nu)$, we continue by using the Gaussian decomposition results for further analyses.

Finally, we note that Equations (6) and (15) could be biased against very high optical depths as they result in saturated absorption spectra, e.g., 4C +32.14, the source we had to exclude from our analyses due to its highly saturated absorption profile. In addition, the equations are based on our explicit assumption of the linear relation between f and $\log_{10}(N_{\text{exp}}/10^{20})$. The HI emission and absorption measurements obtained by HT03a and HT03b along many random lines of sight through the Galaxy show that the linear relation indeed describes the observations well up to $\log_{10}(N_{\text{exp}}/10^{20}) \sim 1.5$, and the fitted coefficients are consistent with Equations (6) and (15) within uncertainties (Appendix B). There is some interesting deviation from the linear relation, however, particularly for a few sources with $\log_{10}(N_{\text{exp}}/10^{20}) \gtrsim 1$. This deviation could suggest a non-linear relation at high column densities, and its significance needs to be further examined with more HI absorption measurements. In the future, it will also be important to study the dense CNM using alternative tracers such as CI and CII fine-structure lines (e.g., Pineda et al. 2013).

4.4. Comparison with Previous Studies

Dickey et al. (2000). The correction factor calculated by Dickey et al. (2000) for the SMC using full line-of-sight information can be rewritten as $f = \log_{10}(N_{\text{exp}}/10^{20})0.667 + 0.066$ for $N_{\text{exp}} > 2.5 \times 10^{21} \text{ cm}^{-2}$. The range of the HI column density we probe for Perseus barely overlaps with that in Dickey et al. (2000), as the HI column density in the low-metallicity SMC is significantly higher compared to the Galaxy. While the difference with Dickey et al. (2000) in f values depends on N_{exp} , our correction factor is only $\sim 4\%$ higher than what Dickey et al. (2000) suggest when extrapolated to their maximum HI column density of 10^{22} cm^{-2} . Similarly, Dickey et al. (2003) used HI emission and absorption data from the Southern Galactic Plane Survey (McClure-Griffiths et al. 2005) in combination with the isothermal method and estimated $f = \sim 1.4\text{--}1.6$ for sources located at $326^\circ < l < 333^\circ$ and $|b| \lesssim 1^\circ$. This correction factor is comparable with what we find for Perseus.

Heiles & Troland (2003a, 2003b). HT03a and HT03b performed Gaussian decomposition of 79 HI emission/absorption spectral line pairs and derived the correction factor, which they called $R_{\text{raw}} = 1/f$. They found that R_{raw} ranges from ~ 0.3 to ~ 1.0 , corresponding to $f = \sim 1.0\text{--}3.0$ (Appendix B). In particular, they estimated $f \sim 1.3$ for the Taurus/Perseus region, which is similar to what we find for Perseus.

Braun et al. (2009). Our correction factor is smaller than what Braun et al. (2009) and Braun (2012) claimed for M31, M33, and the LMC based on the modeling of HI emission spectra. They found that the correction exceeds an order of magnitude in many cases and increases the global HI mass by $\sim 30\%$. Even when considering the correction factor per velocity channel, we find the maximum $f_{\text{chan}}(\nu) \sim 4$ for only one source and $f_{\text{chan}}(\nu) \lesssim 3$ for the rest of our sources.

Chengalur et al. (2013). Our correction factor can also be compared with predictions from Chengalur et al. (2013). This study performed Monte Carlo simulations where observationally motivated input parameters such as the column density, the spin temperatures of the CNM and the WNM, and the fraction of gas in each of the different phases were provided for ISM models. Our correction factor versus integrated optical depth plots, Figures 3 (right) and 5 (right), can be directly compared with Figure 1(A) in Chengalur et al. (2013). We find that the correction factor by Chengalur et al. (2013) is significantly higher than our estimate, although the general trend of increasing correction with the integrated optical depth is similar. For example, Chengalur et al. (2013) expect $f \sim 20$ when $\int \tau d\nu \sim 10 \text{ km s}^{-1}$, while we find only $f < 1.5$. Similarly, our Figures 3 (left) and 5 (left) can be compared with Figure 2 (A) in Chengalur et al. (2013). We find that our estimate is consistent with the correction factor by Chengalur et al. (2013) for the column density less than 10^{21} cm^{-2} , while Chengalur et al. (2013) overestimate at the high end of our column density range. If we extrapolate our relation to 10^{22} cm^{-2} , we expect a ~ 10 times lower correction factor than what Chengalur et al. (2013) suggest. The reason for their very high correction factor could be the inclusion of extremely high column densities ($10^{23}\text{--}10^{24} \text{ cm}^{-2}$) in their ISM models, although it is not clear why it less affects the isothermal estimate of the HI column density. Another possible cause for the high correction factor is the neglect of galactic dynamics (e.g., differential rotation) in their Monte Carlo simulations. As recently shown by Kim et al. (2014), consideration of large-scale gas motions is critical to model an overlap between CNM clouds in velocity space. Without proper modeling of dynamics, therefore, CNM clouds could be overcrowded along the velocity axis, resulting in high correction factors.

Liszt (2014). Using the HI absorption data compiled by Liszt et al. (2010), Liszt (2014) estimated the correction factor for radio continuum sources located at high Galactic latitudes. While they did not provide detailed information about how exactly the derivation was done, their correction factor was small with f less than 1.2 for $E(B - V) \lesssim 0.5$ mag. This is comparable to our finding.

Fukui et al. (2014, 2015). Finally, Fukui et al. (2015) estimated the correction factor for the Galaxy at $|b| > 15^\circ$ by exploring the relation between τ_{353} and $N(\text{HI})$ at $33'$ resolution. Their Figure 13 shows that the correction factor distribution ranges from ~ 1.0 to ~ 3.0 and peaks at ~ 1.5 . Using the same methodology, Fukui et al. (2014) found that the correction increases the total HI mass of the high-latitude molecular clouds MBM 53, MBM 54, MBM 55, and HLCG 92-35 by a factor of ~ 2 . In general, the correction factor by Fukui et al. (2014, 2015) appears higher than our estimate for Perseus. While we do not perform a detailed comparison with Fukui et al. (2014, 2015), we test their claim of the optically thick HI as an alternative of the “CO-dark” gas in Section 7.

5. APPLYING THE CORRECTION FOR HIGH OPTICAL DEPTH TO PERSEUS

We apply the correction derived in Section 4.2 to the $N(\text{HI})$ image of Perseus from Lee et al. (2012) for pixels with $\log_{10}(N_{\text{exp}}/10^{20}) > 0.6$ (where the correction factor is higher than 1). The correction factor and corrected $N(\text{HI})$ images are shown in Figures 6 and 7. In addition, we present the normalized

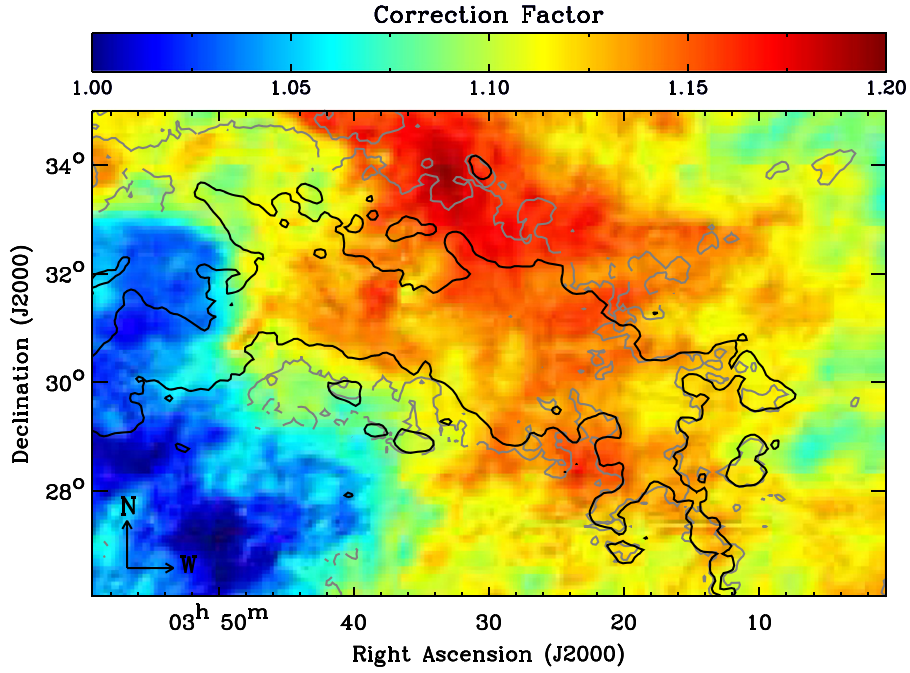


Figure 6. Correction factor at 4/3 resolution estimated in Section 4.2 (Equation (6)). In addition, the 3σ contour of the “old” $N(\text{H}_2)$ from Lee et al. (2012; before the correction for high optical depth) is overlaid in gray, while the 3σ contour of the CfA I_{CO} is shown in black. The resolutions of the $N(\text{H}_2)$ and I_{CO} images are 4/3 and 8/4.

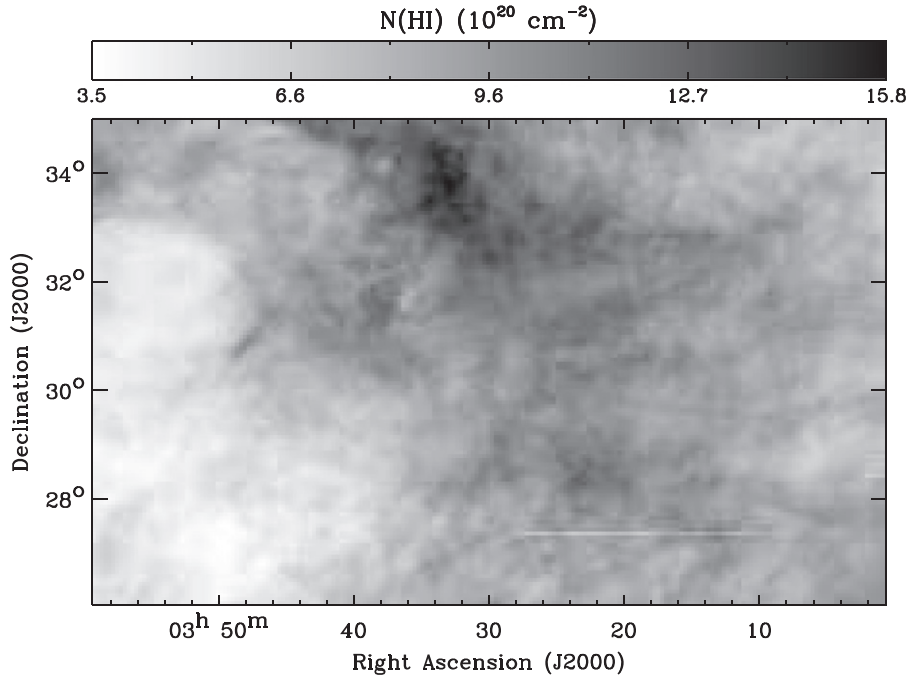


Figure 7. Corrected $N(\text{HI})$ image at 4/3 resolution.

histograms of the two $N(\text{HI})$ images, before and after the correction, in Figure 8.

As Figure 8 indicates, the correction does not make a significant change in $N(\text{HI})$. To be specific, the median $N(\text{HI})$ increases by a factor of ~ 1.1 from $\sim 7.9 \times 10^{20} \text{ cm}^{-2}$ to $\sim 8.7 \times 10^{20} \text{ cm}^{-2}$, while the maximum $N(\text{HI})$ increases by a factor of ~ 1.2 from $\sim 1.3 \times 10^{21} \text{ cm}^{-2}$ to $\sim 1.6 \times 10^{21} \text{ cm}^{-2}$. In terms of the total HI mass, the correction results in a $\sim 10\%$ increase from $\sim 2.3 \times 10^4 M_\odot$ to $\sim 2.5 \times 10^4 M_\odot$. This increase

in the HI mass is comparable to what Dickey et al. (2000) found for the SMC, but is smaller than the value estimated by Braun et al. (2009) and Braun (2012) for M31, M33, and the LMC. In addition, we note that our correction factor image looks smooth compared to a granulated appearance of the corrected $N(\text{HI})$ images by Braun et al. (2009) and Braun (2012), although Perseus ($\sim 80 \text{ pc} \times 50 \text{ pc}$) would be unresolved or only marginally resolved in their studies. Finally, the HI mass increase due to the optically thick gas in Perseus is

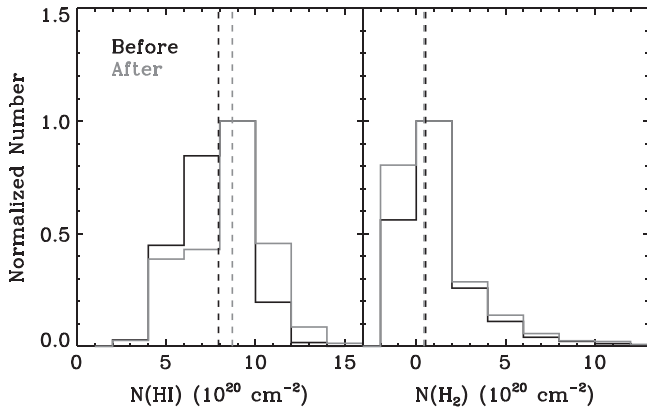


Figure 8. Left: normalized histograms of the two $N(\text{HI})$ images, before (black) and after (gray) the correction for high optical depth. The median of each histogram is shown as the dashed line. Right: same as the left panel, but for $N(\text{H}_2)$.

smaller than what Fukui et al. (2014, 2015) derived for Galactic molecular clouds.

6. HOW DOES THE HIGH OPTICAL DEPTH AFFECT THE HI SATURATION IN PERSEUS?

6.1. Rederiving $N(\text{H}_2)$

To investigate the impact of high optical depth on the observed HI saturation in Perseus, we first rederive the $N(\text{H}_2)$ image using the corrected $N(\text{HI})$. In essence, we use the same methodology as Lee et al. (2012): the A_V image is derived using the IRIS 60/100 μm and 2MASS A_V data, and a local D/G is adopted to estimate $N(\text{H}_2)$. We refer to Section 4 of Lee et al. (2012) for details on the method for deriving $N(\text{H}_2)$ and its limitation, and we summarize here main results.

1. Contamination from VSGs: to estimate T_{dust} from the ratio of I_{60} to I_{100} , the contribution from stochastically heated VSGs to I_{60} must be removed. For this purpose, we compare our IRIS-based T_{dust} with the DIRBE-based T_{dust} from Schlegel et al. (1998) and find that the contribution from VSGs to I_{60} is 78%. This is the same as what Lee et al. (2012) found.
2. Zero-point calibration for τ_{100} : we refine the zero point of the τ_{100} image by assuming that the dust column density traced by τ_{100} is proportional to $N(\text{HI})$ for atomic-dominated regions. Based on the zero point of the τ_{100} – $N(\text{HI})$ relation, we add 1.1×10^{-4} to the τ_{100} image. This is slightly smaller than what Lee et al. (2012) added, i.e., 1.8×10^{-4} .
3. Conversion from τ_{100} to A_V : we convert τ_{100} into A_V by adopting $X = 740$ for $A_V = X\tau_{100}$ that results in the best agreement between our IRIS-based A_V and the 2MASS-based A_V from Ridge et al. (2006). This is slightly higher than $X = 720$ used by Lee et al. (2012). We compare our rederived A_V with the A_V image from Lee et al. (2012) and find that the ratio of the new A_V to the old A_V ranges from ~ 0.93 to ~ 1.02 .
4. Deriving a local D/G and $N(\text{H}_2)$: we examine the A_V – $N(\text{HI})$ relation and find that the slope of $A_V/N(\text{HI}) = 1 \times 10^{-21} \text{ mag cm}^2$ is a good measure of D/G for Perseus. This is slightly lower than what Lee et al. (2012) estimated, i.e., $1.1 \times 10^{-21} \text{ mag cm}^2$, and makes

sense considering that our rederived A_V is essentially the same with the A_V image from Lee et al. (2012), while the corrected $N(\text{HI})$ is slightly higher than the uncorrected $N(\text{HI})$. Finally, we derive $N(\text{H}_2)$ using Equation (2) and mask pixels with possible contaminations (point sources, the Taurus molecular cloud, and the “warm dust ring”), following what Lee et al. (2012) did. We show the final $N(\text{H}_2)$ image in Figure 9, as well as the normalized histograms of the new (this study) and old (Lee et al. 2012) $N(\text{H}_2)$ images in Figure 8 (right).

In conclusion, in our rederivation of $N(\text{H}_2)$, all parameters are identical to or only slightly different from what Lee et al. (2012) used. As a result, the new $N(\text{H}_2)$ is comparable to the old $N(\text{H}_2)$, shown as a good agreement between the two histograms in Figure 8 (right). The rederived $N(\text{H}_2)$ ranges from $-1.5 \times 10^{20} \text{ cm}^{-2}$ to $5.1 \times 10^{21} \text{ cm}^{-2}$ with a median of $4.6 \times 10^{19} \text{ cm}^{-2}$, and $\sim 83\%$ of the pixels whose signal-to-noise ratio is higher than 1 have the new $N(\text{H}_2)$ differing from the old $N(\text{H}_2)$ by only 10%. Finally, we note that $\sim 30\%$ of all finite pixels have negative $N(\text{H}_2)$ values of mostly around $-(1-5) \times 10^{19} \text{ cm}^{-2}$, which are very close to zero considering the median uncertainty in $N(\text{H}_2)$ of $\sim 5 \times 10^{19} \text{ cm}^{-2}$ (Section 6.2).

6.2. Uncertainty in $N(\text{H}_2)$

As Lee et al. (2012) did, we perform a series of Monte Carlo simulations to estimate the uncertainty in $N(\text{H}_2)$. In these simulations, we assess the errors in $N(\text{HI})$ and A_V and propagate them together.

For the uncertainty in $N(\text{HI})$, we combine two terms in quadrature: the error from using a fixed velocity width ($\sigma_{\text{HI},1}$), and the error from the correction for high optical depth ($\sigma_{\text{HI},2}$). To estimate $\sigma_{\text{HI},1}$, we produce 1000 $N(\text{HI})$ images using 1000 velocity widths randomly drawn from a Gaussian distribution that peaks at 20 km s^{-1} with 1σ of 4 km s^{-1} . The standard deviation of the simulated $N(\text{HI})$ is then computed for $\sigma_{\text{HI},1}$. This $\sigma_{\text{HI},1}$ is what Lee et al. (2012) used as their uncertainty in $N(\text{HI})$. Similarly, to derive $\sigma_{\text{HI},2}$, we generate 1000 $N(\text{HI})$ images by applying the correction to the $N(\text{HI})$ image from Lee et al. (2012) using 1000 combinations of a and b in Equation (6). These a and b values are again drawn from Gaussian distributions whose peaks and widths correspond to the fitted a and b values and their uncertainties. We find that the median 1σ of $N(\text{HI})$ in our study is $\sim 8.1 \times 10^{19} \text{ cm}^{-2}$.

For the uncertainty in A_V , we repeat the exercise done by Lee et al. (2012): deriving a number of A_V images by changing input conditions (1σ noises of the IRIS 60/100 μm images, β , zero-point calibration for τ_{100}), and estimating the minimum and maximum A_V values for each pixel. In this exercise, we find that the contribution from VSGs to I_{60} varies from 76% to 88%, while X varies from 655 to 855.

Finally, we propagate the uncertainty in $N(\text{HI})$ and the minimum/maximum A_V values through a Monte Carlo simulation in order to produce 1000 $N(\text{H}_2)$ images. The distribution of the simulated $N(\text{H}_2)$ is then used to estimate the uncertainty in $N(\text{H}_2)$ on a pixel-by-pixel basis. We find that the median 1σ of $N(\text{H}_2)$ in our study is $\sim 4.8 \times 10^{19} \text{ cm}^{-2}$.

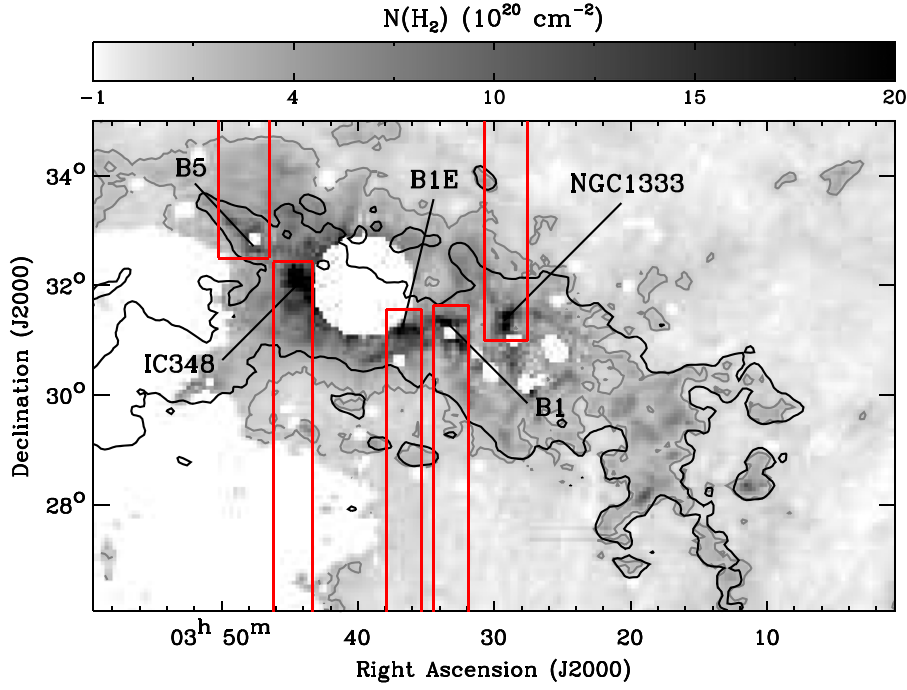


Figure 9. Rederived $N(\text{H}_2)$ image. The blank pixels correspond to regions with possible contaminations (point sources, the Taurus molecular cloud, and the “warm dust ring”; see Section 4 of Lee et al. 2012 for details). The 3σ contour of the new $N(\text{H}_2)$ is overlaid in gray, while the 3σ contour of the CfA I_{CO} is shown in black. The resolutions of the $N(\text{H}_2)$ and I_{CO} images are $4''.3$ and $8''.4$. The red rectangular boxes show the boundaries of the selected dark (B5, B1E, and B1) and star-forming (IC 348 and NGC 1333) regions.

6.3. R_{H_2} versus $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$ and Σ_{HI} versus $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$

From the rederived $N(\text{HI})$ and $N(\text{H}_2)$ images, we estimate Σ_{HI} and Σ_{H_2} by

$$\begin{aligned} \Sigma_{\text{HI}} (M_{\odot} \text{ pc}^{-2}) &= \frac{N(\text{HI}) (\text{cm}^{-2})}{1.25 \times 10^{20}} \\ \Sigma_{\text{H}_2} (M_{\odot} \text{ pc}^{-2}) &= \frac{N(\text{H}_2) (\text{cm}^{-2})}{6.25 \times 10^{19}}. \end{aligned} \quad (16)$$

We find that Σ_{HI} varies by only a factor of ~ 2.6 from ~ 4.8 to $\sim 12.7 M_{\odot} \text{ pc}^{-2}$. On the other hand, Σ_{H_2} ranges from -2.4 to $81.8 M_{\odot} \text{ pc}^{-2}$, although $\sim 98\%$ of the pixels have $\Sigma_{\text{H}_2} < 15 M_{\odot} \text{ pc}^{-2}$. As a result, $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$ has a small dynamic range of $\sim 8\text{--}30 M_{\odot} \text{ pc}^{-2}$ across most of the cloud.

To compare with the KMT09 predictions aiming at revisiting the HI saturation in Perseus, we focus on the individual dark (B5, B1E, and B1) and star-forming (IC 348 and NGC 1333) regions. The boundaries of each region were determined based on the ^{13}CO emission (see Section 5 of Lee et al. 2012 for details) and are shown as the red rectangular boxes in Figure 9. Using all finite pixels in the rectangular boxes, we plot Σ_{HI} and R_{H_2} as a function of $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$ for each region and present the results for IC 348 and B1E in Figures 10 and 11. Additionally, we show the same plots from Lee et al. (2012) for comparison. Note that both this study and Lee et al. (2012) include negative Σ_{H_2} values by using all finite pixels in the rectangular boxes. Almost all ($\sim 90\%$) of these negative Σ_{H_2} values fluctuate around zero within uncertainties.

6.4. Comparison to the KMT09 Predictions

As in Lee et al. (2012), the following KMT09 predictions are used to fit the observed R_{H_2} versus $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$ profiles:

$$R_{\text{H}_2} = \frac{4\tau_c}{3\psi} \left[1 + \frac{0.8\psi\phi_{\text{mol}}}{4\tau_c + 3(\phi_{\text{mol}} - 1)\psi} \right] - 1 \quad (17)$$

where

$$\tau_c = \frac{3}{4} \left(\frac{\Sigma_{\text{comp}} \sigma_d}{2.5 + \chi} \right), \quad (18)$$

$$\psi = \chi \frac{2.5 + \chi}{2.5 + \chi e}, \quad (19)$$

and

$$\chi = 2.3 \frac{1 + 3.1Z'^{0.365}}{\phi_{\text{CNM}}}. \quad (20)$$

Here τ_c is the dust optical depth a spherical cloud would have if its HI and H_2 are uniformly mixed, and χ is the ratio of the rate at which LW photons are absorbed by dust grains to the rate at which they are absorbed by H_2 . In addition, Σ_{comp} is the total gas column density, σ_d is the dust absorption cross section per hydrogen nucleus in the LW band, μ_{H} is the mean mass per hydrogen nucleus, Z' is the metallicity normalized to the value in the solar neighborhood, ϕ_{CNM} is the ratio of the CNM density to the minimum CNM density at which the CNM can be in pressure balance with the WNM, and finally ϕ_{mol} is the ratio of the H_2 density to the CNM density. We refer to Section 6 of Lee et al. (2012) for a detailed summary of the KMT09 model.

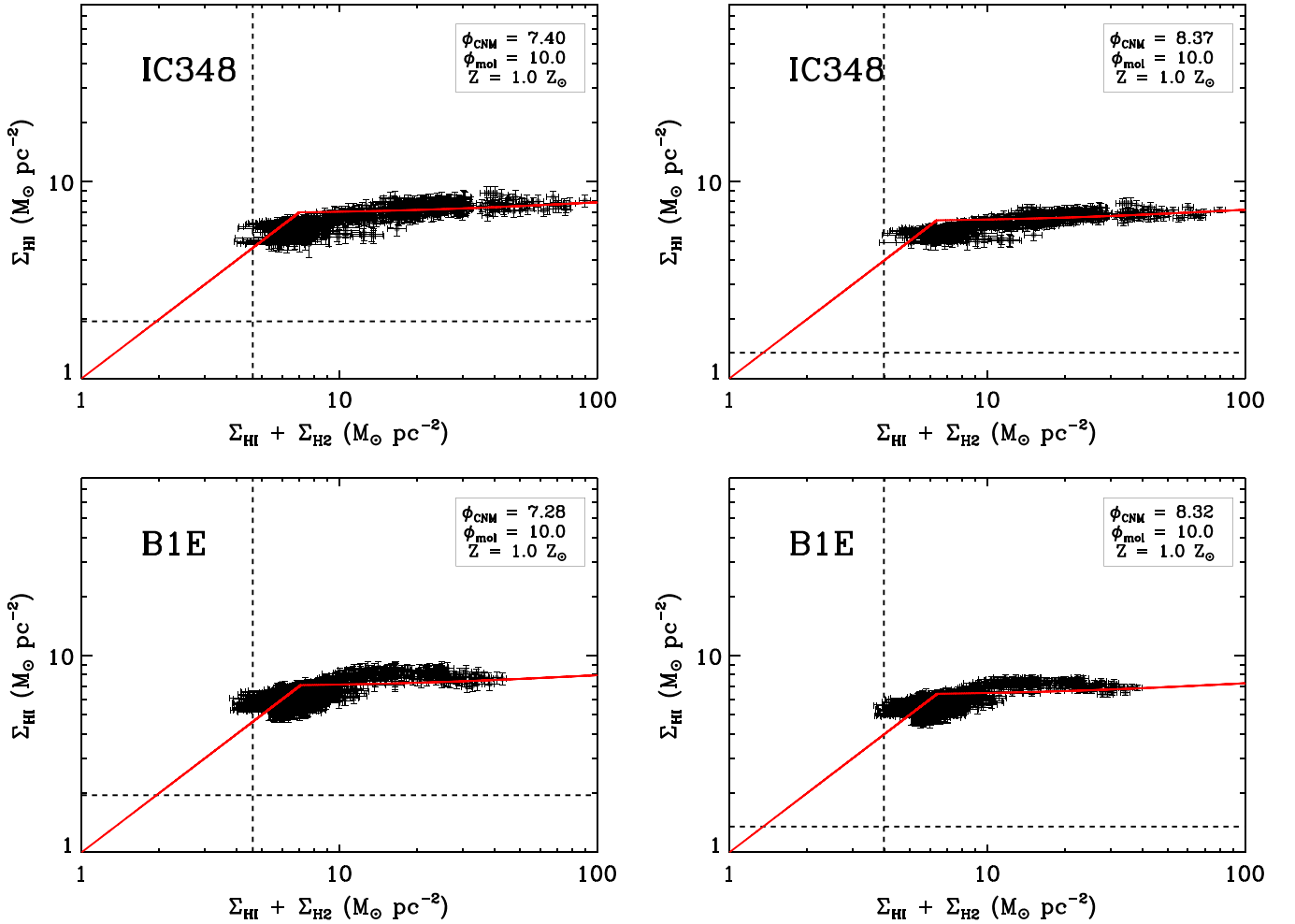


Figure 10. Left: Σ_{HI} vs. $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$ for IC 348 and B1E (this study). All finite pixels in the rectangular boxes shown in Figure 9 are used for plotting. The median 3σ values of Σ_{HI} and $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$ are indicated as the black dashed lines, while the best-fit curves determined in Section 6.4 are overlaid in red. The best-fit parameters are shown in the top right corner of each plot. Right: plots from Lee et al. (2012).

As Equations (17)–(20) suggest, R_{H_2} in the KMT09 model is simply a function of total gas column density, metallicity, ϕ_{CNM} , and ϕ_{mol} and is independent of the strength of the radiation field in which the cloud is embedded. Following Lee et al. (2012), we adopt $Z' = 1$ and $\phi_{\text{mol}} = 10$ (fiducial value by KMT09)¹³ and constrain ϕ_{CNM} by finding the best-fit model for the observed R_{H_2} versus $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$. For this purpose, we perform Monte Carlo simulations where the uncertainties in R_{H_2} and $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$ are taken into account for model fitting. In these simulations, we add random offsets to R_{H_2} and $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$ based on their uncertainties and determine the best-fit curve by setting $Z' = 1$ and $\phi_{\text{mol}} = 10$ and finding ϕ_{CNM} that results in the minimum sum of squared residuals. We repeat this process 1000 times and estimate the best-fit ϕ_{CNM} by calculating the median ϕ_{CNM} among the simulated ϕ_{CNM} . The derived ϕ_{CNM} for each region is summarized in Table 1, and the best-fit curves are shown in red in Figures 10 and 11.

In the KMT09 model, χ measures the relative importance of dust shielding and H_2 self-shielding and is predicted to be ~ 1 for a wide range of galactic environments. In this case, a certain

amount of Σ_{HI} is required to shield H_2 against photodissociation, and H_2 forms out of HI once this minimum shielding column density is obtained. The KMT09 model predicts the minimum shielding column density of $\sim 10 M_{\odot} \text{pc}^{-2}$ for solar metallicity, and this is indeed consistent with what we observe in Perseus: Σ_{HI} saturates at $\sim 7\text{--}9 M_{\odot} \text{pc}^{-2}$ for all five regions. The level of the HI saturation changes between the regions, though, from $\sim 7 M_{\odot} \text{pc}^{-2}$ for B5 to $\sim 9 M_{\odot} \text{pc}^{-2}$ for NGC 1333. In comparison with Lee et al. (2012), our Σ_{HI} values are slightly higher due to the correction for high optical depth. This correction brings a closer agreement with the KMT09 model.

The excellent agreement with the KMT09 model is also evident from R_{H_2} versus $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$ in Figure 11. For all five regions, we find that R_{H_2} steeply rises at small $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$, turns over at $R_{\text{H}_2} \sim 1$, and then slowly increases at large $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$. In fact, this common trend on a log-linear scale is entirely driven by the almost constant Σ_{HI} , and therefore it is not surprising to find such a good agreement with the KMT09 model where the Σ_{HI} saturation is predicted. By fitting the KMT09 predictions to the observed R_{H_2} versus $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$ profiles, we constrain $\phi_{\text{CNM}} \sim 5\text{--}9$, which is consistent with what Lee et al. (2012) estimated within uncertainties. In the KMT09 model, ϕ_{CNM} determines the CNM density by

¹³ We note that R_{H_2} is not sensitive to ϕ_{mol} . For example, with our median ϕ_{CNM} value of ~ 7 (Table 1), R_{H_2} at $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2} = 100 M_{\odot} \text{pc}^{-2}$ varies by only a factor of ~ 1.1 for $\phi_{\text{mol}} = 10\text{--}50$.

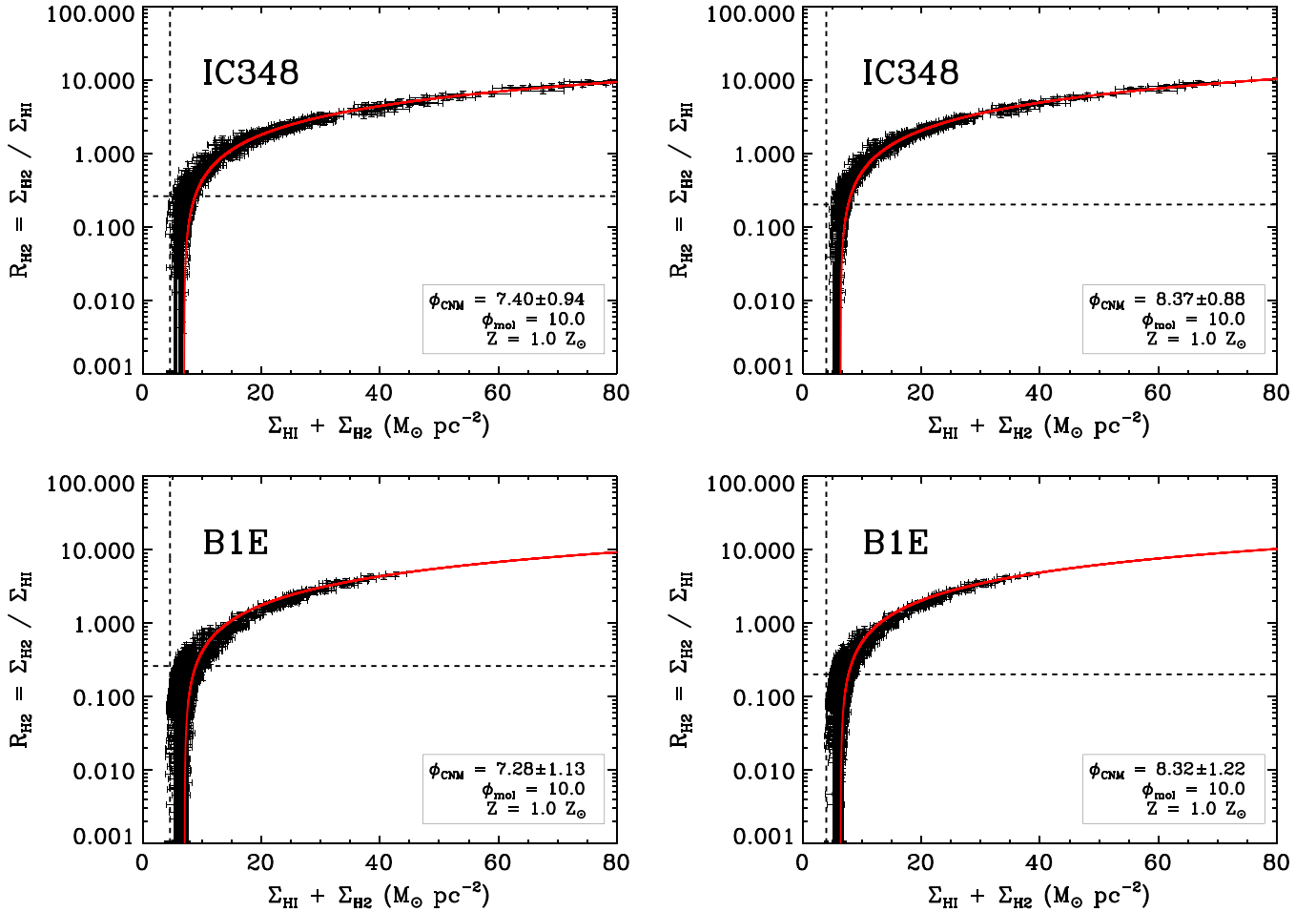


Figure 11. Left: R_{H_2} vs. $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$ for IC 348 and B1E (this study). All finite pixels in the rectangular boxes shown in Figure 9 are used for plotting. The median 3σ values of R_{H_2} and $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$ are indicated as the black dashed lines, while the best-fit curves determined in Section 6.4 are overlaid in red. The best-fit parameters are shown in the bottom right corner of each plot. Right: plots from Lee et al. (2012).

Table 1
Fitting Results for R_{H_2} vs. $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$

Region	Best-fit ϕ_{CNM}
B5	8.75 ± 1.35
IC 348	7.40 ± 0.94
B1E	7.28 ± 1.13
B1	6.93 ± 1.00
NGC 1333	5.28 ± 0.85

Note. The uncertainty in ϕ_{CNM} is estimated from the distribution of the simulated 1000 ϕ_{CNM} values.

$n_{\text{CNM}} = \phi_{\text{CNM}} n_{\text{min}}$, where n_{min} is the minimum CNM density at which the CNM can be in pressure balance with the WNM, and $\phi_{\text{CNM}} = 5\text{--}9$ translates into $T_{\text{CNM}} = 60\text{--}80$ K (Equation (19) of KMT09). This T_{CNM} range is consistent with what we observationally constrained for the HI environment of Perseus via the HI absorption measurements (Paper I): T_{CNM} mostly ranges from ~ 10 to 200 K, and its distribution peaks at ~ 50 K. Finally, we note that ϕ_{CNM} systematically decreases toward the southwest of Perseus, reflecting the observed region-to-region variations in Σ_{HI} . The difference in ϕ_{CNM} , however, is not significant, and this suggests similar χ values for all dark and star-forming regions (Equation (20)). We indeed find

$\chi = \sim 1.1\text{--}1.8$, in agreement with the KMT09 prediction of comparable dust shielding and H_2 self-shielding for H_2 formation.

7. OPTICALLY THICK HI: ALTERNATIVE TO THE “CO-DARK” GAS?

Recently, Fukui et al. (2014, 2015) suggested that the “CO-dark” gas in the Galaxy, referring to the interstellar gas undetectable in both the 21 cm HI and 2.6 mm CO emission (e.g., Bolatto et al. 2013), could be dominated by the optically thick HI. As our HI absorption measurements provide an independent estimate of the optical depth, we can test the validity of their claim against Perseus on sub-pc scales. To do so, we utilize our old and new $N(\text{H}_2)$ images (before and after the correction for high optical depth), as well as the CfA I_{CO} data.

First, we examine how the “CO-dark” gas and the optically thick HI gas are spatially distributed. In order to identify the “CO-dark” gas, we use the 3σ contours of the old $N(\text{H}_2)$ and CfA I_{CO} images, following the definition by Lee et al. (2012). These contours are overlaid on our correction factor image (Figure 6) and show that the relative distribution of H_2 (or simply the gas not probed by the HI emission) and CO changes across the cloud. For example, H_2 and CO trace each other in the southwest, while H_2 is more extended than CO elsewhere.

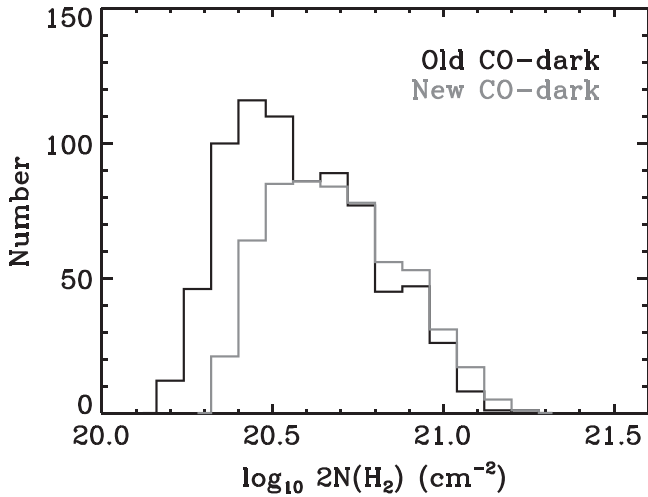


Figure 12. Histograms of the “CO-dark” gas column density. The “CO-dark” gas from the old $N(\text{H}_2)$ image is in black, while that from the new $N(\text{H}_2)$ image is in gray. Both histograms are constructed using the data smoothed to $8''.4$ resolution.

Lee et al. (2012) compared H_2 and CO radial profiles for Perseus and estimated that H_2 is on average ~ 1.4 times more extended than CO, suggesting a substantial amount of the “CO-dark” gas. We then find that the distributions of the “CO-dark” gas (traced by the difference between the H_2 and CO 3σ contours) and the optically thick HI gas (traced by the correction factor) generally disagree with each other. For example, the region around B5, where the discrepancy between the H_2 and CO distributions is greatest, shows relatively low correction factors. On the other hand, the regions with high correction factors at (R.A., decl.) $\sim (3^{\text{h}}33^{\text{m}}, +34^{\circ}30')$ and $\sim (3^{\text{h}}23^{\text{m}}, +29^{\circ})$ do not have a large amount of the “CO-dark” gas.

While our previous evaluation was based on the visual examination of the 3σ contours, here we more rigorously investigate whether or not the optically thick HI gas can explain the “CO-dark” gas by comparing the old and new $N(\text{H}_2)$ images. We first smooth the old and new $N(\text{H}_2)$ images to $8''.4$ resolution, as well as their uncertainties, to match the resolution of the CfA I_{CO} image. We then find the “CO-dark” gas from the smoothed $N(\text{H}_2)$ images by utilizing the 3σ contours of $N(\text{H}_2)$ and I_{CO} . Essentially, a pixel is classified as the “CO-dark” gas when $N(\text{H}_2)$ is above the 3σ level, but I_{CO} is less than the 3σ noise. For the selected pixels, we calculate the “CO-dark” gas column density, $2N(\text{H}_2)$, and present two histograms in Figure 12. The “CO-dark” gas from the old $N(\text{H}_2)$ is in black, while that from the new $N(\text{H}_2)$ is in gray. Figure 12 shows that the two histograms are comparable regarding their minimum, maximum, and median values (different by less than a factor of 2), although the gray histogram has a smaller number of pixels due to the larger uncertainty of the new $N(\text{H}_2)$ image. Given that the optically thick HI gas was already taken into consideration in the derivation of the new $N(\text{H}_2)$ image, the comparable black and gray histograms suggest that the increased column density due to the optically thick HI gas is small (up to $\sim 2 \times 10^{20} \text{ cm}^{-2}$), and the “CO-dark” gas still exists in Perseus. In terms of mass, the additional contribution from the optically thick HI only accounts for $\sim 20\%$ of the observed “CO-dark” gas.

While our results are in contrast with Fukui et al. (2014, 2015), who found that the optically thick HI adds the

column density of $\sim 10^{20}\text{--}10^{22} \text{ cm}^{-2}$ and possibly explains the “CO-dark” gas in the Galaxy, there are multiple factors that could affect the comparison, e.g., spatial coverage ($\sim 500 \text{ deg}^2$ for Perseus versus whole Galactic sky at $|b| > 15^\circ$) and method for deriving $N(\text{H}_2)$ (IRIS/2MASS versus *Planck*). In particular, we note that this study and Fukui et al. (2014, 2015) probe very different scales: our results are based on pencil-beam HI absorption measurements on $3''.5$ scales, while Fukui et al. (2014, 2015) estimated the correction factor on $33''$ scales. If the CNM is highly structured with a low filling factor, this could affect the estimate of the correction factor in both studies. In the future, it will be important to compare the results from Fukui et al. (2014, 2015) with a large sample of HI absorption measurements and numerical simulations (e.g., Audit & Hennebelle 2005; Kim et al. 2014) to investigate how the derivation of the correction factor depends on different methodologies and CNM properties.

8. SUMMARY

In this paper, we investigate the impact of high optical depth on the HI column density distribution across the Perseus molecular cloud. We use Arecibo HI emission and absorption measurements obtained toward 26 background sources (Paper I) in order to derive the properties of CNM and WNM components along each line of sight via the Gaussian decomposition approach (HT03a). The derived properties are then used to estimate the correction factor for high optical depth, and the correction is applied to the HI column density image computed in the optically thin approximation.

To revisit the HI saturation in Perseus observed by Lee et al. (2012), we rederive the H_2 column density image by adopting the same methodology as Lee et al. (2012), but using the corrected HI column density image. The final HI and H_2 column density images at $\sim 0.4 \text{ pc}$ resolution are then compared with the KMT09 predictions. Finally, we investigate whether the observed “CO-dark” gas in Perseus is dominated by the optically thick HI gas.

We summarize our main results as follows.

1. We estimate the correction factor for high optical depth (f), which is defined as the ratio of the true total HI column density (N_{tot}) to the HI column density derived in the optically thin approximation (N_{exp}), and express it as a function of N_{exp} : $f = \log_{10}(N_{\text{exp}}/10^{20})(0.32 \pm 0.06) + (0.81 \pm 0.05)$. We use two different methods, Gaussian decomposition and isothermal approximation methods, and find that they are consistent within uncertainties. This is likely due to the relatively low optical depth and insignificant contribution from the diffuse radio continuum emission for Perseus.
2. We estimate that the correction factor in/around Perseus is small (up to ~ 1.2), and the total HI mass increases by only $\sim 10\%$ from $\sim 2.3 \times 10^4 M_\odot$ to $\sim 2.5 \times 10^4 M_\odot$ due to the inclusion of the optically thick HI gas.
3. The H_2 column density image rederived using the corrected HI column density image is comparable to the original one by Lee et al. (2012), confirming the minor correction for high optical depth.
4. For individual dark and star-forming regions in Perseus (B5, B1E, B1, IC 348, and NGC 1333), the HI surface density is relatively uniform with $\sim 7\text{--}9 M_\odot \text{ pc}^{-2}$. This is

slightly higher than what Lee et al. (2012) found due to the correction for high optical depth. The correction brings a closer agreement with the KMT09 model, where the minimum HI surface density of $\sim 10 M_\odot \text{pc}^{-2}$ is predicted for shielding H_2 against photodissociation in the solar metallicity environment.

5. Driven by the uniform $\Sigma_{\text{HI}} \sim 7\text{--}9 M_\odot \text{pc}^{-2}$, R_{H_2} versus $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$ on a log-linear scale shows remarkably consistent results for all dark and star-forming regions: R_{H_2} sharply rises at small $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$ and then gradually increases toward large $\Sigma_{\text{HI}} + \Sigma_{\text{H}_2}$ with the transition at $R_{\text{H}_2} \sim 1$.
6. The mass increase due to the optically thick HI only accounts for $\sim 20\%$ of the observed “CO-dark” gas in Perseus. The spatial distributions of the “CO-dark” gas and the optically thick HI gas do not generally coincide with each other, and the “CO-dark” gas still exists even after the optically thick HI is considered in the derivation of H_2 .

Our study is one of the first attempts to examine the properties of the cold and warm HI in molecular cloud environments and their relation with the HI-to- H_2 transition. While HI envelopes surrounding molecular clouds have been frequently found (e.g., Knapp 1974; Wannier et al. 1983; Elmegreen & Elmegreen 1987; Andersson et al. 1991; Rogers et al. 1995; Williams & Maddalena 1996; Imara & Blitz 2011; Lee et al. 2012, 2014; Motte et al. 2014), HI has traditionally been considered less important for the formation and evolution of molecular clouds. The excellent agreement between our observations of Perseus on sub-pc scales and the KMT09 model, on the other hand, suggests the significance of HI as one of the key ingredients for the HI-to- H_2 transition and consequently for star formation.

While our data agree with the KMT09 model, further theoretical developments are required. For example, the KMT09 model considers only the CNM as a source of shielding against H_2 photodissociation. Our HI emission and absorption measurements, however, show that the CNM and the WNM have comparable column density contributions for Perseus (Paper I). Clearly, the WNM needs to be taken into consideration in the models of H_2 formation (e.g., Bialy et al. 2015). In addition, the KMT09 model ignores the effect of internal radiation field. This is valid for Perseus, as there are no early-type stars producing a significant amount of dissociating radiation (Lee et al. 2012). However, the role of internal radiation field would be important for massive molecular clouds containing a large number of OB stars. The discrepancy with the KMT09 model recently found for the W43 molecular cloud complex in the Galactic plane (Motte et al. 2014) and the low-metallicity SMC (Welty et al. 2012) suggests that additional physical ingredients (e.g., better description of H_2 formation and photodissociation in the low-metallicity ISM and strong shear/turbulence at galactic bars) would be necessary for extreme environments. Finally, vertical thermal and dynamical equilibrium in a galactic disk is another important aspect, as recently explored by Kim et al. (2014). Future comparisons between theoretical models and HI emission/absorption observations of molecular clouds in a wide range of ISM environments will be important for a deep understanding of HI properties and

their role in the formation and evolution of molecular clouds.

We thank the anonymous referee for suggestions that improved this work. We also thank B.-G. Andersson, Shmuel Bialy, John Dickey, Yasuo Fukui, Chang-Goo Kim, Mark Krumholz, Bill Langer, Vianney Lebouteiller, Franck Le Petit, Harvey Liszt, Suzanne Madden, Bill Reach, Evelyne Roueff, and Amiel Sternberg for stimulating discussion, and telescope operators at the Arecibo Observatory for their help in conducting our HI observations. M.-Y.L. acknowledges support from the DIM ACAV of the Region Ile de France, and the SYMPATICO grant (ANR-11-BS56-0023) of the French Agence Nationale de la Recherche. S.S. acknowledges support from the NSF Early Career Development (CAREER) Award AST-1056780. We also acknowledge the NSF REU grant AST-1004881, which funded summer research of J.M. For this work, we have made use of the KARMA visualization software (Gooch 1996) and NASA’s Astrophysics Data System (ADS).

APPENDIX A COMPARISON BETWEEN METHODS 1 AND 2

In order to investigate why the Gaussian decomposition and isothermal methods result in comparable correction factors despite their apparent differences (Sections 4.2 and 4.3), we start with Equation (14) and compare it with Equation (5):

$$N_{\text{tot}} (\text{cm}^{-2}) = C_0 \int f_{\text{chan}}(v) T_{\text{exp}}(v) dv \quad (21)$$

where $C_0 = 1.823 \times 10^{18} \text{cm}^{-2}/(\text{K km s}^{-1})$. Combined with Equation (12), Equation (21) then becomes

$$\begin{aligned} N_{\text{tot}} (\text{cm}^{-2}) &= C_0 \int \frac{T_s(v)}{T_s(v) - T_{\text{sky}}} \frac{\tau(v)}{1 - e^{-\tau(v)}} T_{\text{exp}}(v) dv \\ &\sim C_0 \int \frac{\tau(v)}{1 - e^{-\tau(v)}} T_{\text{exp}}(v) dv. \end{aligned} \quad (22)$$

The approximation in Equation (22) can be made due to the small $T_s(v)/[T_s(v) - T_{\text{sky}}]$ of $\sim 1.0\text{--}1.2$ (Section 4.3).

While an optical depth spectrum is primarily determined by the CNM, an “expected” emission spectrum has contributions from both the CNM and the WNM. Hence, $T_{\text{exp}}(v)$ can be expressed as (Equations (2)–(4) of Paper I)

$$T_{\text{exp}}(v) = T_{\text{B,CNM}}(v) + T_{\text{B,WNM}}(v) \quad (23)$$

with

$$\begin{aligned} T_{\text{B,CNM}}(v) &= \sum_0^{N-1} T_{s,n} (1 - e^{-\tau_n(v)}) e^{-\sum_0^{M-1} \tau_m(v)} \\ T_{\text{B,WNM}}(v) &= \sum_0^{K-1} [F_k + (1 - F_k) e^{-\tau(v)}] T_{0,k} e^{-[(v-v_{0,k})/\delta v_k]^2} \end{aligned}$$

where $T_{\text{B,CNM}}(v)$ is the emission from N CNM components, $T_{\text{B,WNM}}(v)$ is the emission from K WNM components, the components with subscript n refer to the CNM, the components with subscript k refer to the WNM, τ_m is the optical depth of m CNM component that lies in front of n cloud, F_k is the fraction

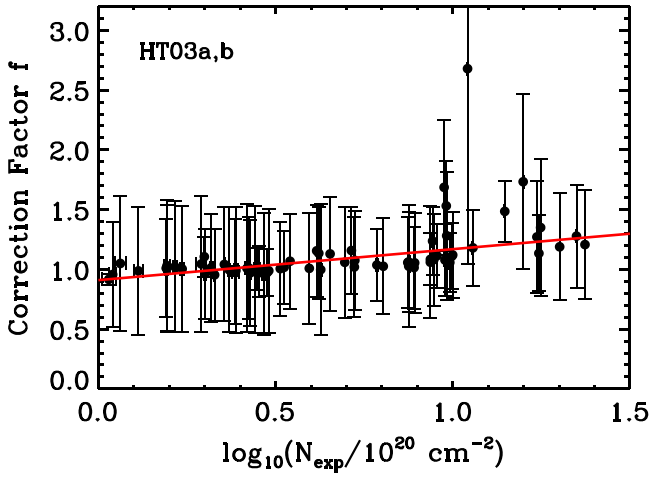


Figure 13. $f = N_{\text{tot}}/N_{\text{exp}}$ as a function of $\log_{10}(N_{\text{exp}}/10^{20})$ for 68 sources in HT03a, HT03b. Eleven lines of sight probing the Galactic plane ($|b| < 10^\circ$) are excluded from the analysis to make a comparison with our study of Perseus. Both the true and optically thin HI column densities (N_{tot} and N_{exp}) are calculated using the published full line-of-sight information, and the linear fit to all 68 data points is overlaid in red: $f = \log_{10}(N_{\text{exp}}/10^{20})(0.26 \pm 0.08) + (0.91 \pm 0.04)$.

of k WNM component that lies in front of all CNM clouds, T_0 is the peak brightness temperature, v_0 is the central velocity, and finally δv is the $1/e$ width of the component. Equation (22) then becomes

$$N_{\text{tot}} (\text{cm}^{-2}) \sim C_0 \int \frac{\tau(v)}{1 - e^{-\tau(v)}} \times \left\{ \sum_0^{N-1} T_{s,n} (1 - e^{-\tau_n(v)}) e^{-\sum_0^{M-1} \tau_m(v)} + \sum_0^{K-1} [F_k + (1 - F_k) e^{-\tau(v)}] T_{0,k} e^{-[(v-v_{0,k})/\delta v_k]^2} \right\} dv. \quad (24)$$

We examine Equation (24) under two extreme circumstances: (I) $F_k = 0$ and $\tau \ll 1$ and (II) $F_k = 1$ and $\tau \ll 1$. In both cases, the first part of Equation (24) becomes

$$\begin{aligned} A &= C_0 \int \frac{\tau(v)}{1 - e^{-\tau(v)}} \left\{ \sum_0^{N-1} T_{s,n} (1 - e^{-\tau_n(v)}) e^{-\sum_0^{M-1} \tau_m(v)} \right\} dv \\ &\sim C_0 \int \frac{\tau(v)}{1 - e^{-\tau(v)}} \left\{ \sum_0^{N-1} T_{s,n} (1 - e^{-\tau_n(v)}) \right\} dv \\ &\sim C_0 \int \frac{\tau(v)}{1 - (1 - \tau(v))} \left\{ \sum_0^{N-1} T_{s,n} (1 - (1 - \tau_n(v))) \right\} dv \\ &= C_0 \int \sum_0^{N-1} T_{s,n} \tau_n(v) dv \\ &= C_0 \int \sum_0^{N-1} T_{s,n} \tau_{0,n} e^{-[(v-v_{0,n})/\delta v_n]^2} dv. \end{aligned} \quad (25)$$

On the other hand, the second part of Equation (24) becomes (I) $F_k = 0$ (all WNM components are behind the CNM) and

$\tau \ll 1$:

$$\begin{aligned} B &= C_0 \int \frac{\tau(v)}{1 - e^{-\tau(v)}} \left\{ \sum_0^{K-1} [F_k + (1 - F_k) e^{-\tau(v)}] \right. \\ &\quad \times \left. T_{0,k} e^{-[(v-v_{0,k})/\delta v_k]^2} \right\} dv \\ &= C_0 \int \frac{\tau(v) e^{-\tau(v)}}{1 - e^{-\tau(v)}} \left\{ \sum_0^{K-1} T_{0,k} e^{-[(v-v_{0,k})/\delta v_k]^2} \right\} dv \\ &\sim C_0 \int \frac{\tau(v)}{1 - (1 - \tau(v))} \left\{ \sum_0^{K-1} T_{0,k} e^{-[(v-v_{0,k})/\delta v_k]^2} \right\} dv \\ &= C_0 \int \sum_0^{K-1} T_{0,k} e^{-[(v-v_{0,k})/\delta v_k]^2} dv. \end{aligned} \quad (26)$$

(II) $F_k = 1$ (all WNM components are in front of the CNM) and $\tau \ll 1$:

$$\begin{aligned} B &= C_0 \int \frac{\tau(v)}{1 - e^{-\tau(v)}} \left\{ \sum_0^{K-1} [F_k + (1 - F_k) e^{-\tau(v)}] \right. \\ &\quad \times \left. T_{0,k} e^{-[(v-v_{0,k})/\delta v_k]^2} \right\} dv \\ &= C_0 \int \frac{\tau(v)}{1 - e^{-\tau(v)}} \left\{ \sum_0^{K-1} T_{0,k} e^{-[(v-v_{0,k})/\delta v_k]^2} \right\} dv \\ &\sim C_0 \int \frac{\tau(v)}{1 - (1 - \tau(v))} \left\{ \sum_0^{K-1} T_{0,k} e^{-[(v-v_{0,k})/\delta v_k]^2} \right\} dv \\ &= C_0 \int \sum_0^{K-1} T_{0,k} e^{-[(v-v_{0,k})/\delta v_k]^2} dv. \end{aligned} \quad (27)$$

As a result, we have

$$\begin{aligned} N_{\text{tot}} (\text{cm}^{-2}) &= A + B \\ &\sim C_0 \int \left\{ \sum_0^{N-1} T_{s,n} \tau_{0,n} e^{-[(v-v_{0,n})/\delta v_n]^2} + \sum_0^{K-1} T_{0,k} e^{-[(v-v_{0,k})/\delta v_k]^2} \right\} dv, \end{aligned} \quad (28)$$

which is essentially Equation (5).

These examples of the two extreme cases suggest that the Gaussian decomposition and isothermal methods would result in comparable correction factors ($f = N_{\text{tot}}/N_{\text{exp}}$) regardless of F_k when $\tau \ll 1$ and $T_{\text{sky}} \ll T_s$.

APPENDIX B CORRECTION FACTOR IN HEILES & TROLAND (2003A, 2003B)

To examine the validity of the linear relation we assume for the correction factor and the optically thin HI column density (Sections 4.2 and 4.3), we produce the same plot as Figure 3 (left) for HT03a, HT03b. Among the total 79 sources in HT03a, HT03b, 11 sources with $|b| < 10^\circ$ are excluded for a fair comparison with our study. The published “expected” emission spectra and total HI column densities are used to estimate f and $\log_{10}(N_{\text{exp}}/10^{20})$ in the same way as we do for our 26 sources, and the results are presented in Figure 13.

We find that Figure 13 is very similar to Figures 3 (left) and 5 (left), suggesting that our assumption of the linear relation between f and $\log_{10}(N_{\text{exp}}/10^{20})$ for Perseus is reasonable. Interestingly, however, a few sources with $\log_{10}(N_{\text{exp}}/10^{20}) \gtrsim 1$ show some deviation from the linear relation, which requires a further examination. See Section 4.3 for a more detailed discussion.

REFERENCES

- Andersson, B.-G., Wannier, P. G., & Morris, M. 1991, *ApJ*, **366**, 464
- Audit, E., & Hennebelle, P. 2005, *A&A*, **433**, 1
- Bialy, S., Sternberg, A., Lee, M.-Y., Le Petit, F., & Roueff, E. 2015, *ApJ*, in press (arXiv:1505.06200)
- Bally, J., Walawender, J., Johnstone, D., Kirk, H., & Goodman, A. 2008, in *Handbook of Star Forming Regions*, Vol. 1, ed. B. Reipurth (San Francisco, CA: ASP), 308
- Barraault, L., Juncas, G., Lockman, F. J., & Martin, P. G. 2010, *MNRAS*, **407**, 2645
- Bigiel, F., Leroy, A., Walter, F., et al. 2008, *AJ*, **136**, 2846
- Blitz, L., & Rosolowsky, E. 2006, *ApJ*, **650**, 933
- Bolatto, A. D., Leroy, A. K., Jameson, K., et al. 2011, *ApJ*, **741**, 12
- Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, *ARA&A*, **51**, 207
- Braun, R. 2012, *ApJ*, **749**, 87
- Braun, R., Thilker, D. A., Walterbos, R. A. M., & Corbelli, E. 2009, *ApJ*, **695**, 937
- Braun, R., & Walterbos, R. A. M. 1992, *ApJ*, **386**, 120
- Browning, M. K., Tumlinson, J., & Shull, J. M. 2003, *ApJ*, **582**, 810
- Černis, K. 1990, *Ap&SS*, **166**, 315
- Chengalur, J. N., Kanekar, N., & Roy, N. 2013, *MNRAS*, **432**, 3074
- Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, *AJ*, **115**, 1693
- Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, *ApJ*, **547**, 792
- Dickey, J. M., & Benson, J. M. 1982, *AJ*, **87**, 278
- Dickey, J. M., McClure-Griffiths, N. M., Gaensler, B. M., & Green, A. J. 2003, *ApJ*, **585**, 801
- Dickey, J. M., Mebold, U., Stanimirović, S., & Staveley-Smith, L. 2000, *ApJ*, **536**, 756
- Douglas, K. A., & Taylor, A. R. 2007, *ApJ*, **659**, 426
- Draine, B. T., & Bertoldi, F. 1996, *ApJ*, **468**, 269
- Elmegreen, B. G. 1993, *ApJ*, **411**, 170
- Elmegreen, B. G., & Elmegreen, D. M. 1987, *ApJ*, **320**, 182
- Fukui, Y., Okamoto, R., Kaji, R., et al. 2014, *ApJ*, **796**, 59
- Fukui, Y., Torii, K., Onishi, T., et al. 2015, *ApJ*, **798**, 6
- Genzel, R., Tacconi, L. J., Kurk, J., et al. 2013, *ApJ*, **773**, 68
- Gillmon, K., Shull, J. M., Tumlinson, J., & Danforth, C. 2006, *ApJ*, **636**, 891
- Glassgold, A. E., & Langer, W. D. 1974, *ApJ*, **193**, 73
- Glover, S. C. O., Federrath, C., Mac Low, M.-M., & Klessen, R. S. 2010, *MNRAS*, **404**, 2
- Goldsmith, P. F., Li, D., & Krčo, M. 2007, *ApJ*, **654**, 273
- González Hernández, J. I., Iglesias-Groth, S., Rebolo, R., et al. 2009, *ApJ*, **706**, 866
- Goodman, A. A., Pineda, J. E., & Schnee, S. L. 2009, *ApJ*, **692**, 91
- Gould, R. J., & Salpeter, E. E. 1963, *ApJ*, **138**, 393
- Haslam, C. G. T., Salter, C. J., Stoffel, H., & Wilson, W. E. 1982, *A&AS*, **47**, 1
- Heiles, C., & Troland, T. H. 2003a, *ApJS*, **145**, 329
- Heiles, C., & Troland, T. H. 2003b, *ApJ*, **586**, 1067
- Henderson, A. P., Jackson, P. D., & Kerr, F. J. 1982, *ApJ*, **263**, 116
- Herbig, G. H., & Jones, B. F. 1983, *AJ*, **88**, 1040
- Imara, N., & Blitz, L. 2011, *ApJ*, **732**, 78
- Kennicutt, R. C., Jr. 1989, *ApJ*, **344**, 685
- Kennicutt, R. C., & Evans, N. J. 2012, *ARA&A*, **50**, 531
- Kim, C.-G., Ostriker, E. C., & Kim, W.-T. 2013, *ApJ*, **776**, 1
- Kim, C.-G., Ostriker, E. C., & Kim, W.-T. 2014, *ApJ*, **786**, 64
- Knapp, G. R. 1974, *AJ*, **79**, 527
- Koyama, H., & Inutsuka, S.-I. 2002, *ApJL*, **564**, L97
- Krumholz, M. R., McKee, C. F., & Tumlinson, J. 2009, *ApJ*, **693**, 216
- Lada, C. J., Lombardi, M., & Alves, J. F. 2010, *ApJ*, **724**, 687
- Lee, M.-Y., Stanimirović, S., Douglas, K. A., et al. 2012, *ApJ*, **748**, 75
- Lee, M.-Y., Stanimirović, S., Wolfire, M. G., et al. 2014, *ApJ*, **784**, 80
- Leroy, A. K., Walter, F., Brinks, E., et al. 2008, *AJ*, **136**, 2782
- Liszt, H. 2014, *ApJ*, **783**, 17
- Liszt, H. S. 2007, *A&A*, **461**, 205
- Liszt, H. S., Pety, J., & Lucas, R. 2010, *A&A*, **518**, A45
- 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., Green, A. J., Dickey, J. M., et al. 2001, *ApJ*, **551**, 394
- McKee, C. F., & Ostriker, J. P. 1977, *ApJ*, **218**, 148
- Meyerdierks, H., & Heithausen, A. 1996, *A&A*, **313**, 929
- Miville-Deschênes, M.-A., & Lagache, G. 2005, *ApJS*, **157**, 302
- Motte, F., Nguyễn Luong, Q., Schneider, N., et al. 2014, *A&A*, **571**, A32
- Offner, S. S. R., Bisbas, T. G., Viti, S., & Bell, T. A. 2013, *ApJ*, **770**, 49
- Peek, J. E. G., Heiles, C., Douglas, K. A., et al. 2011, *ApJS*, **194**, 20
- Pineda, J. L., Langer, W. D., Velusamy, T., & Goldsmith, P. F. 2013, *A&A*, **554**, A103
- Rachford, B. L., Snow, T. P., Tumlinson, J., et al. 2002, *ApJ*, **577**, 221
- Reach, W. T., Koo, B.-C., & Heiles, C. 1994, *ApJ*, **429**, 672
- Ridge, N. A., di Francesco, J., Kirk, H., et al. 2006, *AJ*, **131**, 2921
- Rogers, C., Heyer, M. H., & Dewdney, P. E. 1995, *ApJ*, **442**, 694
- Sancisi, R., Goss, W. M., Anderson, C., Johansson, L. E. B., & Winnberg, A. 1974, *A&A*, **35**, 445
- Savage, B. D., Bohlin, R. C., Drake, J. F., & Budich, W. 1977, *ApJ*, **216**, 291
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, **500**, 525
- Schmidt, M. 1959, *ApJ*, **129**, 243
- Schruba, A., Leroy, A. K., Walter, F., et al. 2011, *AJ*, **142**, 37
- Spaans, M., & Neufeld, D. A. 1997, *ApJ*, **484**, 785
- Spitzer, L., Jr. 1948, *ApJ*, **107**, 6
- Stanimirović, S., & Heiles, C. 2005, *ApJ*, **631**, 371
- Stanimirović, S., Murray, C. E., Lee, M.-Y., Heiles, C., & Miller, J. 2014, *ApJ*, **793**, 132
- Stanimirović, S., Putman, M., Heiles, C., et al. 2006, *ApJ*, **653**, 1210
- Stanimirović, S., Staveley-Smith, L., Dickey, J. M., Sault, R. J., & Snowden, S. L. 1999, *MNRAS*, **302**, 417
- Sternberg, A. 1988, *ApJ*, **332**, 400
- Sternberg, A., Le Petit, F., Roueff, E., & le Bourlot, J. 2014, *ApJ*, **790**, 10
- Tacconi, L. J., Genzel, R., Neri, R., et al. 2010, *Natur*, **463**, 781
- van Dishoeck, E. F., & Black, J. H. 1986, *ApJS*, **62**, 109
- Wannier, P. G., Lichten, S. M., & Morris, M. 1983, *ApJ*, **268**, 727
- Welty, D. E., Xue, R., & Wong, T. 2012, *ApJ*, **745**, 173
- Williams, J. P., & Maddalena, R. J. 1996, *ApJ*, **464**, 247
- Wilson, C. D., Warren, B. E., Israel, F. P., et al. 2009, *ApJ*, **693**, 1736
- Wolfire, M. G., McKee, C. F., Hollenbach, D., & Tielens, A. G. G. M. 2003, *ApJ*, **587**, 278
- Wong, T., & Blitz, L. 2002, *ApJ*, **569**, 157
- Wong, T., Hughes, A., Fukui, Y., et al. 2009, *ApJ*, **696**, 370
- Wong, T., Xue, R., Bolatto, A. D., et al. 2013, *ApJL*, **777**, L4