



HAL
open science

Mapping a Lower Limit on the Mass Fraction of the Cold Neutral Medium Using Fourier-transformed H i 21 cm Emission Line Spectra: Application to the DRAO Deep Field from DHIGLS and the HI4PI Survey

Antoine Marchal, Peter G Martin, Marc-Antoine Miville-Deschênes, Naomi M Mcclure-Griffiths, Callum Lynn, Andrea Bracco, Luka Vujeva

► To cite this version:

Antoine Marchal, Peter G Martin, Marc-Antoine Miville-Deschênes, Naomi M Mcclure-Griffiths, Callum Lynn, et al.. Mapping a Lower Limit on the Mass Fraction of the Cold Neutral Medium Using Fourier-transformed H i 21 cm Emission Line Spectra: Application to the DRAO Deep Field from DHIGLS and the HI4PI Survey. *The Astrophysical Journal*, 2024, 961 (2), pp.161. 10.3847/1538-4357/ad0f21 . hal-04780848

HAL Id: hal-04780848

<https://hal.science/hal-04780848v1>

Submitted on 13 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Mapping a Lower Limit on the Mass Fraction of the Cold Neutral Medium Using Fourier-transformed HI 21 cm Emission Line Spectra: Application to the DRAO Deep Field from DHIGLS and the HI4PI Survey

Antoine Marchal^{1,2} , Peter G. Martin¹ , Marc-Antoine Miville-Deschênes³ , Naomi M. McClure-Griffiths² , Callum Lynn² ,
Andrea Bracco⁴ , and Luka Vujeva⁵

¹ Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada; antoine.marchal@anu.edu.au

² Research School of Astronomy & Astrophysics, Australian National University, Canberra ACT 2610 Australia

³ AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, F-91191 Gif-sur-Yvette, France

⁴ Laboratoire de Physique de l'École Normale Supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris, F-75005 Paris, France

⁵ Niels Bohr International Academy, Niels Bohr Institute, Blegdamsvej 17, DK-2100 Copenhagen, Denmark

Received 2023 February 13; revised 2023 November 20; accepted 2023 November 21; published 2024 January 23

Abstract

We develop a new method for spatially mapping a lower limit on the mass fraction of the cold neutral medium by analyzing the amplitude structure of $\hat{T}_b(k_\nu)$, the Fourier transform of $T_b(\nu)$, the spectrum of the brightness temperature of the HI 21 cm line emission with respect to the radial velocity ν . This advances a broader effort exploiting 21 cm emission line data alone (without absorption line data, τ) to extract integrated properties of the multiphase structure of the HI gas and to map each phase separately. Using toy models, we illustrate the origin of interference patterns seen in $\hat{T}_b(k_\nu)$. Building on this, a lower limit on the cold gas mass fraction is obtained from the amplitude of \hat{T}_b at high k_ν . Tested on a numerical simulation of thermally bi-stable turbulence, the lower limit from this method has a strong linear correlation with the “true” cold gas mass fraction from the simulation for a relatively low cold gas mass fraction. At a higher mass fraction, our lower limit is lower than the “true” value, because of a combination of interference and opacity effects. Comparison with absorption surveys shows a similar behavior, with a departure from linear correlation at $N_{\text{HI}} \gtrsim 3\text{--}5 \times 10^{20} \text{ cm}^{-2}$. Application to the DRAO Deep Field from DHIGLS reveals a complex network of cold filaments in the Spider, an important structural property of the thermal condensation of the HI gas. Application to the HI4PI survey in the velocity range $-90 < \nu < 90 \text{ km s}^{-1}$ produces a full sky map of a lower limit on the mass fraction of the cold neutral medium at 16.2 resolution. Our new method has the ability to extract a lower limit on the cold gas mass fraction for massive amounts of emission line data alone with low computing time and memory, pointing the way to new approaches suitable for the new generation of radio interferometers.

Unified Astronomy Thesaurus concepts: [Interstellar medium \(847\)](#); [Cold neutral medium \(266\)](#); [Neutral hydrogen clouds \(1099\)](#); [Fast Fourier transform \(1958\)](#)

1. Introduction

In the interstellar medium (ISM), thermal instability is thought to be the main process that leads to thermal condensation of the warm neutral phase, producing a thermally unstable lukewarm medium and a dense cold neutral medium (CNM; Field 1965; Field et al. 1969; Wolfire et al. 1995; Hennebelle & Péroul 1999, 2000; Wolfire et al. 2003; Audit & Hennebelle 2005). Understanding the properties of the multiphase neutral ISM is key to understanding this initial step leading to the atomic-to-molecular (HI-to-H₂) transition and the formation of molecular clouds (Lee et al. 2012; Stanimirović et al. 2014; Burkhart et al. 2015; Lee et al. 2015; Pingel et al. 2018; Bialy & Sternberg 2019; Wang et al. 2020b; Kalberla et al. 2020; Syed et al. 2020; Marchal & Miville-Deschênes 2021; Rybarczyk et al. 2022; Seifried et al. 2022).

In recent decades, huge efforts have been made to map the 21 cm emission of Galactic HI (e.g., Taylor et al. 2003; Kalberla et al. 2005; Stil et al. 2006; McClure-Griffiths et al. 2009; Peek et al. 2011, 2018; Martin et al. 2015; Beuther et al. 2016; HI4PI Collaboration et al. 2016; Winkel et al. 2016; Blagrave et al.

2017; Wang et al. 2020a), as well as HI in nearby galaxies (e.g., Kim et al. 1998; Walter et al. 2008; Koch et al. 2018; Pingel et al. 2022). However, the spatial information about the multiphase and multiscale nature of the neutral ISM contained in these large hyperspectral position–position–velocity (PPV) cubes has remained elusive due to the difficulty in separating the emission from the different phases along each line of sight (see Marchal et al. 2019, and references within).

In recent decades, methods aimed at phase separation of 21 cm emission-only data have been developed to recover information about the physical properties of each phase. Among them, Gaussian decomposition algorithms have focused on modeling the contribution of each phase to HI PPV cubes (Haud 2000; Nidever et al. 2008; Lindner et al. 2015; Kalberla & Haud 2018; Marchal et al. 2019; Riener et al. 2019). Through the use of spatial regularization of the Gaussian model, the code ROHSA (Marchal et al. 2019) was developed to model each phase with spatial coherence across velocities and allowed the thermal and turbulent properties of the WNM to be studied with emission-line data only (Marchal & Miville-Deschênes 2021). However, these methods are often computationally expensive and difficult to apply to large data sets.

Furthering this ongoing effort, Murray et al. (2020) have recently introduced a novel approach based on a 1D convolutional neural network (CNN) trained on a library of synthetic observations of emission-line and absorption spectra



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

of three-dimensional hydrodynamical simulations of the multi-phase ISM (Kim et al. 2013, 2014). In contrast to Gaussian decomposition algorithms, Murray et al. (2020) focused on extracting the total cold gas mass fraction, and the correction to the optically thin estimate of HI column density because of optical depth along each spectrum, losing information on the velocity field. In this paper, although focusing on a similar goal (i.e., extracting the total cold gas mass fraction), we develop a new method by analyzing the amplitude structure of $\hat{T}(k_v)$, the Fourier transform (hereafter FT) of $T_b(v)$, the spectrum of the brightness temperature of HI 21 cm line emission with respect to the radial velocity v .

Previous studies performing spectral modeling in the k_v domain, such as the Velocity Coordinate Spectrum (VCS; Lazarian & Pogosyan 2000, 2006, 2008; Lazarian 2009), have shown that the power spectrum of spatially averaged 21 cm data cubes contains information about the multiphase structure of the gas (Chepurnov & Lazarian 2006; Chepurnov et al. 2010, 2015), but the VCS loses spatial information on the plane of the sky and precludes mapping out integrated properties of each phase. Our new method using Fourier-transformed spectra exploits valuable information on the amplitude in the k_v domain along individual lines of sight and we obtain a spatial map of a lower limit on the contribution of the cold phase to the total line emission (i.e., the gas cold mass fraction).⁶

The paper is organized as follows. In Section 2, we present the interference patterns seen in the FT of emission data from the 21 cm Spectral Line Observations of Neutral Gas with the VLA (21-SPONGE) survey (Murray et al. 2015, 2018) and use toy models to understand their origins. Building on this, we develop a methodology to evaluate a lower limit on the mass fraction of gas below an arbitrary kinetic temperature. Evaluation using a realistic numerical simulation of the neutral ISM is presented in Section 3. In Section 4.1, we apply our new method to emission data alone from 21-SPONGE and compare our lower limits to the cold gas mass fraction obtained by Murray et al. (2018) by joint analysis of emission and absorption data. Application to emission-line data for the DRAO Deep Field (hereafter DF) from the DHIGLS survey and the HI4PI survey in Sections 6 and 7, respectively, shows that the information contained at high k_v in PPK cubes⁷ of Fourier-transformed emission-line spectra is intimately related to the emission in small spatial scale features often seen in individual channel maps of PPV cubes. A discussion and a summary are provided in Sections 9 and 10, respectively.

2. Fourier Transform of 21 cm Data

2.1. Fourier Transform and $k_v = 0$ Normalization

The discrete FT of brightness temperature data $T_b(v)$ as a function of the radial velocity v is

$$\hat{T}_b(k_v) = \sum_v T_b(v) \exp(-2\pi j v k_v) dv, \quad (1)$$

where j is the imaginary unit. At $k_v = 0$, $\hat{T}_b(0) = \sum_v T_b(v) dv$, which is the estimator of the total HI column density in the optically thin limit, N_{HI}^* , divided by $C = 1.82243 \times$

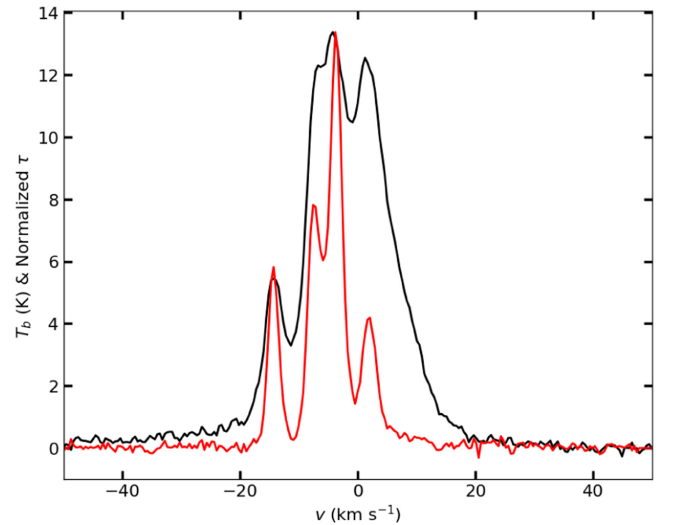


Figure 1. Brightness temperature T_b of the absorption source J2232 from the 21-SPONGE survey (black). Its corresponding optical depth τ , normalized at the peak of T_b , is shown in red.

$10^{18} \text{ cm}^{-2} (\text{K km s}^{-1})^{-1}$. We use the information contained in the first mode to define the $k_v = 0$ normalization of the FT data

$$\hat{T}'_b(k) = \hat{T}_b(k) / \hat{T}_b(0), \quad (2)$$

whose amplitude has values in the range $[0, 1]$.

2.2. A Case from the 21-SPONGE Survey

To illustrate the transformation of the data from the velocity domain to the k_v domain, we used observations from 21-SPONGE.⁸ In particular, we used observations of the 21 cm absorption-line optical depth $\tau(v)$ against the radio continuum source J2232 located at $(\alpha, \delta) = (22^{\text{h}}32^{\text{m}}36^{\text{s}}.4, 11^{\circ}43'50''.9)$ or $(l, b) = (77^{\circ}.438, -38^{\circ}.582)$ obtained with the Karl G. Jansky Very Large Array (VLA). The matching emission-line spectrum $T_b(v)$ was interpolated from off-source observations with the Arecibo telescope with an angular resolution of approximately $3'.5$ and a velocity resolution of 0.16 km s^{-1} (see details about the matching procedure in Section 2.4 of Murray et al. 2015). These complementary spectra are shown in Figure 1, with the $\tau(v)$ spectrum scaled to the same maximum as $T_b(v)$ to highlight the relationship between the two.

Multiple components of cold gas were identified by Murray et al. (2018) using a joint multi-Gaussian decomposition of T_b and τ . Inferred optical depth, velocity, velocity dispersion, spin temperature, and column density of each component can be found in their Table 5. The total mass fraction of cold gas that they detect in absorption along the line of sight against J2232 is 0.30 ± 0.10 (see their Table 2).

Figure 2 shows the amplitude of k_0 normalized Fourier-transformed spectra of T_b and τ on a logarithmic scale, using the same color code as in Figure 1. Several ‘‘dips’’ at similar k_v are observed in the two amplitude spectra, marked by the vertical dashed lines. Because the absorption spectrum (τ) is revealing only cold gas along the line-of-sight, these dips observed in both the FT of τ and T_b suggest that CNM gas plays an important role in shaping the FT of the emission spectrum as well, even though emission includes contributions

⁶ A notebook that illustrates the method is available at <https://github.com/antoinemarchal/FFT-21cm>.

⁷ We will refer to the hyperspectral data cube of Fourier-transformed emission-line spectra in the k_v domain as a position–position– k_v (PPK) cube, akin to a PPV cube that describes the data in the velocity domain.

⁸ Publicly available on the 21-SPONGE [Dataverse](#).

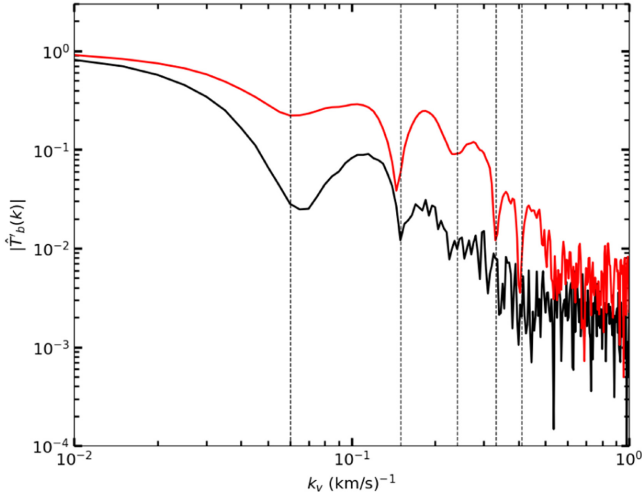


Figure 2. Amplitude of the normalized FT of T_b (black) and τ (red) of the absorption source J2232 from the 21-SPONGE survey. The vertical dashed lines show the position of dips observed in both amplitude spectra at similar k_v .

from HI at all temperatures. Note also that at high k_v , noise dominates the signal in both spectra. However, in the FT of τ , the position of the dips can be traced unambiguously up to relatively high k_v (low velocity dispersion), approximately $k_v \gtrsim 0.4 \text{ (km/s)}^{-1}$.

2.3. Understanding Interference Patterns in the Amplitude of the FT Spectrum

To understand the origin of the observed features in the Fourier-transformed emission-line spectra, we used a toy model based on the mixing of multiple Gaussians with various amplitudes, velocities, and velocity dispersions (an arbitrary mixture of kinetic and turbulent broadening) to generate simplified mock observations of 21 cm emission data.

2.3.1. Toy Model

The generative model is

$$T_b(v, \theta) = \sum_{n=1}^N a_n \exp\left(-\frac{(v - \mu_n)^2}{2\sigma_n^2}\right), \quad (3)$$

where each of the N Gaussians is parameterized by $\theta_n = (a_n, \mu_n, \sigma_n)$, with amplitude $a_n \geq 0$, mean velocity μ_n , and velocity dispersion σ_n . Taking the Fourier transform and using the shift theorem, the normalized FT of the model is

$$\hat{T}_b'(k_v, \theta'_n) = \sum_{n=1}^N f_n \exp\left(-\frac{k_v^2}{2\sigma_{k_v,n}^2}\right) e^{-2\pi j \mu_n k_v}, \quad (4)$$

where $\theta'_n = (f_n, \mu_n, \sigma_{k_v,n})$, $f_n = \sqrt{2\pi} a_n \sigma_n (C/N_{\text{HI}}^*)$ is the mass fraction of each Gaussian n (bounded between zero and unity) relative to the optically thin estimator of the total column density, and $\sigma_{k_v,n} = 1/(2\pi\sigma_n)$.

2.3.2. Single-Gaussian Case

For a simplified model with $N=1$ the complex exponential that defines the phase of the signal, parametrized by μ , simplifies when one considers the normalized amplitude of

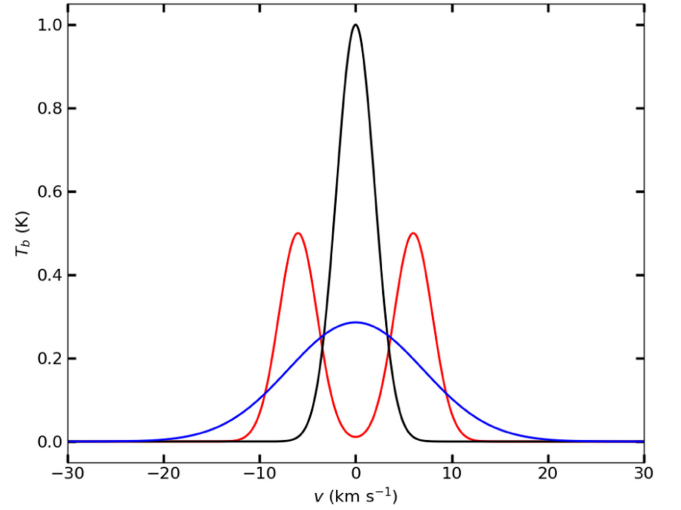


Figure 3. Brightness temperature spectrum T_b of mock models. The black and blue curves show single-Gaussian models with no radial velocity (i.e., centered at $v = 0 \text{ km s}^{-1}$) and $\sigma = 2 \text{ km s}^{-1}$ and $\sigma = 7 \text{ km s}^{-1}$, respectively. The red curve shows a double-Gaussian model with velocity separation $\Delta_v = 12 \text{ km s}^{-1}$.

the FT,

$$|\hat{T}_b'(k_v, \theta')| = f \exp\left(-\frac{k_v^2}{2\sigma_{k_v}^2}\right), \quad (5)$$

and can be fully parametrized by a single $f=1$ and the corresponding σ_{k_v} , dropping the subscript $n=1$ in Equation (4). The amplitude of the FT of a single Gaussian is also a Gaussian, centered on $k_v=0$ with a width inversely proportional to the velocity dispersion of the original Gaussian.

Therefore, a narrow (broad) Gaussian describing cold (warm) gas $T_b(v, \theta)$ spectrum will appear broad (narrow) in its FT amplitude spectrum. This is illustrated in Figure 3 by the narrow ($\sigma = 2 \text{ km s}^{-1}$) and broad ($\sigma = 7 \text{ km s}^{-1}$) Gaussians with $\mu = 0 \text{ km s}^{-1}$ (i.e., centered at $v = 0 \text{ km s}^{-1}$; black and blue lines, respectively), chosen to have the same area, representing the same column density of the gas. Their corresponding normalized FT amplitude spectra are shown in Figure 4.

2.3.3. Two Centered Gaussians

Here the spectral model is the sum of the narrow (black) and broad (blue) Gaussians in Figure 3, though not shown there. Thus, there is twice as much gas, half of which could be classified as cold; i.e., the cold gas mass fraction is 0.5. The normalized amplitude of the FT is given by

$$|\hat{T}_b'(k_v, \theta')| = f \sum_{n=1}^{N=2} \exp\left(-\frac{k_v^2}{2\sigma_{k_v,n}^2}\right), \quad (6)$$

where $f=f_1=f_2=0.5$. Because $\mu_1 = \mu_2$ there is no information from the phase of the FT (see Equation (7)), and Equation (6) has the same functional form as the spectral model (i.e., a sum of two Gaussians). However, now the broad component comes from the cold gas (black, $\sigma_{k_v,1} = 0.079577 \text{ (km s}^{-1}\text{)}^{-1}$), while the narrow component originates from the warm gas (blue, $\sigma_{k_v,2} = 0.022736 \text{ (km s}^{-1}\text{)}^{-1}$). The normalized amplitude is plotted as the cyan curve in Figure 4.

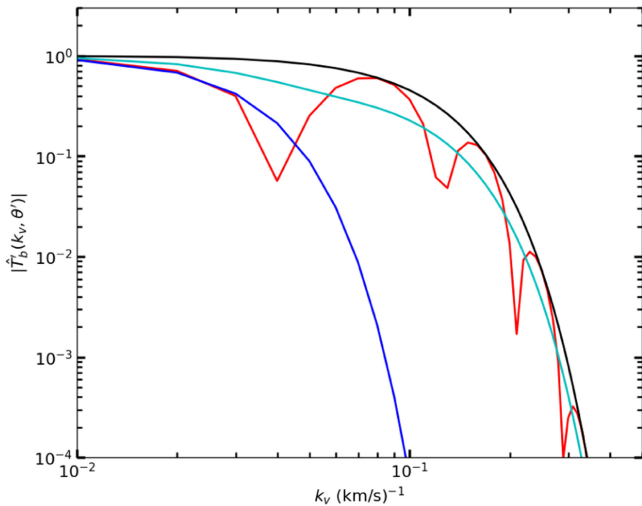


Figure 4. Normalized amplitude of the FT of the brightness temperature spectrum T_b of toy models. Colors are the same as in Figure 3, with the addition of the cyan curve for the model with two centered Gaussians in Section 2.3.3.

2.3.4. Analogy to Fraunhofer Diffraction

It is useful to make an analogy to Fraunhofer diffraction. When a plane wave passes through an aperture, the patterns on a distant screen can be calculated by the interference of secondary wavelets according to the Huygens–Fresnel principle. The diffraction pattern calculated for a sharply bounded transparent aperture has characteristic oscillations, but the diffraction pattern calculated for an aperture with a Gaussian transmission profile is a smooth Gaussian. This can be generalized to the case of multiple combined Gaussian transmission profiles of different widths but all centered at the same position.

Alternatively, in Fourier optics, the Fraunhofer diffraction can be written as the FT of the aperture function, whose functional form again determines the shape of the diffraction pattern observed on the distant screen. Here, the brightness temperature $T_b(v)$ can be seen as analogous to a 1D aperture function and so the amplitude of its FT is analogous to the amplitude of the resulting 1D Fraunhofer diffraction pattern.

2.3.5. A “Double-slit” Analogy

When a plane wave passes through two slits, there is further interference of the secondary wavelets to be accounted for. The resulting diffraction can again be calculated with the Fourier transform. This can be appreciated with a simple double-Gaussian model with $a_1 = a_2 = a = 0.5$, and $\sigma_1 = \sigma_2 = 2 \text{ km s}^{-1}$, and $\mu_1 = -\mu_2 = 6 \text{ km s}^{-1}$. Note that this model, red in Figure 3, has the same column density of gas as the single-Gaussian model (black) with the same σ .

Equation (4) simplifies to

$$|\hat{T}_b'(k_v, \theta')| = f \exp\left(-\frac{k_v^2}{2\sigma^2}\right) |\cos(\pi \Delta_v k_v)|, \quad (7)$$

where $f=1$ and $\Delta_v = \mu_1 - \mu_2 = 12 \text{ km s}^{-1}$ is the velocity separation between the two Gaussians. This is the same as Equation (5) for the single Gaussian, except now modulated by $|\cos(\pi \Delta_v k)|$. The normalized amplitude of its FT is shown by the red curve in Figure 4. The spectral resolution was set to 0.1 km s^{-1} . This modulation affects the Gaussian function only destructively and so the single-Gaussian case (black) forms an

envelope to the amplitude of the FT of the double-Gaussian model (red), as is familiar in the Fraunhofer diffraction pattern of any double-slit experiment.

2.3.6. Multiple Gaussians

Using Equation (3), we can generate more realistic synthetic spectra of a turbulent fluid with multiple components of various widths representing multiple thermal phases. In particular, we built six-Gaussian models with one Gaussian assigned to the velocity dispersion range $6 \text{ km s}^{-1} < \sigma < 8 \text{ km s}^{-1}$ (approximately $4400 \text{ K} < T_{k,\text{max}} < 7800 \text{ K}$), two in the range $3\text{--}6 \text{ km s}^{-1}$ (approximately $1100 \text{ K} < T_{k,\text{max}} < 4400 \text{ K}$), and three in the range $0.7\text{--}2 \text{ km s}^{-1}$ (about $70 \text{ K} < T_{k,\text{max}} < 500 \text{ K}$). The six amplitudes a_n were chosen randomly, in the range $0\text{--}16 \text{ K}$.

A baseline model was made with all components centered at $\mu_n = 0$ to produce a nondestructive reference. Next, five more models were built with six velocities μ_n randomly chosen in the velocity range $-6 \text{ km s}^{-1} < \mu < 6 \text{ km s}^{-1}$, to be different for each of these five models. However, the six models share the same amplitudes a_n and velocity dispersions σ_n and hence have the same mass fractions associated with each of the six individual Gaussians.

Figure 5 shows the brightness temperature of all six models for a case with a cold gas mass fraction (the three Gaussians with $\sigma < 3 \text{ km s}^{-1}$) of approximately 0.5. The spectral resolution was again set to 0.1 km s^{-1} . The black dashed curve shows the baseline model with $\mu_n = 0$. The five-colored curves show models with varying μ_n . Figure 6 shows their FT amplitude spectra, using the same color code. As expected from the single and double-Gaussian models, interference patterns are seen in the five models with $\mu_n \neq 0$, resulting from the multiple peaks in their T_b spectra. As in the double-Gaussian model, interference appears only destructively and the $\mu_n = 0$ baseline model provides an envelope for the other five models, regardless of the specific patterns produced by the random positions of each Gaussian.

2.4. A Lower Limit on the Cold Gas Mass Fraction

Building on our understanding of patterns produced in the FT amplitude spectrum by the multicomponent nature (phase and clouds/turbulent structures) of 21 cm data, here we develop a new method to set a lower limit on the mass fraction of gas with an arbitrary maximum velocity dispersion (maximum Kinetic temperature) in PPV cubes of HI data. In particular, we focused our analysis on the cold gas mass fraction.

Figure 7 shows the normalized amplitude of the FT of single-Gaussian models with $0.5 < \sigma < 8 \text{ km s}^{-1}$. Gaussians with larger σ appear narrower in the k_v domain. In general, the presence of cold gas (narrow spectral lines) is detectable by the fact that the amplitude of the FT remains high at relatively large k_v . Furthermore, if there is warm gas contributing to the spectrum, it will not be detectable in the amplitude of the FT at high k_v , but it will affect the normalization. Therefore, what we can estimate at high k_v is a lower limit to the cold gas mass fraction of the emitting gas.

Focusing on colder gas, at the vertical black line $k_v = k_{v,\text{lim}} = 0.12 \text{ (km s}^{-1}\text{)}^{-1}$ gas characterized by a Gaussian with $\sigma = \sigma_{\text{lim}} = 3 \text{ km s}^{-1}$ can be detected in the FT, though only at a threshold fractional level $t = t_{\text{lim}} = 0.05$ of its true total mass. The value of the normalized amplitude is thus an

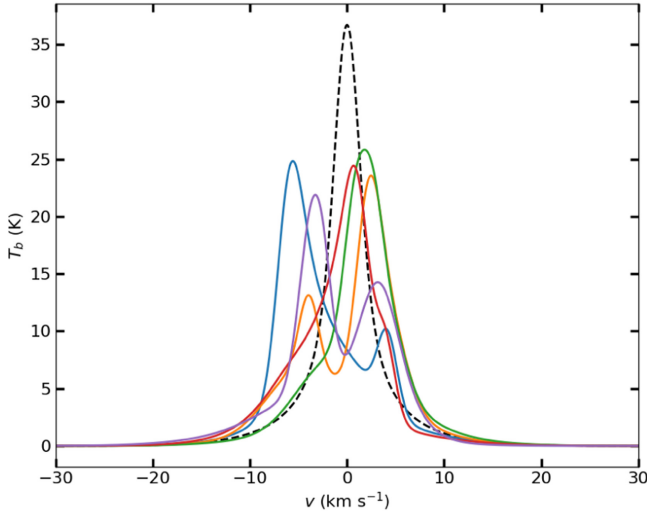


Figure 5. Brightness temperature spectrum T_b of six-Gaussian toy models of a multi-Gaussian component multiphase gas. The black dashed curve shows T_b of the baseline model with $\mu_n = 0$, and randomly selected a_n and σ_n (see text). Colored lines show the T_b of five more models with the same a_n and σ_n but with μ_n chosen randomly in the range $-6 \text{ km s}^{-1} < \mu < 6 \text{ km s}^{-1}$.

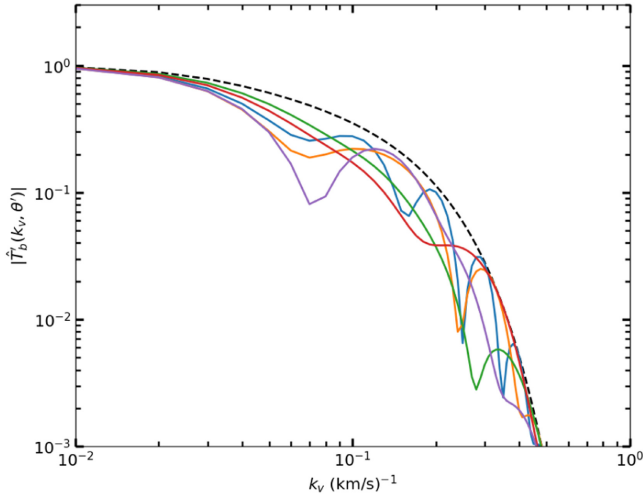


Figure 6. Normalized amplitude of the FT of the brightness temperature spectrum T_b of toy models of a multi-Gaussian component multiphase gas. Annotations are as in Figure 5.

underestimate of the total gas present in an amount dependent on k_v and σ .

Without measurements of the spin temperature T_s , which is close to the Kinetic temperature T_k in cold gas, we need to consider line widths in the T_b spectra. Using the approximation

$$T_{k,\text{max}} \simeq 121 \sigma_{\text{lim}}^2 \quad (8)$$

in units of K, $\sigma = 3 \text{ km s}^{-1}$ corresponds to gas with a maximum Kinetic temperature of approximately 1000 K. In reality, accounting for turbulent broadening, T_k is probably much less. This can be assessed by looking at the data in the BIGHICAT meta-catalog compiled by McClure-Griffiths et al. (2023; Section 4.4). Figure 8 shows a histogram of σ for all Gaussian components identified in absorption with closely matching profiles in emission. These analyses typically use $T_s < 250 \text{ K}$ to define CNM, not the line width (because of turbulent

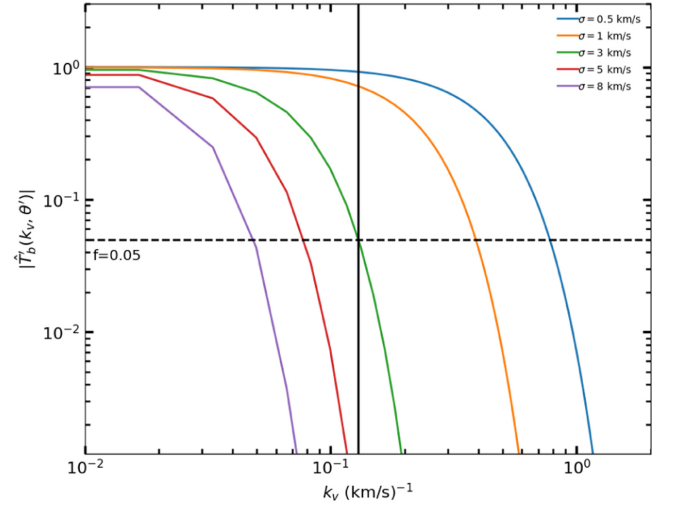


Figure 7. Normalized amplitude of the FT spectrum of single Gaussians with various widths σ . The horizontal dashed line denotes the mass fraction $f = 0.05$ and the vertical line its corresponding k_v for the Gaussian with $\sigma = 3 \text{ km s}^{-1}$.

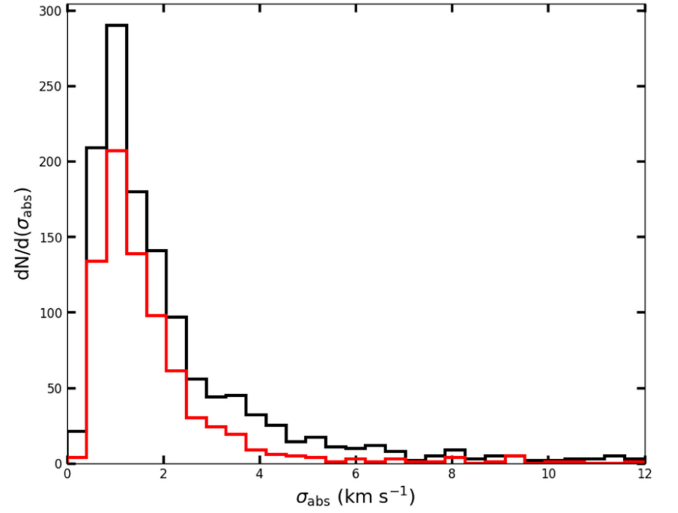


Figure 8. Histogram of σ of absorption-line components in the BIGHIAT catalog (black). The histogram for the CNM subset, defined as having $T_s < 250 \text{ K}$, is also plotted (red).

broadening). With that definition, we overplot the histogram for the subset that are CNM components. This shows that we want an estimator that detects gas with σ up to a few km s^{-1} . Most of this gas will be CNM and any unstable LNM will be suppressed.

As first introduced in Section 2.3.5, destructive interference patterns caused by multiple peaks in the T_b spectra can reduce the amplitude of the FT at $k_v = k_{v,\text{lim}}$, lowering the detectability of the contribution to the T_b signal coming from Gaussians with $\sigma \leq \sigma_{\text{lim}}$. Note that because of the modulation, higher values of the amplitude might occur at higher $k_v > k_{v,\text{lim}}$, improving detectability, though still providing a lower limit on the gas mass fraction.

Following this discussion, we define the estimator of the lower limit on the cold gas mass fraction to be

$$f_{\text{low}}^{k_{v,\text{lim}}} = \max [\hat{T}'_b(k_v > k_{v,\text{lim}})]. \quad (9)$$

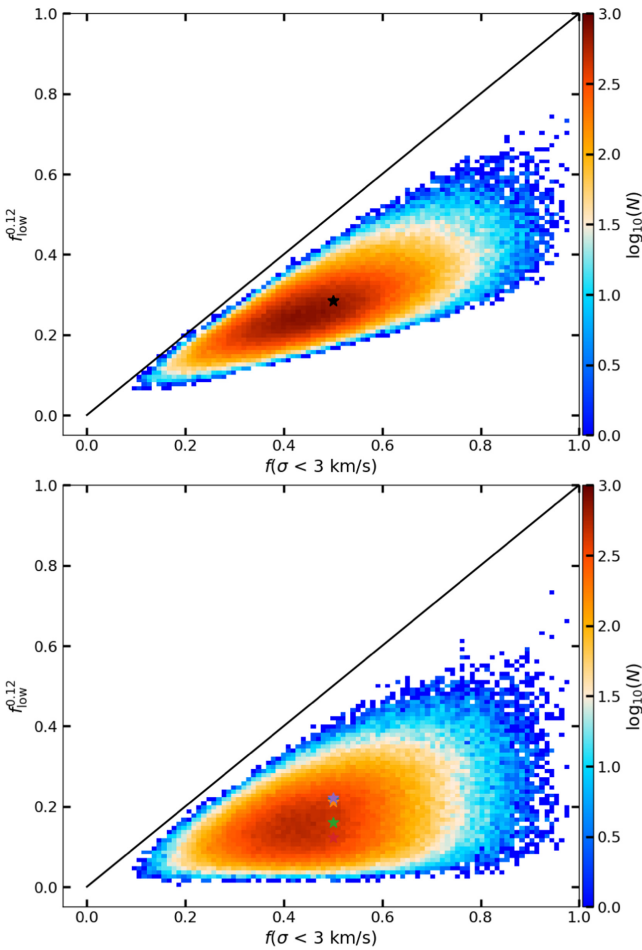


Figure 9. 2D histograms of the estimator $f_{\text{low}}^{0.12}$ and the actual $f(\sigma < 3 \text{ km s}^{-1})$ for 2^{18} six-Gaussian toy models with random amplitudes and dispersions in the parameter ranges used in Figure 5 (see text). Top: spectra modified to have all $\mu_n = 0$. The black star is for the baseline model in Figure 5. Bottom: spectra retaining random μ_n in the range $-6 \text{ km s}^{-1} < \mu < 6 \text{ km s}^{-1}$. The colored stars are for the other five spectra that have interference in Figure 6, which lowers the estimator.

Except where specified in the appendices, in the rest of this paper we have evaluated this estimator of the cold gas mass fraction with $k_{v,\text{lim}} = 0.12 \text{ (km s}^{-1})^{-1}$. For compactness, this is denoted $f_{\text{low}}^{0.12}$. For other applications based on the maximum Kinetic temperature one wishes to probe in the data (e.g., probing the structure of cold gas and unstable gas simultaneously), $k_{v,\text{lim}}$ can be chosen judiciously following the methodology previously described pairing σ_{lim} and t_{lim} .

2.4.1. Assessing $f_{\text{low}}^{0.12}$ for Toy Models

First, we evaluated $f_{\text{low}}^{0.12}$ for the instructive cases in Figure 4. In the case of a simple spectrum with a single peak, the amplitude of the FT (blue, black) gives a direct quantitative estimate. Our estimator evaluates as 1×10^{-7} for the warm gas (blue); i.e., warm is not detectable. For the cold gas (black) the estimate is 0.26 for the selected $k_{v,\text{lim}}$, thus detectable but lower than the true value of the cold gas mass fraction, which is 1.0. In this simple model, only for σ even smaller than 2 km s^{-1} does this estimator approach unity (Figure 7). For the mixture of warm and cold gas (cyan), the warm gas is not detectable directly but affects the normalization and we find 0.13

compared to the true value of 0.5. For the double-Gaussian model (red), which has the same amount of cold gas as the black envelope because of the modulation, by a factor depending on the (range of) k_v examined. Because of this modulation, the estimate of the cold gas mass fraction in more complex spectra can be even lower than the true value; for the red curve we find 0.14, compared to 0.26 for the black curve and the true 1.0.

Similarly, for the more realistic six-Gaussian models in Figure 6, we have evaluated $f_{\text{low}}^{0.12}$. Our estimator evaluates as 0.26 for the baseline model (black line in Figure 6), approximately half of the true input cold gas mass fraction, which is approximately 0.5. For the related five models with random μ_n in the range $-6 \text{ km s}^{-1} < \mu < 6 \text{ km s}^{-1}$ (colored lines), our estimator fluctuates due to interference but is lower than the value inferred from the baseline model.

We have expanded this six-Gaussian experiment to a sample of 2^{18} spectra with randomly selected amplitudes and dispersions (hence cold gas mass fractions) and velocities μ_n in the same parameter ranges described in connection with Figure 5. The top panel of Figure 9 shows the 2D histogram of the estimator $f_{\text{low}}^{0.12}$ versus the actual $f(\sigma < 3 \text{ km s}^{-1})$ for modified spectra resulting from setting all $\mu_n = 0$, like the baseline model in Figure 5, which is shown here as the black star. For a constant $f(\sigma < 3 \text{ km s}^{-1})$, there is a vertical spread of $f_{\text{low}}^{0.12}$ though still consistent with being a lower limit. This illustrates well that our estimator is sensitive to the specific distribution of the underlying amplitudes and dispersions for all phases.

The bottom panel is based on the original 2^{18} spectra including the nonzero μ_n . The colored stars are for the other five spectra in Figure 6. This illustrates again that destructive interference lowers $f_{\text{low}}^{0.12}$ but our estimator is still aptly called a lower limit. We also note that even with interference, $f_{\text{low}}^{0.12}$ retains a good correlation with $f(\sigma < 3 \text{ km s}^{-1})$.

3. Assessing $f_{\text{low}}^{0.12}$ Using a Numerical Simulation

Because of the simplicity of the toy models, it is unclear if $f_{\text{low}}^{0.12}$ also retains information about the spatial structure of cold gas in a realistic situation. This is investigated here with the high-resolution hydrodynamical simulation of thermally bi-stable turbulence performed by Saury et al. (2014) using the three-dimensional radiation hydrodynamics code HERACLES (González et al. 2007).

3.1. Numerical Simulation

In HERACLES, heating and cooling processes were implemented based on the prescription of Wolfire et al. (1995, 2003), including cooling by [C II], [O I], recombination onto positively charged interstellar grains, and Ly α , and heating dominated by the photoelectric effect on small dust grains and PAHs. In their simulation suite, Saury et al. (2014) considered a spatially uniform radiation field with the spectrum and intensity of the Habing field ($G_0/1.7$, where G_0 is the Draine flux: Habing 1968; Draine 1978).

In particular, we used part of their 1024N01 simulation (1024^3 pixels and a physical size of the box of 40 pc) characterized by an initial volume density $n_0 = 1 \text{ cm}^{-3}$, a large-scale velocity $v_s = 12.5 \text{ km s}^{-1}$ and a spectral weight $\zeta = 0.2$ that controls the modes of the turbulent mixing (here a majority of compressible modes). The initial density corresponds to the

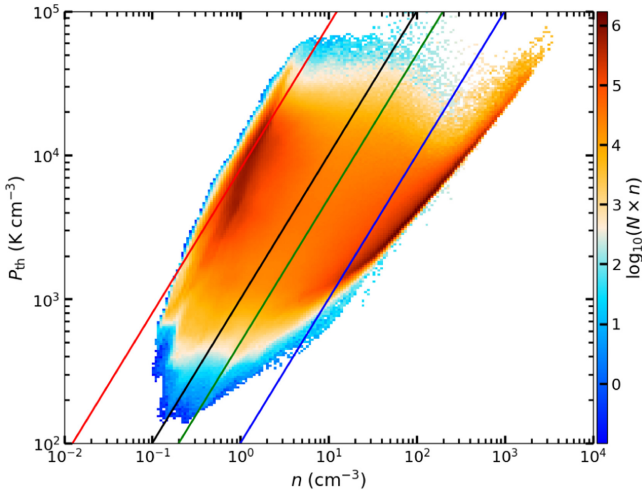


Figure 10. Pressure vs. volume density histogram weighted by volume density. The red, black, green, and blue lines indicate isothermal curves at 8000, 1000, 500, and 100 K, respectively.

typical density of the WNM before thermal instability and condensation, and the large-scale velocity represents the amplitude given to the field that generates large-scale turbulent motions in the box. A stationary state was reached after running the simulation for approximately 16 Myr out of a total of 40 Myr. The snapshot used here was taken after 16 Myr (Saury et al. 2014).

Figure 10 shows the 2D histogram of thermal pressure P_{th} and volume density n for cells in the numerical simulation, weighted by n . The red, black, green, and blue lines indicate isothermal curves at 8000, 1000, 500, and 100 K, respectively.

To explore the effect of opacity we concentrate our analysis on cells in a $512 \times 512 \times 1024$ pixel region with a large range of cold gas mass fraction.

3.2. 21 cm Line Synthetic Observations

Following Marchal et al. (2019), we generated synthetic 21 cm HI observations both in the optically thin limit and considering the full radiative transfer, which is sensitive to the optical depth of the line. We used an effective velocity resolution of 0.8 km s^{-1} in the velocity range $-40 < v < 40 \text{ km s}^{-1}$.

In a given cell at spatial coordinates \mathbf{r} , the turbulent velocity field of the gas along the line of sight (taken to be the z -axis) is $v_z(\mathbf{r}, z)$. The 1D distribution function of the z -component of the Maxwellian velocity distribution in that cell is given by $\phi_{v_z}(\mathbf{r}, z)$, a normal distribution (Gaussian) shifted by $v_z(\mathbf{r}, z)$,

$$\begin{aligned} \phi_{v_z}(\mathbf{r}, z) dv' &= \frac{1}{\sqrt{2\pi} \Delta(\mathbf{r}, z)} \times \exp\left(-\frac{(v' - v_z(\mathbf{r}, z))^2}{2\Delta^2(\mathbf{r}, z)}\right) dv', \quad (10) \end{aligned}$$

where $\Delta(\mathbf{r}, z) = \sqrt{k_B T(\mathbf{r}, z)/m_H}$ is the thermal broadening of the 21 cm line, m_H is the mass of the hydrogen atom, k_B is the Boltzmann constant, and $T(\mathbf{r}, z)$ is the gas temperature.

The general case for the computation of the 21 cm brightness temperature $T_b(v_z, \mathbf{r})$ is based on the following radiative transfer equation:

$$T_b(v_z, \mathbf{r}) = \sum_z T_s(\mathbf{r}, z) [1 - e^{-\tau(v_z, \mathbf{r}, z)}] e^{-\sum_{z' < z} \tau(v_z, \mathbf{r}, z')}, \quad (11)$$

where $\tau(v_z, \mathbf{r}, z)$ is the optical depth of the 21 cm line across the cell defined as

$$\tau(v_z, \mathbf{r}, z) = \frac{1}{C} \frac{\rho(\mathbf{r}, z) \phi_{v_z}(\mathbf{r}, z)}{T_s(\mathbf{r}, z)} dz, \quad (12)$$

where $\rho(\mathbf{r}, z)$ is the gas density. In this representation, emission by gas in the cell at z is absorbed by gas in foreground cells at positions $z' < z$.

A complete derivation of $T_s(\mathbf{r}, z)$ requires the determination of the level populations of the hyper-fine state of the 21 cm line (see, e.g., Section 2.3 in Kim et al. 2014). The three dominant mechanisms are: collisional transitions, direct radiative transitions by 21 cm photons, and indirect radiative transitions (e.g., Ly α scattering). To simplify our calculation of the opacity, we assume here that $T_s(\mathbf{r}, z) = T(\mathbf{r}, z)$, motivated by the fact that collisions dominate the transition between levels in the cold dense phase. In the warmer phase T_s might depart from T if the rate of collisional transitions is insufficient to thermalize the gas, resulting in low values of T_s in the range 1000–4000 K (Liszt 2001). However, the detection of a widespread warm component in 21-SPONGE studies with $\langle T_s \rangle = 7200_{-1200}^{+1800}$ K (Murray et al. 2014) or higher,⁹ suggests that additional excitation mechanisms such as resonant Ly α scattering are present, making T_s close to the Kinetic temperature of the gas.

In the optically thin limit where the opacity is negligible (i.e., $\tau(v_z, \mathbf{r}, z) \ll 1$ everywhere), the 21 cm brightness temperature is

$$T_b^*(v_z, \mathbf{r}) dv' = \frac{1}{C} \int_0^L dz \rho(\mathbf{r}, z) \phi_{v_z}(\mathbf{r}, z) dv', \quad (13)$$

where L is the total depth of atomic gas along the line of sight. Of course, $T_b^*(v_z, \mathbf{r})$ can be calculated even when the gas is not optically thin, and the difference of $T_b(v_z, \mathbf{r})$ with respect to this reveals the effect of opacity.

We generated two additional cubes, $T_b(v_z=0, \mathbf{r})$ and $T_b^*(v_z=0, \mathbf{r})$, both with the turbulent velocity field fixed at 0 km s^{-1} for each cell in the simulation. These four cubes were used to explore the independent effects of opacity and of interference patterns induced by the turbulent velocity field on the Fourier-transformed 21 cm emission line.

Figure 11 shows the brightness temperature spectra of a randomly selected line of sight within these four variants of the simulated PPV cube to illustrate the shape of the simulated spectra and the effects of opacity and velocity field on the line.

3.3. Testing the Lower Limit

As a baseline, Figure 12 shows a projection of the mass fraction of cold gas in the simulation, obtained by integration of cells with $T < T_{k,\text{max}} = 1089 \text{ K}$. This is the “true” cold gas mass fraction, denoted $f(T < T_{k,\text{max}})$.

Figure 13 shows maps of $f_{\text{low}}^{0.12}$ evaluated using Equation (9) on the four different synthetic spectral cubes. For all cases, the map of the lower limit on the cold gas mass fraction strongly resembles the map of the true mass fraction obtained directly from the density field alone ($f(T < T_{k,\text{max}})$), shown in Figure 12. The spatial structure of the cold gas mass fraction is highlighted well.

⁹ Following spectral line modeling, Murray et al. (2018) stacked residual absorption in bins of residual emission and detected a significant absorption feature with a harmonic mean T_s of approximately 10^4 K .

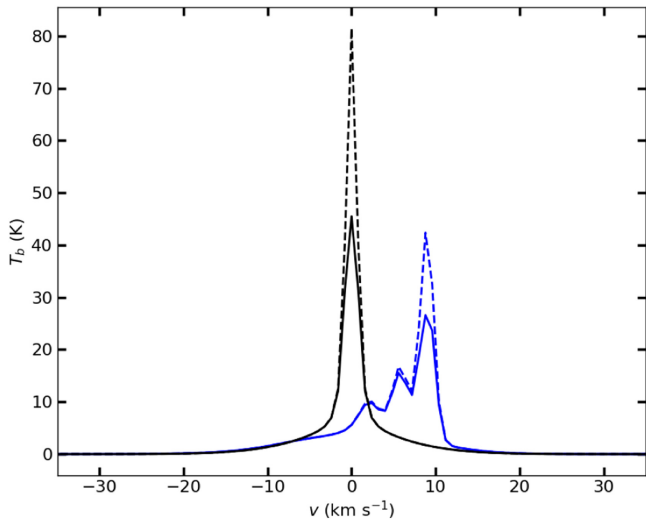


Figure 11. Brightness temperature spectra of a randomly selected line of sight within the four variants of the simulated PPV cube: $T_b^*(v_z = 0, \mathbf{r})$ (dashed black), $T_b^*(v_z, \mathbf{r})$ (dashed blue), $T_b(v_z = 0, \mathbf{r})$ (black), and $T_b(v_z, \mathbf{r})$ (blue).

This similarity is quantified in Figure 14 which shows 2D histograms of $f_{\text{low}}^{0.12}$ versus $f(T < T_{k,\text{max}})$ for the four different cases. Table 1 summarizes four statistics of the correlations: the Pearson correlation coefficient r_p , the slope and intercept, and the scale-averaged¹⁰ cross-correlation coefficient $\langle cc(k) \rangle$.

Regardless of the opacity regime (top row or bottom row in Figures 13 and 14; columns (2) and (3) or (4) and (5) in Table 1), r_p is the highest when $v_z(\mathbf{r}, z) = 0$ (left column in figures; columns (2) and (4) in table). Interference patterns induced by the turbulent velocity field (right column in figures; columns (3) and (5) in table) lower the slope and spread the linear correlation with increasing “true” cold gas mass fraction, resulting in lower r_p . The synthetic spectra are indeed expected to be more complex (possibly showing multiple peaks) and their FTs to be more susceptible to significant interference patterns as the mass fraction of cold gas rises.

Regardless of the velocity field (left or right column), the effect of opacity (bottom row) translates into a saturation of the upper envelope to $f_{\text{low}}^{0.12}$ at approximately 0.5. This saturation is intrinsic to the simulated spectra where the intensity of their peaks is lowered (compared to the optically thin treatment) by the opacity of the line. Its effect is all the more important when there is more cold gas along the line of sight.

It can be seen visually in the successive panels in Figure 13 that despite the effects of both opacity and the velocity field, information about the projected spatial structure in Figure 12 is retained. The scale-averaged cross-correlation coefficients $\langle cc(k) \rangle$ between $f_{\text{low}}^{0.12}$ and $f(T < T_{k,\text{max}})$ for the four cases are tabulated in Table 1. Across scales and for both treatments of opacity, the imprint of the turbulent velocity field on $f_{\text{low}}^{0.12}$ removes less than 20% of the spatial correlation with $f(T < T_{k,\text{max}})$. Qualitatively, the impact of the turbulent velocity field appears as local attenuation on $f_{\text{low}}^{0.12}$. This is likely to change the statistical properties of the field (hence the statistics of the corresponding column density field).

¹⁰ The scale-averaging was performed by taking the mean of all cross-correlation coefficients. Using the median (not tabulated here) yields a three percent variation around each $\langle cc(k) \rangle$.

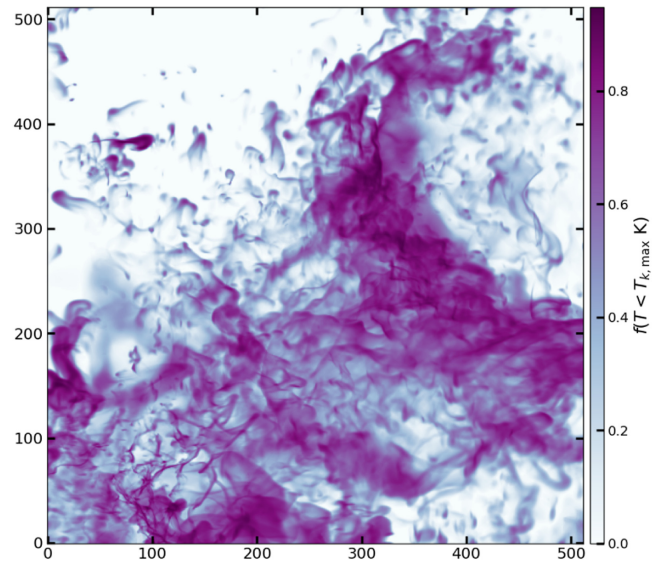


Figure 12. Cold gas mass fraction with $T < T_{k,\text{max}}$ obtained directly from the simulated density field.

4. Assessing $f_{\text{low}}^{0.12}$ with Pointed Observations

4.1. Using T_b and τ Spectra from 21-SPONGE

For an initial exploration of HI observations, we return to the line of sight to J2232 (Galactic latitude $-38^\circ 582$) discussed in Section 2.2, whose 21-SPONGE survey τ and interpolated T_b spectra and FTs are displayed in Figures 1 and 2. Murray et al. (2018) carried out a fairly classical joint analysis of τ and T_b spectra, including Gaussian decomposition and exploiting the fact that the optical depth is inversely dependent on the spin temperature T_s , to obtain a cold gas mass fraction of $f_{21} = 0.30 \pm 0.10$. For our simple estimate, using only the FT of the interpolated emission spectrum T_b , Equation (9) for the lower limit evaluates to $f_{\text{low,Arecibo}}^{0.12} = 0.08$, indeed lower than f_{21} , which is expected to be closer to the true value.

The 21-SPONGE data set covers lines of sight over a range of Galactic latitudes with spectra displaying a variety of complexity. We extended the comparison made for J2232 to the entire data set. A scatter plot is shown in Figure 15, where dots are color coded by absolute Galactic latitude and their sizes encode the complexity of the τ spectra (the number of absorption components, ranging from 0 to 13, identified in the Gaussian decomposition by Murray et al. 2018). The dot with the red edges denotes the line of sight to J2232.

Most of the $f_{\text{low,Arecibo}}^{0.12}$ values are below the 1:1 line shown by the solid black line (i.e., $\langle f_{21} \rangle$), which supports our proposition that $f_{\text{low}}^{0.12}$ would provide a lower limit. The lower limit can approach the true value in some circumstances: at high Galactic latitudes and relatively low cold gas mass fraction ($f_{21} \lesssim 0.2$) and spectral complexity (see encoding of dots), we observe a fair agreement between $f_{\text{low,Arecibo}}^{0.12}$ and f_{21} . See also the discussion of Figure 16 below. However, for $f_{21} \gtrsim 0.2$ and increasing spectral complexity, $f_{\text{low,Arecibo}}^{0.12}$ saturates below approximately 0.1. At lower Galactic latitudes the trend is similar, but the values of $f_{\text{low,Arecibo}}^{0.12}$ are systematically lower for a given f_{21} and saturate below approximately 0.03. This relative behavior of $f_{\text{low,Arecibo}}^{0.12}$ is plausibly due to a combination of increasing opacity coupled with a higher complexity of the emission line (i.e., more components seen in

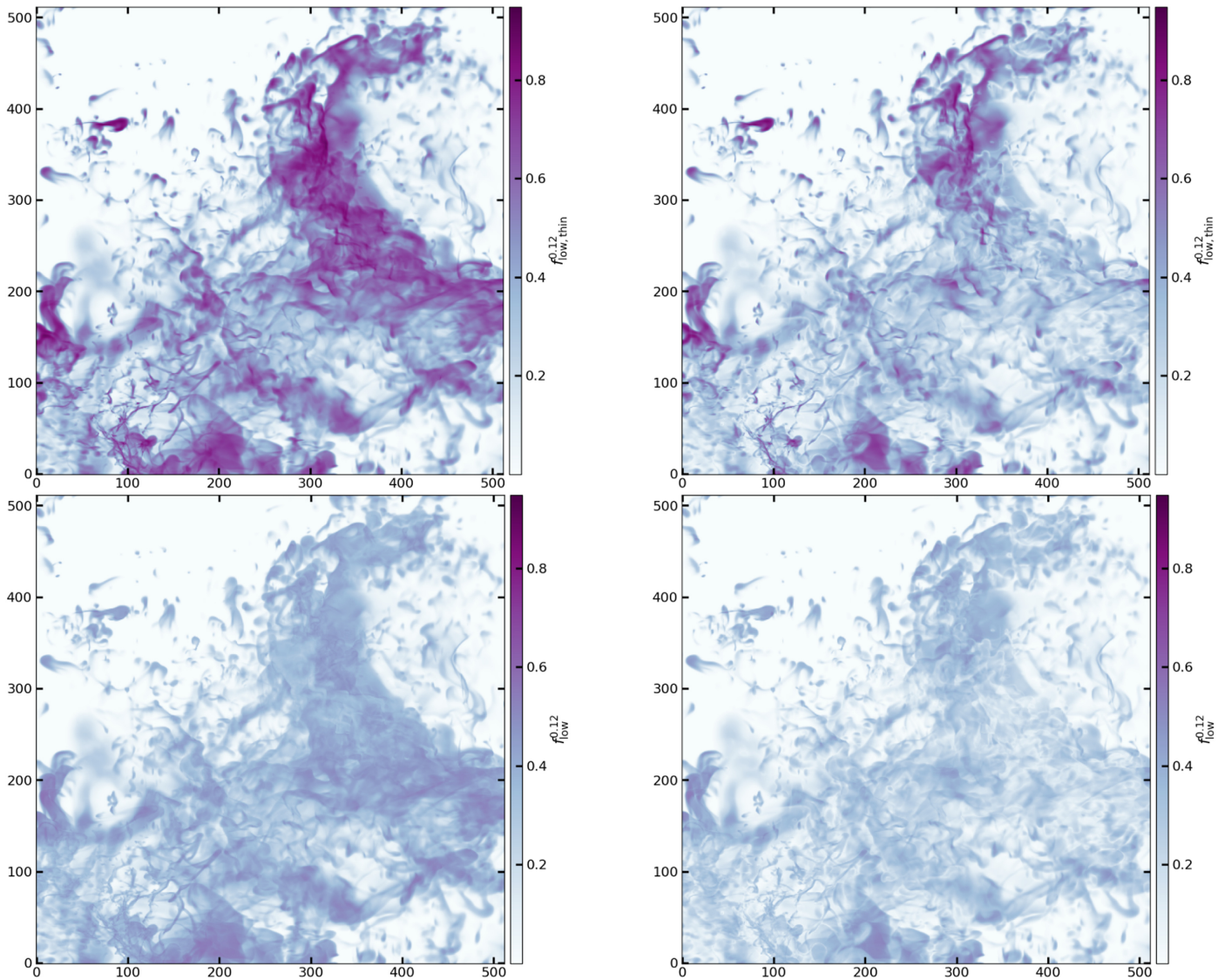


Figure 13. Map of lower limit on the cold gas mass fraction $f_{\text{low}}^{0.12}$ evaluated with Equation (9) for $T_b^*(v_z = 0, r)$ (top left), $T_b^*(v_z, r)$ (top right), $T_b(v_z = 0, r)$ (bottom left), and $T_b(v_z, r)$ (bottom right). Compare to $f(T < T_{k,\text{max}})$ in Figure 12.

absorption, as shown by the size of each dot, is likely associated with a more complex line profile in emission as well), leading to more interference in the FT.

In the Taurus and Gemini Regions, Nguyen et al. (2019) observed a strong dependency of the correction to the optically thin HI approximation, $\mathcal{R} = N_{\text{HI}}/N_{\text{HI}}^*$, on the cold gas mass fraction. The authors found that for the CNM mass fraction below 0.2, N_{HI}^* and N_{HI} are consistent within the errors. However, for mass fraction higher than 0.2 (and lower than 0.75), \mathcal{R} rises from 1 to 2.

A scatter plot of the difference $f_{21} - f_{\text{low,Arecibo}}^{0.12}$ versus total column density N_{HI} so corrected for opacity is shown in Figure 16. Values of f_{21} and $f_{\text{low,Arecibo}}^{0.12}$ are fairly similar for $N_{\text{HI}} \lesssim 3\text{--}5 \times 10^{20} \text{ cm}^{-2}$, which is close to the typical column density, $N_{\text{HI}} \approx 5 \times 10^{20} \text{ cm}^{-2}$, where opacity effects start to become apparent in various absorption-line surveys (see, e.g., Figure 2 in Nguyen et al. 2018, showing \mathcal{R} versus the opacity-corrected N_{HI} for data in the Arecibo Millennium Survey (Heiles & Troland 2003) and 21-SPONGE). That means that at low column density either our lower limit is close to the true value, which it would tend to be for narrow lines from cold gas

such as are favored by absorption-line measurements (also low turbulent broadening), or the true value f_{21} is also an underestimate. The dot with the red edges denotes the line of sight to J2232 (Figure 2), near this threshold column density. The effect of the combination of increasing opacity coupled with higher complexity of the line (dot size), especially at low latitude, can be appreciated here again.¹¹

Although values of f_{21} and $f_{\text{low,Arecibo}}^{0.12}$ are fairly similar at low column density, we do note that seven lines of sight have $f_{21} - f_{\text{low,Arecibo}}^{0.12} < 0$, which seems to suggest an inconsistency between our lower limit and the cold gas mass fraction measured in absorption. We compared values of this difference with the uncertainties on the cold gas mass fraction calculated by Murray et al. (2020, see Table D1). Four of them are consistent within errors but three lines of sight remain unexplained, namely, 3C245A, 1055 + 018, and 3C273. We have explored this further in Section 4.4 (validation on BIGHICAT) where a significant

¹¹ We observe a similar trend (not shown here) when plotted instead against the total column density N_{HI}^* computed in the optically thin limit.

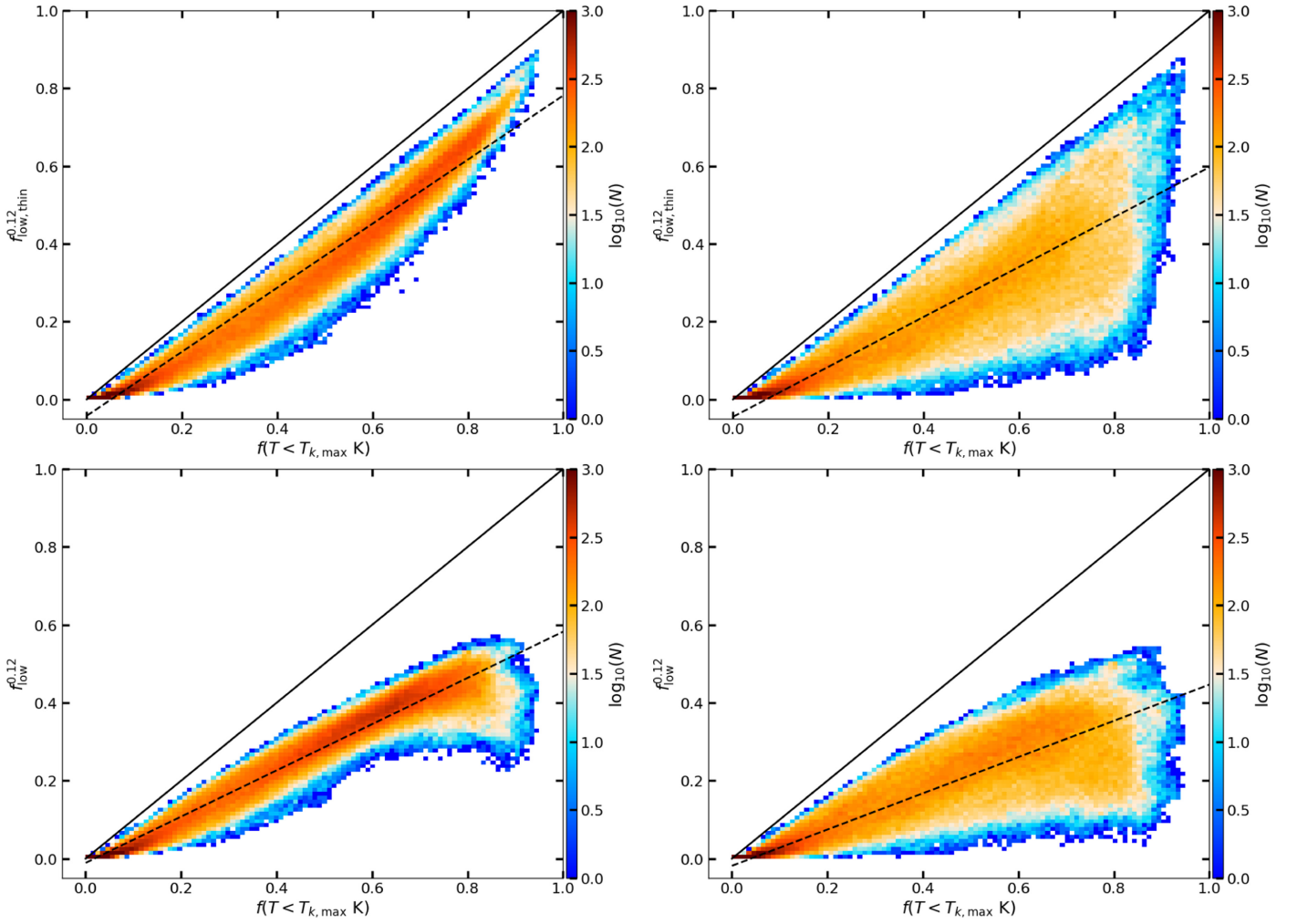


Figure 14. 2D distribution function of $f_{\text{low}}^{0.12}$ evaluated with Equation (9) vs. $f(T < T_{k,\text{max}})$, for $T_b^*(v_z = 0, r)$ (top left), $T_b^*(v_z, r)$ (top right), $T_b(v_z = 0, r)$ (bottom left), and $T_b(v_z, r)$ (bottom right). The solid black line shows the 1:1 line. Dashed lines show linear fits using the bisector estimator.

Table 1

Summary Statistics of the Correlation between $f_{\text{low}}^{0.12}$ and $f(T < T_{k,\text{max}})$ for the Four Cases Presented in Figures 13 and 14

	$T_b^*(v_z = 0, r)$	$T_b^*(v_z, r)$	$T_b(v_z = 0, r)$	$T_b(v_z, r)$
Pearson r_p	0.98	0.89	0.98	0.88
Slope	0.82	0.64	0.59	0.46
Intercept	-0.04	-0.04	0.01	-0.02
$\langle cc(k) \rangle$	0.88	0.75	0.83	0.71

number of lines of sight show the same behavior and cannot be reconciled within uncertainties.

4.2. Lower Resolution Emission-line Data From HI4PI

We evaluated the applicability of our FT method to 21 cm emission-line data with lower spatial and spectral resolution than 21-SPONGE (3'5), namely, that from the HI4PI survey,¹² which has a 16'2 beam and 1.49 km s^{-1} spectral resolution (HI4PI Collaboration et al. 2016). At the position of each

¹² HI4PI is a combined all-sky product based on data from the Effelsberg-Bonn HI Survey (EBHIS; surveyed with the Effelsberg 100 m radio telescope; Kerp et al. 2011; Winkel et al. 2016) and from the third revision of the Galactic All-Sky Survey (GASS; surveyed with the Parkes telescope; McClure-Griffiths et al. 2009; Kalberla et al. 2010; Kalberla & Haud 2015).

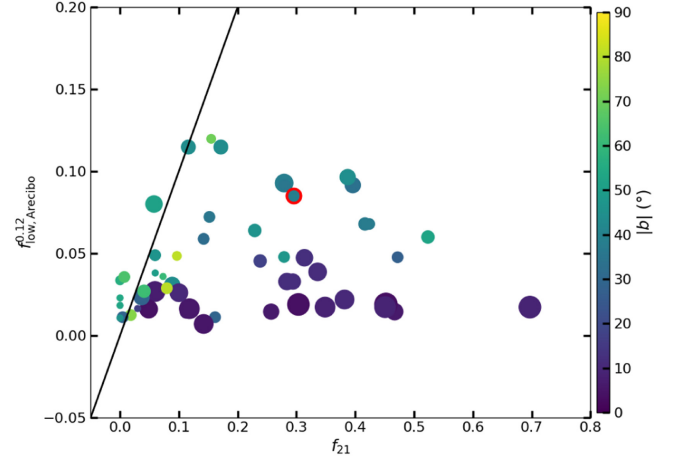


Figure 15. Scatter plot of $f_{\text{low,Areibo}}^{0.12}$ vs. f_{21} for all sources from the 21-SPONGE survey color coded by absolute Galactic latitude. The size of each dot encodes the number of absorption components (ranging from 0 to 13) identified in the Gaussian decomposition of the optical depth spectrum by Murray et al. (2018). The solid black line shows the 1:1 line. The dot with the red edges denotes the line of sight to J2232.

source from 21-SPONGE for which we previously evaluated $f_{\text{low,Areibo}}^{0.12}$, we extracted the on-source T_b spectrum in the spectral range $-90 < v < 90 \text{ km s}^{-1}$, apodizing it with a cosine

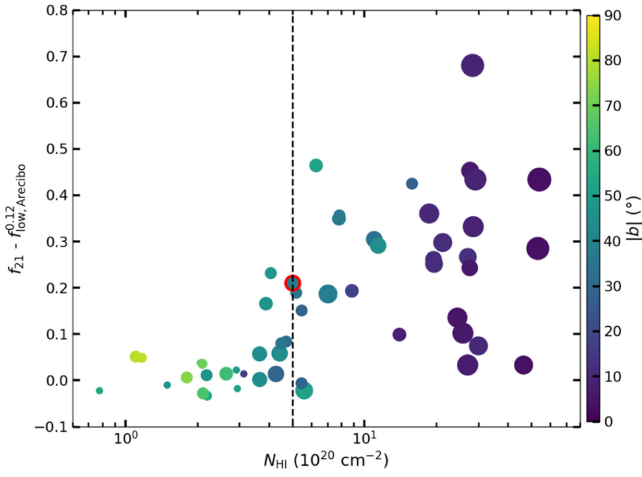


Figure 16. Scatter plot of $f_{21} - f_{\text{low,Arecibo}}^{0.12}$ vs. N_{HI} color coded by absolute Galactic latitude. The dashed vertical line denotes the typical column density where the effects of opacity become significant. The size code is as in Figure 15. The dot with the red edges denotes the line of sight to J2232.

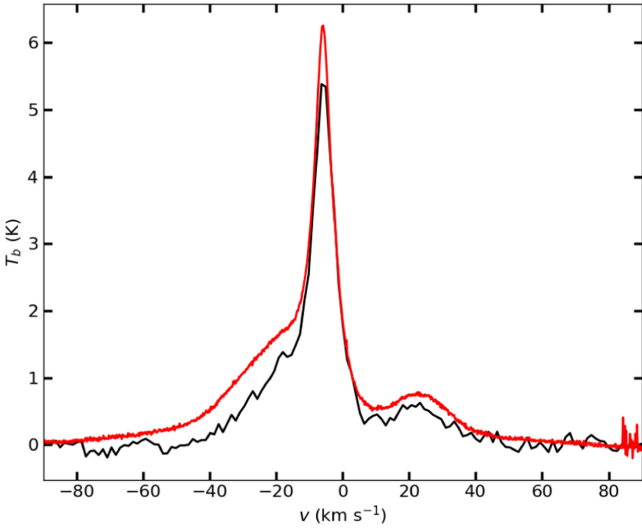


Figure 17. Comparison between the brightness temperature spectrum interpolated around 3C273 (source with the highest T_c) from 21-SPONGE (red) and the on-source spectrum extracted from the HI4PI survey (black).

function at both ends (Section 6.2). On this extracted HI4PI spectrum, we evaluated our FT-based estimate of the lower limit of the cold gas mass fraction, $f_{\text{low, HI4PI}}^{0.12}$.

Using the on-source spectrum rather than an interpolated spectrum assumes that at the low spatial resolution of HI4PI, the peak brightness temperature of the continuum source T_c is diluted so much that the on-source emission spectra do not show the effects of absorption. To quantify this, T_c seen by HI4PI at $16''.2$ is (Blagrave et al. 2017)

$$T_c = 170 \left(\frac{[1']}{16''.2} \right)^2 \frac{S_\nu}{[1 \text{ Jy}]} \text{ K}, \quad (14)$$

where S_ν is the source flux density. The background source is reduced by HI absorption producing a negative-going feature in the continuum-subtracted spectrum with profile $T_c(1 - \exp(-\tau))$. This is to be compared to the emission-line profile, which for a gas of constant spin temperature T_s is $T_b = T_s(1 - \exp(-\tau))$.

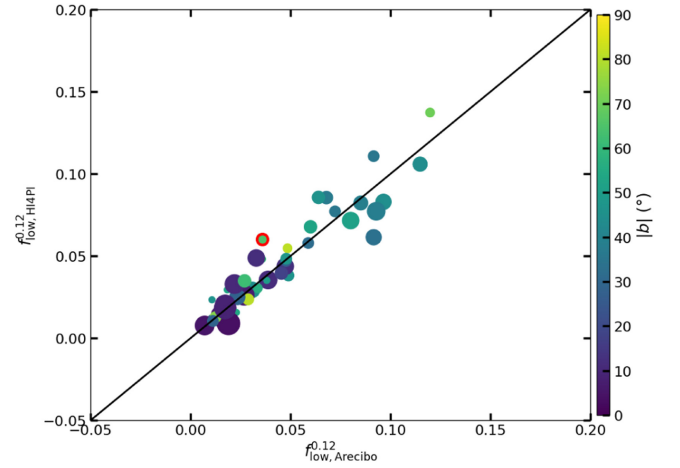


Figure 18. Scatter plot of $f_{\text{low, HI4PI}}^{0.12}$ vs. $f_{\text{low, Arecibo}}^{0.12}$ for lines of sight in 21-SPONGE. Color code, size, and annotations are as in Figure 15. The dot with the red edges denotes the line of sight to 3C273 shown in Figure 17.

Therefore, we want to check that T_c is significantly less than T_s . We adopt $T_s = 88 \text{ K}$ for cold gas (Blagrave et al. 2017) and note that even for the brightest source in 21-SPONGE, 3C273, $T_c = 36 \text{ K}$.

As an empirical check, in Figure 17 we have compared the brightness temperature spectrum interpolated from off-source spectra around 3C273 from 21-SPONGE (red) and the on-source spectrum extracted from the HI4PI survey (black). Their peak brightness temperatures differ by less than 1 K. Importantly for our estimator of the cold gas mass fraction, the value $f_{\text{low, HI4PI}}^{0.12} = 0.06$ for the black spectrum is close to $f_{\text{low, Arecibo}}^{0.12} = 0.036$ for the red spectrum.

In addition, we extracted a 2.56 deg^2 HI4PI brightness temperature PPV cube (not shown here) centered on the sources and found no evidence of attenuation in individual channel maps.

In summary, Figure 18 plots $f_{\text{low, HI4PI}}^{0.12}$ versus $f_{\text{low, Arecibo}}^{0.12}$. Points are scattered around the 1:1 relation shown by the black solid line with a relatively low dispersion that shows that $f_{\text{low}}^{0.12}$ is quite consistent over spatial resolutions varying from $3''.5$ to $16''.2$, at least for lines of sight in 21-SPONGE. Although cold gas emission is likely to arise in small-scale structures, its fractional contribution is not being modified significantly within a $16''.2$ beam. Furthermore, despite the lower velocity resolution of HI4PI, the cold gas with narrow lines is being reliably detected.

4.3. Using a CNN on T_b Spectra

As mentioned in the introduction, Murray et al. (2020) developed a 1D CNN, trained on simulated data that included opacity information and designed to infer an absolute value of the cold gas mass fraction when applied to emission-line data. For the same 21-SPONGE emission data for the J2232 line of sight, using this 1D CNN they obtained $f_{\text{CNN}} = 0.08 \pm 0.03$, similar to the lower limit obtained with our FT method. Both f_{CNN} and $f_{\text{low, Arecibo}}^{0.12}$ are lower than f_{21} .

We used table D1 in Murray et al. (2020) to produce the scatter plot of $f_{\text{CNN}} - f_{\text{low, Arecibo}}^{0.12}$ against N_{HI} shown in Figure 19 for all the lines of sight at high Galactic latitude (i.e., $|b| > 30^\circ$) used in Murray et al. (2020). The color code, size, and annotations are as in Figure 16. Both the CNN and

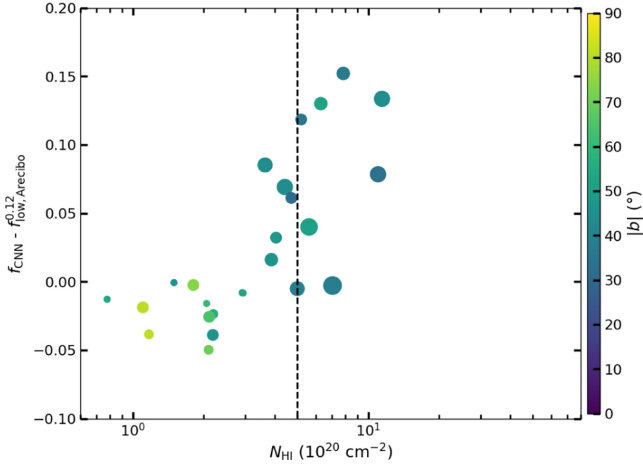


Figure 19. Scatter plot of $f_{\text{CNN}} - f_{\text{low,Arceibo}}^{0.12}$ vs. N_{HI} . Color code, size, and annotations are as in Figure 16.

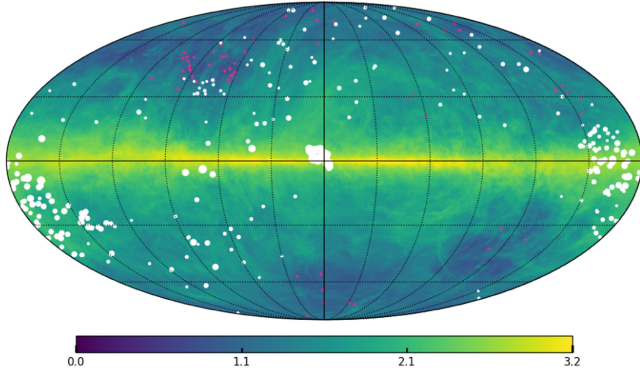


Figure 20. Mollweide projection centered on the Galactic center of the total column density map N_{HI}^* (in units of 10^{19} cm^{-2} , on a logarithmic scale) from the HI4PI survey over the velocity range $-90 < v < 90 \text{ km s}^{-1}$. The purple crosses show the positions of nondetections in the compilation of HI absorption spectra from the BIGHICAT meta-catalog (McClure-Griffiths et al. 2023). The white dots, whose size encodes the number of absorption components (ranging from 1 to 16) in their τ spectrum, show the detections in BIGHICAT.

our FT method are consistent within 10% for $N_{\text{HI}} \lesssim 3\text{--}5 \times 10^{20} \text{ cm}^{-2}$. The FT values are lower than the CNN values above this threshold, which again supports our proposition that $f_{\text{low}}^{0.12}$ is a lower limit, affected by both the opacity of the line and the complexity of the line shape.

4.4. BIGHICAT

We extended the comparison made with 21-SPONGE to the 373 unique lines of sight in the BIGHICAT meta-catalog compiled by McClure-Griffiths et al. (2023).

BIGHICAT combines results from absorption surveys with publicly available spectral Gaussian modeling (Heiles & Troland 2003; Mohan et al. 2004a; Stanimirović et al. 2014; Dénes et al. 2018; Murray et al. 2018, 2021; Nguyen et al. 2019). As McClure-Griffiths et al. (2023) emphasize, BIGHICAT is heterogeneous in terms of sensitivity with median rms noise in the HI optical depth per 1 km s^{-1} channel from 0.0006 (21-SPONGE) to 0.1 (Dénes et al. 2018), angular resolution from $1'$ (CGPS and VGPS; Taylor et al. 2003; Stil et al. 2006) to $36'$ (Mohan et al. 2004a, 2004b; Roy et al. 2013), velocity resolution from 0.1 km s^{-1} (Dénes et al. 2018) to 3.3 km s^{-1} (Mohan et al.

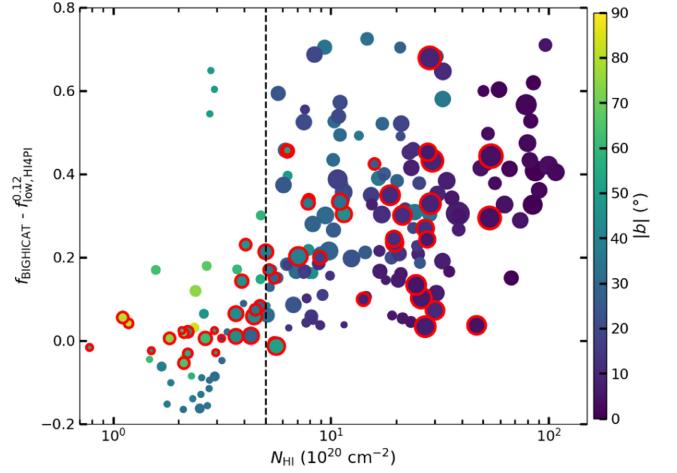


Figure 21. Scatter plot of $f_{\text{BIGHICAT}} - f_{\text{low,HI4PI}}^{0.12}$ vs. N_{HI} color coded by absolute latitude. Color code, size, and annotations are as in Figure 16. The dots with red edges show lines of sight from 21-SPONGE. The locations of lines of sight with a negative difference are shown with green dots in Figure 26.

2004a, 2004b), and the Gaussian decomposition method used to model the data (all but Mohan et al. 2004a and Dénes et al. 2018 follow the methodology developed by Heiles & Troland 2003). While this might affect the homogeneity of the cold gas mass fraction estimates, f_{BIGHICAT} , it provides a larger sample that probes different regions of the sky.

Figure 20 shows the location of the 373 lines of sight in BIGHICAT on a Mollweide projection of the total column density map N_{HI}^* (on a logarithmic scale) from the HI4PI survey over the velocity range $-90 < v < 90 \text{ km s}^{-1}$. The purple crosses show the positions of nondetections and the white dots, whose sizes encode the number of absorption components (ranging from 1 to 16) in their τ spectrum, show the detections.

For comparison to f_{BIGHICAT} , we need an emission spectrum on which to apply our FT method. For a homogeneous product, we used on-source spectra from HI4PI as described in Section 4.2. Figure 21 shows the scatter plot of $f_{\text{BIGHICAT}} - f_{\text{low,HI4PI}}^{0.12}$ versus N_{HI} color coded by absolute latitude for all lines of sight in BIGHICAT. Dots with red edges show lines of sight from 21-SPONGE. Due to the high correlation observed between $f_{\text{low,HI4PI}}^{0.12}$ and $f_{\text{low,Arceibo}}^{0.12}$ (see Figure 18), the distribution for 21-SPONGE lines of sight is similar to that in Figure 16. The comparison with BIGHICAT as a whole shows a similar trend to that of 21-SPONGE only but with many more lines of sight. Again our explanation of the scatter at large column density is that $f_{\text{low,HI4PI}}^{0.12} < f_{\text{BIGHICAT}}$, with the amount of underestimation by our lower limit $f_{\text{low,HI4PI}}^{0.12}$ depending on the opacity and the destructive interference introduced by the complexity of the line profile.

However, at a low column density, there are now many sight lines where our lower limit is higher than the cold gas mass fraction tabulated in BIGHICAT, by nearly 20% (as opposed to less than 10% for 21-SPONGE only). On investigating their distribution in the sky, we found most to be part of the MACH survey (Murray et al. 2021) that targets a small region of the sky located in the first quadrant above the Galactic plane (see the concentration of purple crosses and white dots in Figure 20). This is a region where we found a large cloud of relatively low column density and high cold gas mass fraction

when evaluating Equation (9) on HI4PI over the whole sky (see Section 7 and Figure 26 therein).

Interestingly, in their analysis whose results are incorporated into BIGHICAT Murray et al. (2021) reported a case (source named J14434) where their Gaussian modeling of the spectra yielded no detection of CNM in absorption at local velocities while the emission data from EBHIS at 10'8 resolution show a clear narrow peak at the same velocities. This can be appreciated in their Figure 2 (first column and second row). They suggested that this is likely to reflect the nature of the small-scale structure of the CNM in this region, i.e., a porous medium in which the pencil beam of the absorption-line observation misses the CNM structure in the large beam.¹³ As a result, $f_{\text{BIGHICAT}} = 0$ at local velocities and 0.01 ± 0.02 when considering the full velocity range (incl. an IVC component at approximately -40 km s^{-1}). The case of J14434 suggests that other lines of sight in this region are likely to be affected by the same effect.

EBHIS data appear in HI4PI at degraded 16/2 resolution and it seems reasonable that a 16/2 beam probing the same porous medium would exhibit the same effect in emission. The HI4PI spectrum at the position of the source (not shown here) is similar to that of the EBHIS data, also showing a narrow peak at local velocities. In that specific case, our FT approach finds $f_{\text{low, HI4PI}}^{0.12} = 0.1$. Because of the underlying structure of the phases within the 16/2 beam, $f_{\text{low, HI4PI}}^{0.12} > f_{\text{BIGHICAT}}$ but $f_{\text{low, HI4PI}}^{0.12}$ is still a lower limit at the beam size at which the FT was applied.

5. Assessing $f_{\text{low}}^{0.12}$ Using the ROHSA Thermal Phase Decomposition of LLIV1

Vujeva et al. (2023) used the spectral decomposition code ROHSA (see details in Section 6.3) to model the column density of different thermal phases in the Low-Latitude Intermediate-Velocity Arch 1 (LLIV1) data set from GHIGLS in the velocity range $-81 < v < -27 \text{ km s}^{-1}$ (Vujeva et al. 2023). At each pixel, the ROHSA decomposition is expressed as Gaussian components, each with an amplitude, central velocity (μ), and velocity dispersion (σ). These parameters cluster in a 2D histogram of σ versus μ (the σ - μ diagram, Figure 6 in Vujeva et al. 2023). The Gaussians in clusters with $\sigma \sim 2 \text{ km s}^{-1}$ are interpreted as from CNM gas, leading to a map of the column density of CNM gas, their Figure 7 (upper left); that map, $N_{\text{HI, CNM, ROHSA}}^*$, is the basis for the test here. We have applied the FT method to the same data cube and produced a map of $f_{\text{low}}^{0.12} N_{\text{HI}}^*$, shown in Figure 22.

Visual comparison of these two maps reveals that the FT method retains information about the projected spatial structure, just as we found using simulated data in Section 3.3. However, because $f_{\text{low}}^{0.12}$ is a lower limit, the column density is lower. This is quantified in Figure 23, a 2D histogram correlating the values in the two maps.

The CNM clusters in the ROHSA σ - μ diagram are centered near 2 km s^{-1} . If all CNM Gaussians had this dispersion, then their $f_{\text{low}}^{0.12}$ would give the slope of the dark blue line (second lowest), absent destructive interference effects. However, the clusters have

¹³ Depending on the spatial structure, this seemingly paradoxical situation might be avoided by an emission-line survey at a very high spatial resolution. However, in that case T_c of the radio source would be very high and any absorption would compromise the on-source spectrum. Interpolating emission spectra from the surrounding annulus would reintroduce the issue of sampling different spatial structures.

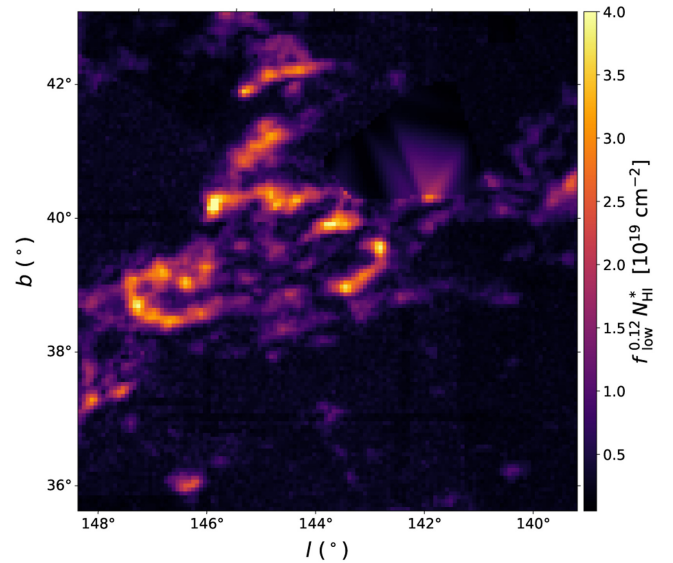


Figure 22. Lower limit on the cold gas column density $f_{\text{low}}^{0.12} N_{\text{HI}}^*$ of Low-Latitude Intermediate-Velocity Arch 1 from GHIGLS in the velocity range $-81 < v < -27 \text{ km s}^{-1}$. The color bar is in units 10^{19} cm^{-2} .

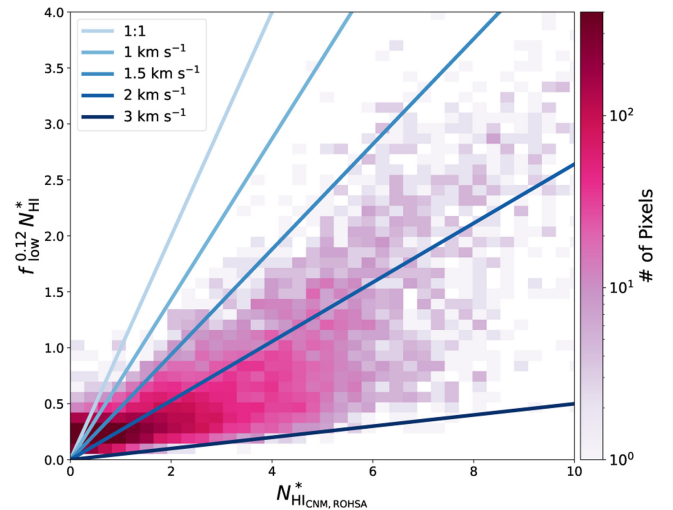


Figure 23. 2D histogram (logarithmic color bar) demonstrating the pixel-by-pixel correlation of the map of $f_{\text{low}}^{0.12} N_{\text{HI}}^*$ with the CNM column density map obtained using ROHSA (Figure 7 upper left of Vujeva et al. 2023). Both axes are in units 10^{19} cm^{-2} . The values from the FT method are clearly below the 1:1 line, by amounts bounded by the other lines discussed in the text.

some spread. Some Gaussians are as broad as $\sigma = 3 \text{ km s}^{-1}$, leading to slope 0.05 (lowest line), the fiducial value discussed above, and some are narrower, a few as small as 1 km s^{-1} . These lines nicely bracket the 2D histogram in Figure 23.

At low $N_{\text{HI, CNM, ROHSA}}^*$ ($\lesssim 2 \times 10^{19} \text{ cm}^{-2}$), the plateau of low but finite $f_{\text{low}}^{0.12} N_{\text{HI}}^*$ (corresponding to the darker regions in Figure 22) arises from map areas with very small values of $f_{\text{low}}^{0.12}$ (typically below 0.05) from broad WNM emission and noise multiplied by significant N_{HI}^* from WNM gas.

6. Application to Mapped 21 cm Data for DHIGLS DF

The main science goal illustrated by the applications below is to search for regions where there has been a thermal phase transition with condensation of cold dense gas. Searches in the future will need to be done over larger and larger data sets and

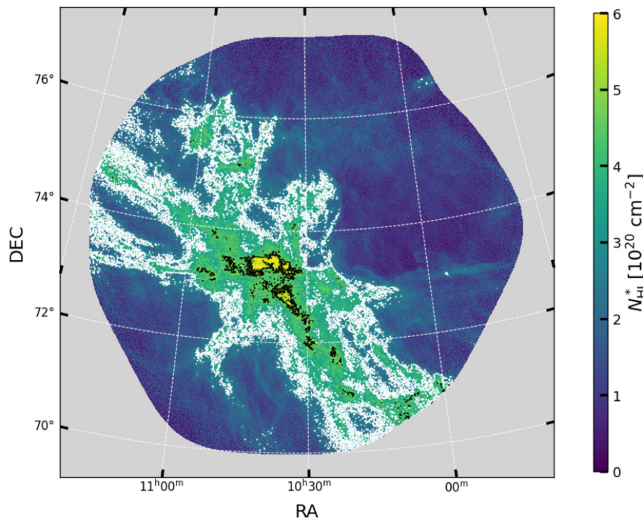


Figure 24. Total column density N_{HI}^* of the DF data set from DHIGLS in the velocity range $-15 < v < 15 \text{ km s}^{-1}$, computed in the optically thin limit. The solid white and black lines show the $N_{\text{HI}}^* = 3 \times 10^{20} \text{ cm}^{-2}$ and $N_{\text{HI}}^* = 5 \times 10^{20} \text{ cm}^{-2}$ contours.

in different spectral ranges and so require an efficient approach, such as we have developed in our FT-based method. Interesting regions found can then be targeted for analysis with more compute-intensive methods like ROHSA (Marchal et al. 2019) to quantify the properties of the multiphase medium.

We have used the FT method on data from GHIGLS and DHIGLS for a rapid evaluation of other regions and velocity ranges to study. The reader can explore this with the notebook supplied (footnote 6). Two applications are described below, here and in Section 7.

6.1. Data

The 57.4 square degrees DF data set used in the application here, located at $(\alpha, \delta) = (10^{\text{h}}30^{\text{m}}, 73^{\circ}48')$, was produced in the DHIGLS¹⁴ HI survey (Blagrove et al. 2017) with the Synthesis Telescope (ST) at the Dominion Radio Astrophysical Observatory. The 256-channel spectrometer, spacing $\Delta v = 0.824 \text{ km s}^{-1}$ sampling data with velocity resolution 1.32 km s^{-1} , was centered at $v_c = -60 \text{ km s}^{-1}$ relative to the local standard of rest (LSR). The spatial resolution of the ST interferometric data was approximately 0.9 . DF, which encompasses the Spider region (named for its prominent “legs” emanating from a central “body”: Marchal & Martin 2023 and references therein), is embedded in the North Celestial Pole Loop (NCPL) mosaic of the GHIGLS HI survey (Martin et al. 2015) with the Green Bank Telescope (GBT), with a spatial resolution of approximately $9'.4$. The DHIGLS DF product has the full range of spatial frequencies, obtained by a rigorous combination of the ST interferometric and GBT single dish data (see Section 5 in Blagrove et al. 2017). The pixel size is $18''$. A map of the total column density in the DF field in the velocity range $-15 < v < 15 \text{ km s}^{-1}$ is shown in Figure 24.

6.2. Apodization

Real measurements of the 21 cm line frequently include emission from intermediate-velocity clouds (IVCs) that overlap

the local velocity component. For example in the NCPL, there are also significant emission bridging velocities in between, which does not show the loop and has some characteristics of IVC, but was distinguished as a Bridge Velocity Component (BVC) by Taank et al. (2022) in their analysis of the multiphase structure of the loop. Building on this, we used T_b in the velocity range $-15 < v < 15 \text{ km s}^{-1}$, effectively setting T_b to zero outside of this range. Introducing this sharp edge (step function) in the signal produces a “ringing” in the Fourier-transformed spectrum (Gibbs phenomenon). However, we suppress this by applying a cosine apodization on four channels of the signal at both ends of the spectral range.

6.3. Denoising

Evaluation of the cold gas mass fraction using the method described in Section 2.4 is sensitive to noise. This can be appreciated in Figure 2 where noise dominates the signal at high $k_{v,\text{lim}}$ values in the amplitude spectrum of the FT of the interpolated T_b spectrum (black) of J2232 from 21-SPONGE. To overcome this limitation, we used the ROHSA algorithm (Marchal et al. 2019) to “denoise” the whole hyperspectral cube of the DF data set.

ROHSA is a regularized optimization algorithm that decomposes PPV data cubes into a sum of Gaussians (Marchal et al. 2019). ROHSA considers the spatial coherence of the emission and its multiphase nature to perform a separation of different thermal phases. Earlier applications of ROHSA were dedicated to mapping out thermal phases in 21 cm data (e.g., Marchal & Miville-Deschênes 2021; Marchal et al. 2021; Taank et al. 2022; Vujeva et al. 2023).

Here we simply made use of the spatial regularization to obtain a spatially coherent model of the data that concurrently reduces noise. The decomposition used in this work was obtained using $N=6$ Gaussians and the hyper-parameters $\lambda_a = \lambda_\mu = \lambda_\sigma = \lambda_{\sigma'} = 10$, which control the strength of the regularization that penalizes small-scale spatial fluctuations of each Gaussian parameter (amplitude, central velocity, dispersion), measured by the energy at high spatial frequencies ($\lambda_a = \lambda_\mu = \lambda_\sigma$), and the regularization of the variance of each velocity dispersion field ($\lambda_{\sigma'}$). This decomposition reproduces the original data with a good χ^2 and we refer to the PPV cube built from the model as “denoised data.”¹⁵

Maps of the gas column density, $f_{\text{low}}^{k_{v,\text{lim}}} N_{\text{HI}}^*$, as a function of $k_{v,\text{lim}}$ computed from the denoised data and the original data are shown in Figures A2 and A3, respectively, in Appendix A.2. At low $k_{v,\text{lim}}$, the maps are similar but the relative proportion of noise compared to the sky signal increases with $k_{v,\text{lim}}$ to a point where filamentary structures are hardly visible in Figure A3. Decomposition with ROHSA concurrently reduces noise at high $k_{v,\text{lim}}$, retaining spatial information about the very cold structures in the Spider in Figure A2.

Application of a denoising algorithm should be considered based on the maximum kinetic temperature one wishes to probe in the data following the methodology described in Section 2.4 pairing σ_{lim} and t_{lim} to motivate the choice of $k_{v,\text{lim}}$. Because the FT is computationally inexpensive (see Section 9) compared to denoising, we recommend the exploratory application of the FT

¹⁴ DRAO HI Intermediate Galactic Latitude Survey: <https://www.cita.utoronto.ca/DHIGLS/>.

¹⁵ This simple decomposition is good enough for denoising DF, and then applying the FT method. However, it has not undergone the rigorous analysis and testing needed to recommend it for thermal phase separation in this complex field.

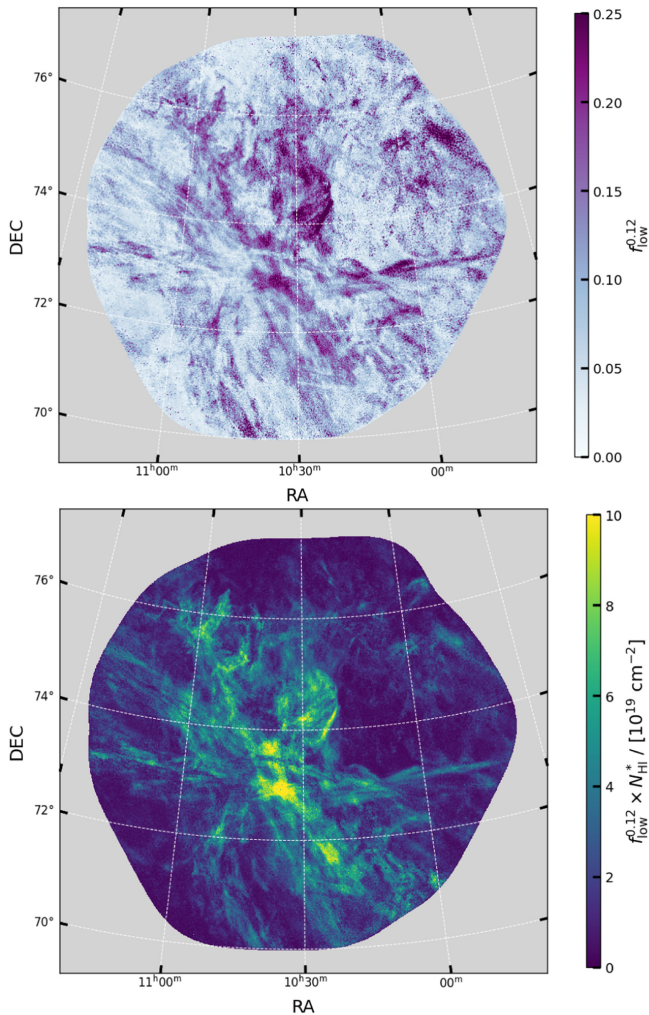


Figure 25. Lower limit on the cold gas mass fraction $f_{\text{low}}^{0.12}$ (top) and corresponding lower limit on the cold gas column density $f_{\text{low}}^{0.12} N_{\text{HI}}^*$ (bottom) of the DF data set from DHIGLS in the velocity range $-15 < v < 15 \text{ km s}^{-1}$.

method without denoising the data to judge the quality of $f_{\text{low}}^{k_{v,\text{lim}}}$ at the selected $k_{v,\text{lim}}$. In our application to Spider with $f_{\text{low}}^{k_{v,\text{lim}}} = 0.12$, denoising the data is not particularly beneficial for most of the map, except for areas such as the upper right where the column density is low or around the perimeter where the noise is larger. In those areas, the image is less noisy and the contrast between green structure and blue is enhanced.

6.4. Results

The top panel of Figure 25 shows $f_{\text{low}}^{0.12}$ (corresponding to $k_{v,\text{lim}} = 0.12 \text{ (km s}^{-1}\text{)}^{-1}$) in DF using the apodized denoised data. The bottom panel of Figure 25 shows the corresponding lower limit on the column density of cold gas, calculated by multiplying the total column density (Figure 24, computed in the optically thin limit) by the cold gas mass fraction (top panel), i.e., $f_{\text{low}}^{0.12} N_{\text{HI}}^*$. Both panels reveal structures attributable to cold gas in the Spider, arranged in a complex network of “filaments.”

Because $f_{\text{low}}^{0.12}$ is a lower limit, we cannot discuss the exact spatial distribution of the cold gas, even in projection let alone 3D. However, our tests in Section 3.3 *n* realistic simulations of HI gas show that information about the projected spatial structure is retained in the FT-based map. Here, applied to

actual data, it seems unlikely that the structure that appears filamentary in projection is created at random; rather it reveals an important structural property of the thermal condensation of the HI gas.

In the top panel, the lower limit $f_{\text{low}}^{0.12}$ varies considerably across the field, ranging from 0 (no CNM) up to 0.3 along the filamentary structures. In the bottom panel, the brightest part of the lower limit on cold gas column density is located near the core of the Spider identified in the total column density map, but the contrast is higher in this map and the structural details differ.

For completeness, we present lower-limit gas mass fraction maps and column density maps for the first nine $k_{v,\text{lim}}$ values obtained from the FT of the denoised data, in Figures A1 and A2, respectively.¹⁶ The PPK cube reveals increasingly colder structures as $k_{v,\text{lim}}$ increases and it can be appreciated visually that at high $k_{v,\text{lim}}$ there is more structure on small spatial scales. Note that we applied a Fourier transform only along the velocity axis of the PPV cube, with no explicit spatial filtering. The relationship between small spatial scale features and the intensity at high $k_{v,\text{lim}}$ can be understood as a direct consequence of the thermal condensation of HI gas.

The network of filaments seen in $f_{\text{low}}^{0.12} N_{\text{HI}}^*$ is similar to the filamentary structures revealed by scanning through the PPV cube of the original data. These structures were first illustrated by Blagrove et al. (2017) in their RGB color image of the HI emission using three distinct channels separated by only 2.47 km s^{-1} (top panel of their Figure 24). The dramatic changes in the filamentary structure of HI channel maps over a small velocity separation are captured at high k_v in the Fourier-transformed data, hence the similarities between $f_{\text{low}}^{0.12} N_{\text{HI}}^*$ and the structure of individual channel maps.

The FT approach and the channel maps provide complementary perspectives on the nature of filaments (also called fibers in the context of the diffuse sky at high latitude, Clark et al. 2014) that are believed to be dominated by cold density structures (Clark et al. 2019; Peek & Clark 2019; Murray et al. 2020).

This qualitative result is consistent with the recent spatial wavelet scattering transform (ST) analysis performed by Lei & Clark (2023). Using a combination of GALFA-HI data (Peek et al. 2011, 2018) and 21-SPONGE data, the authors found that ST-based metrics for small-scale linearity are predictive of the fraction of cold gas along the line of sight. It is noteworthy that the ST approach is based solely on analyzing the multiscale information of HI emission maps while our FT method is solely applied to the spectral axis of the data cube. This complementarity opens up an interesting avenue for combining both the spatial and spectral information that can be extracted from 21 cm data.

A comprehensive statistical characterization of the correlation between $f_{\text{low}}^{0.12} N_{\text{HI}}^*$ (or different $k_{v,\text{lim}}$) and the structure of individual channel maps is beyond the scope of this paper. This could be achieved using a cross-correlation analysis. Similarly, a thorough analysis of the whole PPK cube of DF is beyond the scope of this paper but in the future will provide valuable information about the multiphase and multiscale structure of HI gas in the Spider that is part of the NCPL.

¹⁶ Note that at high $k_{v,\text{lim}}$, denoising the data prior to the FT becomes critical, as is evident in Figure A3, which shows the same column density maps as Figure A2 but computed with the original data.

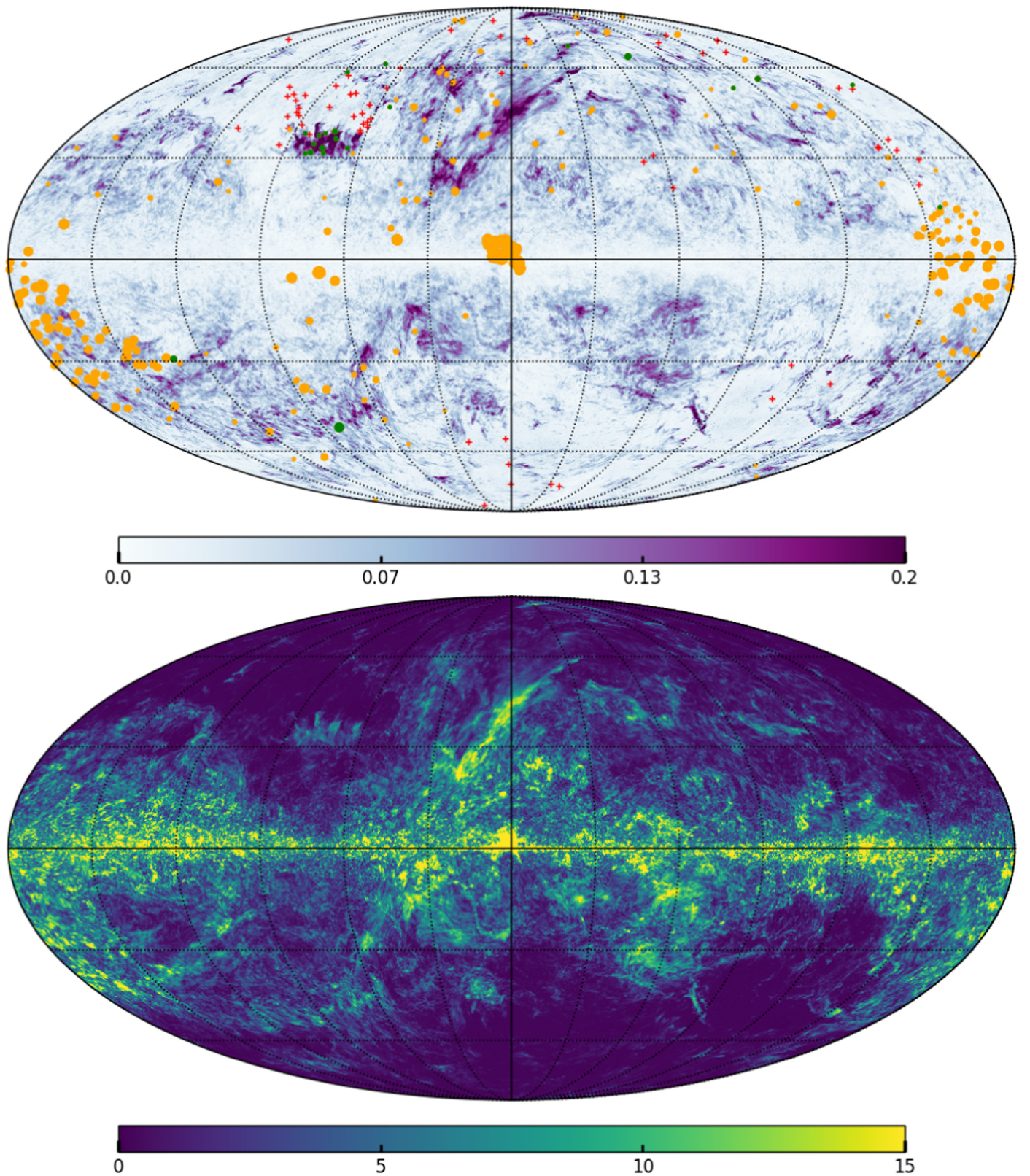


Figure 26. Mollweide projection centered on the Galactic center of the lower limit on the cold gas mass fraction $f_{\text{low, HI4PI}}^{0.12}$ (top), and corresponding lower limit on the cold gas column density $f_{\text{low, HI4PI}}^{0.12} N_{\text{HI}}^*$ in units of 10^{19} cm^{-2} (bottom) of the entire HI4PI data set in the velocity range $-90 < v < 90 \text{ km s}^{-1}$. Similar to the annotations in Figure 20, the red crosses in the top panel show the positions of nondetections in the compilation of H I absorption spectra from the BIGHICAT meta-catalog (McClure-Griffiths et al. 2023) and the orange dots, whose size encodes the number of absorption components (ranging from 1 to 16) in their τ spectrum, show the detections in BIGHICAT. The green dots show where detections in BIGHICAT have $f_{\text{BIGHICAT}} - f_{\text{low, HI4PI}}^{0.12} < 0$.

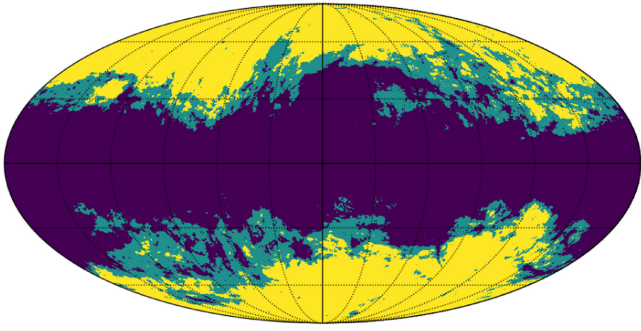


Figure 27. Mollweide projection centered on the Galactic center showing regions with $N_{\text{HI}}^* < 3 \times 10^{20} \text{ cm}^{-2}$ (yellow), $3 \times 10^{20} < N_{\text{HI}}^* < 5 \times 10^{20} \text{ cm}^{-2}$ (green), and $N_{\text{HI}}^* > 5 \times 10^{20} \text{ cm}^{-2}$ (blue) from the HI4PI survey over the velocity range $-90 < v < 90 \text{ km s}^{-1}$.

7. Application to HI4PI Data for the Whole Sky

Using the HI4PI data (Section 4.2), we computed $f_{\text{low, HI4PI}}^{0.12}$ in the velocity range $-90 < v < 90 \text{ km s}^{-1}$ over the full sky. No denoising was applied to the data set prior to the Fourier transform because the noise level of T_b is fairly uniformly low, approximately $\sigma_{\text{rms}} \sim 43 \text{ mK}$. Each spectrum was apodized with a cosine function at both ends of the spectral range.

Figure 26 shows Mollweide projections centered on the Galactic center of $f_{\text{low, HI4PI}}^{0.12}$ (top), and $f_{\text{low, HI4PI}}^{0.12} N_{\text{HI}}^*$ in units of 10^{19} cm^{-2} (bottom). Annotations showing results from BIGHICAT are similar to those in Figure 20.¹⁷ Comparison of the top map of $f_{\text{low, HI4PI}}^{0.12}$ with maps made with increasing values of $k_{v, \text{lim}}$ above 0.12 $(\text{km s}^{-1})^{-1}$ (Figure B1 top in Appendix B) reveals the same morphology but with amplitude decreasing with $k_{v, \text{lim}}$, showing that the features detecting CNM gas are real, not noise dominated.

At high latitudes, these full sky maps show a wide variety of known regions, e.g., the North Polar Spur (Tunmer 1958; Hanbury Brown et al. 1960; Das et al. 2020; West et al. 2021, and reference within), the compressed magnetized shells associated to the Corona Australis molecular cloud (Bracco et al. 2020), or the North Celestial Pole Loop (Meyerdierts et al. 1991; Taank et al. 2022; Marchal & Martin 2023). A comprehensive description is beyond the scope of this paper and will be explored in future works.

Toward the Galactic plane, $f_{\text{low, HI4PI}}^{0.12}$ is low, but due to the high total column density (computed in the optically thin limit), $f_{\text{low, HI4PI}}^{0.12} N_{\text{HI}}^*$ shows a variety of structures with column density comparable to that of the high latitude sky. At these latitudes, a sum of Gaussians as a model to describe the 21 cm line is not a good description of the data, due to the extreme effect of opacity and self-absorption. The Riegel–Crutcher cloud seen toward the Galactic center (Heeschen 1955; Riegel & Crutcher 1972; McClure-Griffiths et al. 2006) is a noteworthy example. HI self-absorption (HISA) seen in its emission spectrum (see, e.g., Figure 2 in McClure-Griffiths et al. 2006) is most likely to impact the amplitude structure of $\hat{T}_b(k_v)$. Narrow emission lines in the optically thin limit and HISA will both be highlighted at high k_v in the Fourier transform of the emission data but whether Equation (9) still provides a lower limit on the amount of cold gas remains unclear and should be investigated in future work. We therefore

¹⁷ For completeness, in Figure B1 we present the same two quantities but for nine $k_{v, \text{lim}}$ values sampled in the range $0.01 < k_v (\text{km s}^{-1})^{-1} < 0.27$.

recommend that $f_{\text{low, HI4PI}}^{0.12}$ at very low Galactic latitude be taken with caution, especially in places where HISA are known to impact strongly the emission lines.

Figure 27 shows masks of regions with $N_{\text{HI}}^* < 3 \times 10^{20} \text{ cm}^{-2}$ (yellow), $3 \times 10^{20} < N_{\text{HI}}^* < 5 \times 10^{20} \text{ cm}^{-2}$ (green), and $N_{\text{HI}}^* > 5 \times 10^{20} \text{ cm}^{-2}$ (blue). In the green and yellow areas (mainly the high latitude sky), where f_{BIGHICAT} and $f_{\text{low, HI4PI}}^{0.12}$ are fairly similar, our lower limit estimate tends toward the absolute value of cold gas mass fraction probed by absorption-line surveys (see discussion of Figure 21 in Section 4.4). In the blue areas, our lower limit should be treated as its definition, saying the cold gas mass fraction is at least $f_{\text{low, HI4PI}}^{0.12}$ but its absolute value remains unknown from this method. We recommend using these column density cuts to determine regions of the sky where $f_{\text{low, HI4PI}}^{0.12}$ tends toward the absolute value of the cold gas mass fraction and where $f_{\text{low, HI4PI}}^{0.12}$ should be strictly used as a lower limit.

8. Assessing $f_{\text{low}}^{0.12}$ in the High Latitude Sky Using a CNN

Hensley et al. (2022) applied the CNN model developed by Murray et al. (2020) to HI4PI data in the same velocity range, $-90 < v < 90 \text{ km s}^{-1}$, but only in the high latitude sky (i.e., $|b| > 30^\circ$), obtaining a map of $f_{\text{CNN}}^{\text{HI4PI}}$ in both Galactic hemispherical caps.

The two estimates of the cold mass gas fraction are compared in Figure 28, an orthographic projection of the difference $f_{\text{CNN}}^{\text{HI4PI}} - f_{\text{low, HI4PI}}^{0.12}$ in the northern (left) and southern (right) Galactic hemispheres. Figure 29 shows the 2D distribution function of the difference versus $\log N_{\text{HI}}^*$.

The two estimates are in fairly good agreement, with the mean and standard deviation of the difference $f_{\text{CNN}}^{\text{HI4PI}} - f_{\text{low, HI4PI}}^{0.12}$ being -0.0065 and 0.033 , respectively. However, there are coherent regions of positive and negative excursions around the mean, predominately in the range $30 < |b| < 60$. Positive values are predominantly in regions of higher column density, where the CNN, through its training, performs better at correcting for opacity effects (see Figure 8 in Murray et al. 2020).

Negative values are found in more localized clouds, prominent at, e.g., $(\ell, b) = (75^\circ, 35^\circ)$ (which corresponds to the region covered by the MACH survey and is discussed in Section 4.4) and $(\ell, b) = (135^\circ, 55^\circ)$. The emission of the first cloud is in the low-velocity range. The second cloud is a bright ($T_b \approx 26 \text{ K}$) IVC at approximately -50 km s^{-1} .¹⁸ The CNN and the FT methods have both been applied to the same emission spectra at $16''$ resolution and so unlike in the comparison between our lower limit and the BIGHICAT estimate based on absorption, this discrepancy cannot be explained by the beam of the observation.

We have investigated the IVC cloud, where the difference between the two methods is the most negative. Its column density map N_{HI}^* in the velocity range $-90 < v < 90 \text{ km s}^{-1}$ from HI4PI is shown in the left panel in Figure 30. Figure 30 (middle) shows a map of $f_{\text{low, HI4PI}}^{0.12}$. It has a structure similar to that of the total column density within the IVC cloud, a shape clearly recognized in Figure 28. We have extracted a spectrum at $(\ell, b) = (136^\circ.3, 54^\circ.7)$ in a bright substructure of the cloud (annotated by the red star in Figure 30) where $f_{\text{low, HI4PI}}^{0.12}$ is quite high (0.28). Figure 31 shows this T_b spectrum in light blue. We have decomposed the

¹⁸ A $4^\circ \times 4^\circ$ subset of this area was mapped in ICRS coordinates in the GHIGLS survey with the GBT at $9''/55$ resolution. There it is called the UM1 field, centered at $(\ell, b) = (134^\circ.95, 54^\circ.13)$ coincidentally near the peak of the IVC-dominated column density.

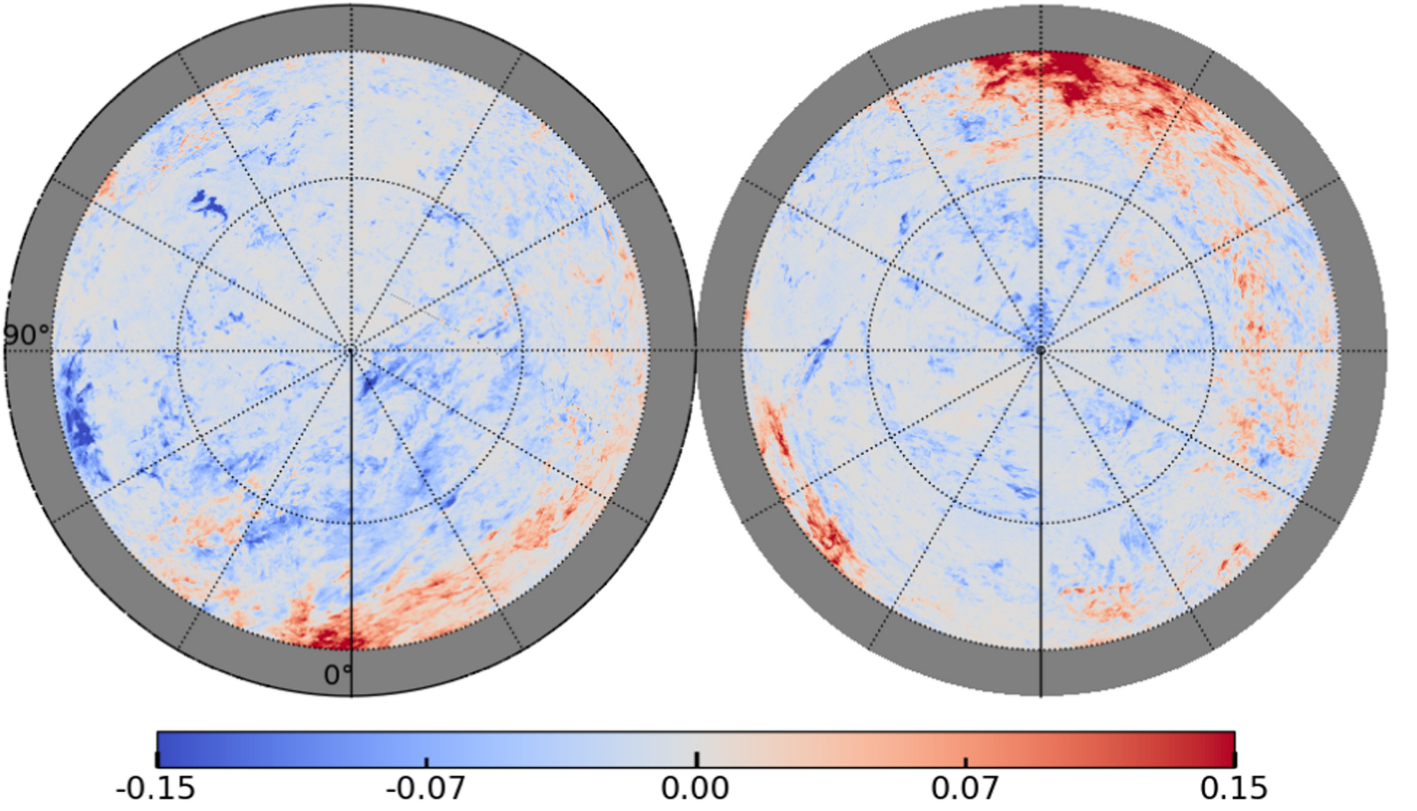


Figure 28. Orthographic projection of the northern (left) and southern (right) Galactic hemisphere of $f_{\text{CNN}}^{\text{HI4PI}} - f_{\text{low, HI4PI}}^{0.12}$. For orientation, $\ell = 0^\circ$ and $\ell = 90^\circ$ are annotated on the left panel.

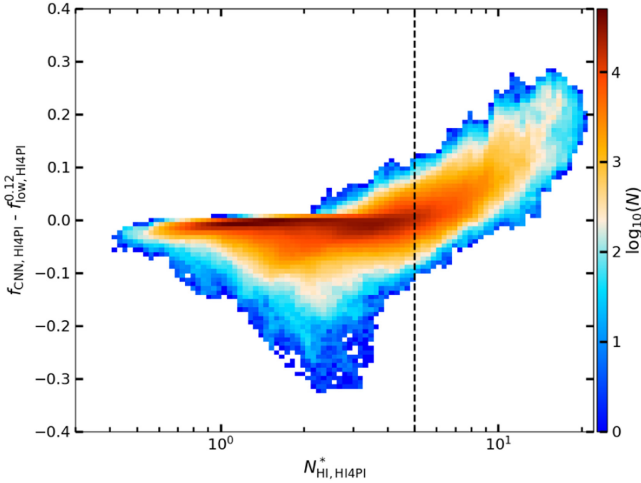


Figure 29. 2D distribution function of $f_{\text{CNN, HI4PI}}^{\text{HI4PI}} - f_{\text{low, HI4PI}}^{0.12}$ vs. $\log N_{\text{HI}}^*$ computed in the optically thin limit with HI4PI data. The dashed vertical line denotes the typical column density where the effects of opacity become significant, as in Figure 16.

spectral data with a Gaussian model, where the number of Gaussian ($N = 10$) was selected based on a reduced chi-square criterion. The total model is shown in black. Individual Gaussians with $\sigma < 3 \text{ km s}^{-1}$ ($\sigma > 3 \text{ km s}^{-1}$) are shown in blue (red). The gas mass fraction in the cold Gaussians (blue) evaluates to 0.67, higher than our lower limit $f_{\text{low, HI4PI}}^{0.12}$.

On the other hand, the map of the estimate $f_{\text{CNN}}^{\text{HI4PI}}$ (Figure 30, right) is overall very close to 0; while it also shows a structure of the same shape, its mean value is curiously negative,

approximately -0.04 , lower than that of the background of approximately 0.01. This failure of the CNN probably arises because the IVC-dominated spectra are far different than the training set used for the CNN. In support of this, if we translate the IVC peak in the spectrum in Figure 31 to be centered on 0 km s^{-1} , then the CNN method does detect CNM gas, with $f_{\text{CNN}}^{\text{HI4PI}} = 0.22$.

9. Discussion

The FT method presented in this work has disadvantages as well as advantages. By focusing on only the amplitude information of the PPK cube, the phase information is inevitably lost and only the average information about the amount of cold gas remains available to some extent. The impact of the velocity field on the amplitude of the FT (i.e., the interference patterns) constrains us from evaluating a lower limit on the mass fraction of cold gas.

However, this partial loss of information on the velocity field, as well as the absence of modeling in comparison to Gaussian decomposition algorithms, makes our approach highly efficient in terms of computing time and a unique tool to explore where the phase transition has taken place. Note that the computing time is not reduced in the case where ROHSA is used to denoise the data prior to the FT but, as discussed in Section 6.3 and Appendix A.2, denoising only becomes critical at high $k_{v, \text{lim}}$, and the cold gas mass fraction map inferred at $k_{v, \text{lim}} = 0.12 \text{ (km s}^{-1}\text{)}^{-1}$ using the original data is still practicable for further analysis.

As a reference, the PPK cube of the DF field with dimensions 60, 1680, and 1764 along the spectral and two spatial axes, respectively, was obtained in approximately a

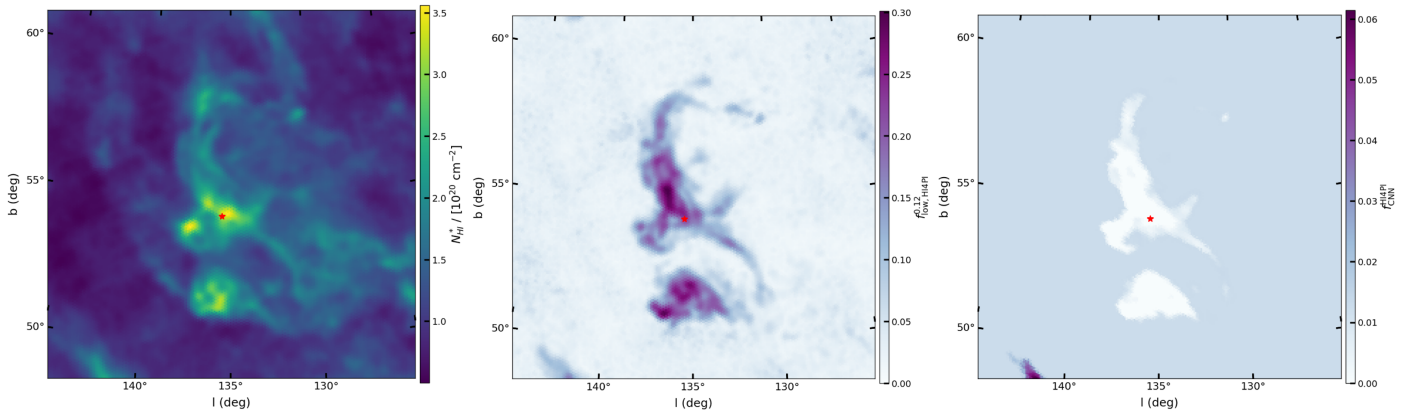


Figure 30. Left: Tangential projection in Galactic coordinates of N_{HI}^* in the velocity range $-90 < v < 90 \text{ km s}^{-1}$ from HI4PI, centered at $(\ell, b) = (135^\circ, 55^\circ)$. Middle: $f_{\text{low, HI4PI}}^{0.12}$. Right: $f_{\text{CNN}}^{\text{HI4PI}}$. The red star marks $(\ell, b) = (136.3, 54.7)$, the line of sight of the spectrum whose Gaussian decomposition is shown in Figure 31.

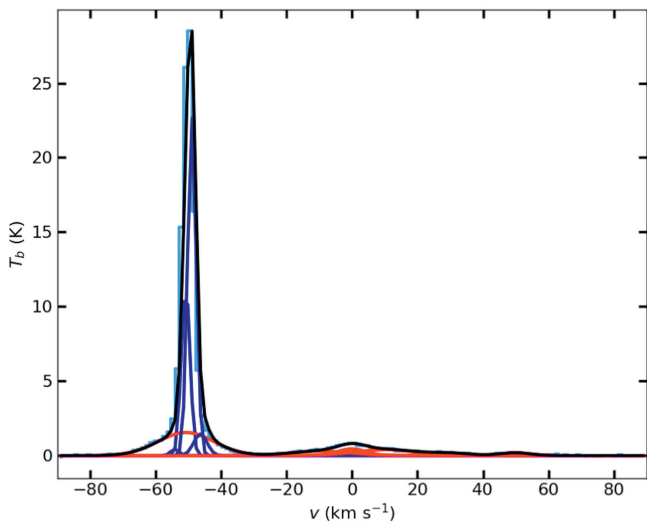


Figure 31. Brightness temperature spectrum extracted from HI4PI at $(\ell, b) = (136.3, 54.7)$, located by the red star in Figure 30. The data (light blue histogram) is well matched by the total model (black curve). Individual Gaussians with $\sigma < 3 \text{ km s}^{-1}$ ($\sigma > 3 \text{ km s}^{-1}$) are shown in blue (red). The gas mass fraction in cold Gaussians evaluates to 0.63 and our lower limit estimate $f_{\text{low, HI4PI}}^{0.12} = 0.28$.

minute on a single CPU using the real fast Fourier transform algorithm implemented in NumPy. The lower resolution full sky HI4PI data were Fourier transformed in approximately 10 minutes on a single CPU. The computation timescales linearly with the number of pixels because each line of sight is treated independently while applying the FT. This offers the possibility of straightforward parallelization on multiple CPUs. Furthermore, almost no additional memory is required as the operation can be performed independently for each spectrum. Both the low computing time and memory make this method easily applicable to massive data sets, pointing the way to new approaches suitable for the new generation of radio interferometers like the Australian Square Kilometer Pathfinder (ASKAP; Dickey et al. 2013; Pingel et al. 2022), the Square Kilometer Array (SKA; de Blok et al. 2015; McClure-Griffiths et al. 2015), as well as the Next Generation Very Large Array (ngVLA; Pisano et al. 2018).

To overcome the loss of information on the velocity field and fully probe the fraction of cold gas as a function of the velocity, it is necessary to go beyond the statistics encoded in the

amplitude of the FT. Future research ought to examine the short-time Fourier transform (STFT) and/or wavelet transform, which extract information from the k_v domain while preserving information in the time domain (i.e., here the velocity domain), to facilitate progress on understanding the multiphase (multi-scale in terms of velocity) nature of the medium traced by 21 cm data.

10. Summary

We develop a new method for spatially mapping a lower limit on the mass fraction of the cold neutral medium by analyzing the amplitude structure of $\hat{T}_b(k_v)$, the Fourier transform of $T_b(v)$, the spectrum of the brightness temperature of HI 21 cm line emission with respect to the radial velocity v . This development is supported by multiple experiments including a basic understanding of the FT spectrum using toy models, evaluation on a realistic simulation of the neutral ISM, validation using absorption-line surveys, and application to real data. The main conclusions are as follows.

1. Using toy models, we illustrate the origin of interference patterns seen in $\hat{T}_b(k_v)$. Building on this, a lower limit on the cold gas mass fraction is obtained from the amplitude of \hat{T}_b at high k_v .
2. Tested on a numerical simulation of thermally bi-stable turbulence, the lower limit from this method has a strong linear correlation with the “true” cold gas mass fraction from the simulation for relatively low cold gas mass fraction. At a higher mass fraction, our lower limit is lower than the “true” value likely due to a combination of interference and opacity effects.
3. Comparison with absorption surveys shows that our lower limit is close to the absolute value obtained from a combination of emission and absorption data for a total column density along the line of sight $N_{\text{HI}} \lesssim 3\text{--}5 \times 10^{20} \text{ cm}^{-2}$, with a departure from linear correlation at higher column densities also due to a combination of interference and opacity effects.
4. Application to the DRAO Deep Field (DF) from DHIGLS reveals a complex network of filaments in the Spider, an important structural property of the thermal condensation of the HI gas. Our estimator $f_{\text{low}}^{0.12}$ varies considerably across the field, ranging from 0 (no CNM) up to 0.3 along the filamentary structures. The brightest parts in the map of the lower limit on the cold gas column

density are located near the core of the Spider identified in the total column density map.

5. Application to the HI4PI survey in the velocity range $-90 < v < 90 \text{ km s}^{-1}$ provides a full sky map of a lower limit on the mass fraction of the cold neutral medium that is fairly consistent with results from the CNN model produced by Murray et al. (2020).

Although highly efficient in terms of computing time, the main limitation of our FT approach is the loss of information encoded in the velocity field. Future research ought to examine the Short-Time Fourier Transform and/or wavelet transform, which extract information from the k_v domain while preserving information in the velocity domain.

Acknowledgments

We acknowledge support from the Natural Sciences and Engineering Research Council (NSERC) of Canada. A.B. acknowledges support from the European Research Council through the Advanced Grant MIST (FP7/2017-2022, No.742719). This research has made use of the NASA Astrophysics Data System. This work began under the program ‘‘Grand Cascade’’ organized and hosted by Institut Pascal at Universit  Paris-Saclay and the Interstellar Institute. We appreciate enlightening conversations with J.E.G. Peek, C. Murray, N. Pingel, and other members of the Interstellar Institute. We appreciate enlightening conversations with Van

Hiep Nguyen. We thank the anonymous referee whose comments and suggestions have improved this manuscript.

Software: Matplotlib (Hunter 2007), NumPy (van der Walt et al. 2011).

Appendix A DHIGLS DF

A.1. Multiple $k_{v,\text{lim}}$ Limits

We present in Figures A1 and A2, respectively, the gas mass fraction maps and column density maps for the first nine $k_{v,\text{lim}}$ values obtained from the FT of the denoised data of the DF field from DHIGLS.

A.2. Denoising the Data Prior to the FT

To provide a baseline for evaluating the impact of the denoising of the DF data using ROHSA on the PPK cube, we computed the FT of the original data. Figure A3 shows the column density $f_{\text{low}}^{k_{v,\text{lim}}} N_{\text{H I}}^*$ as a function of $k_{v,\text{lim}}$. This can be compared panel by panel with Figure A2 for the denoised data, presented with consistent color bars.

At low $k_{v,\text{lim}}$, Figures A2 and A3 are very similar. However, proceeding toward higher $k_{v,\text{lim}}$, the maps in Figure A3 become increasingly noisier, up to a point where the filamentary structures are hard to distinguish from the ambient noise (see especially the map at the highest $k_{v,\text{lim}}$).

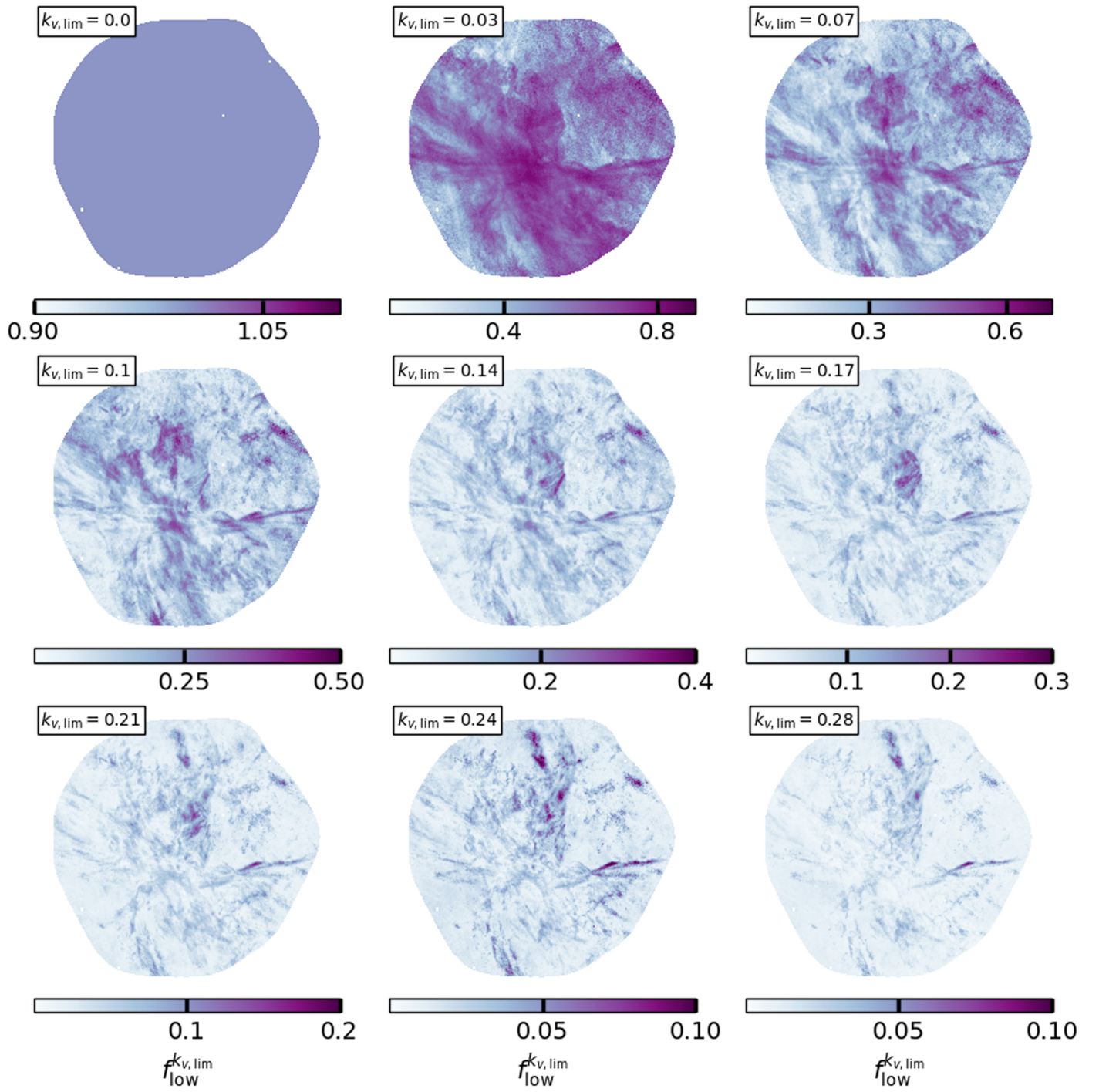


Figure A1. Gas mass fraction $f_{\text{low}}^{k_v, \text{lim}}$ as a function of $k_{v, \text{lim}}$ computed from the denoised data of the DF field from DHIGLS.

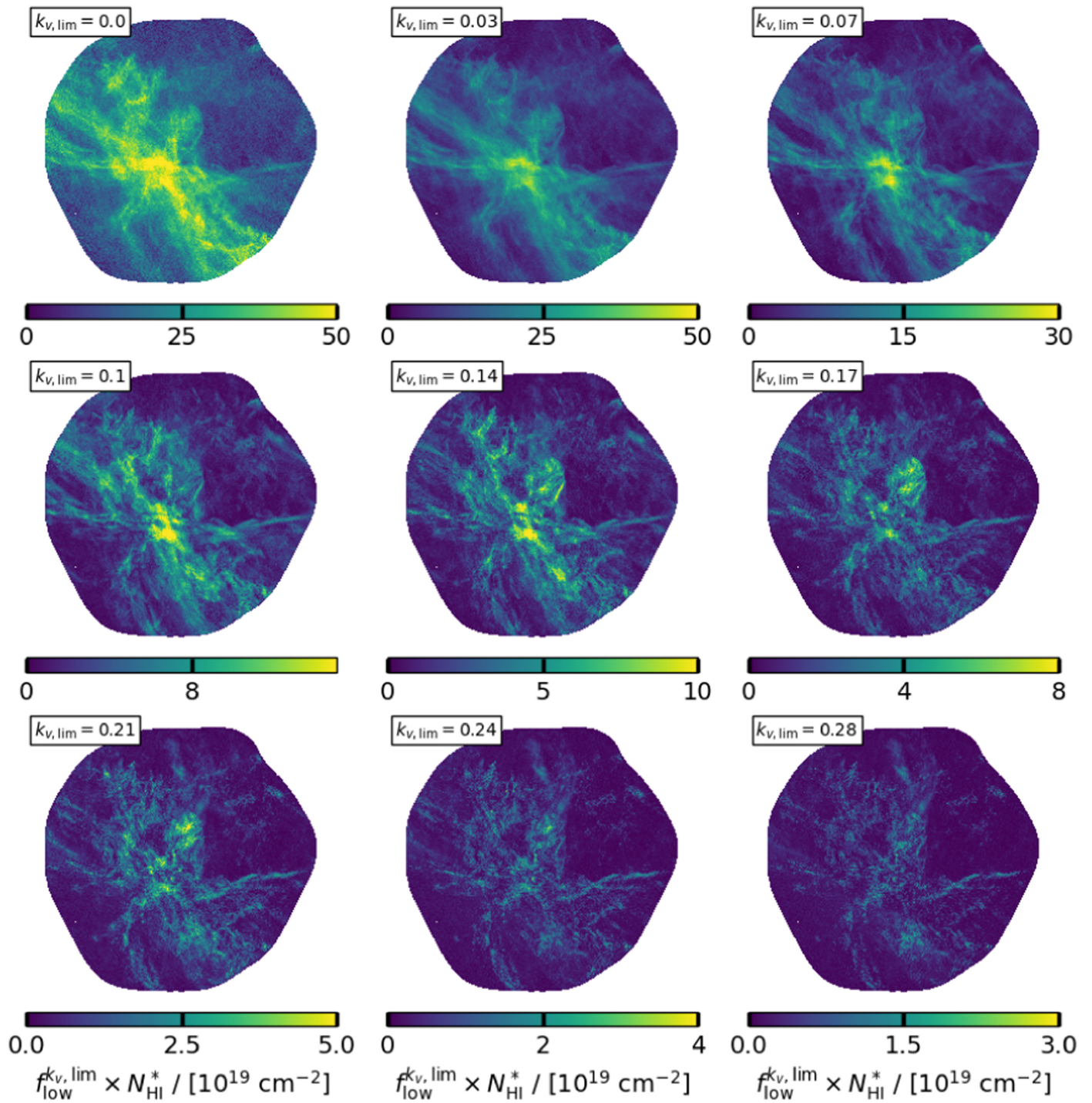


Figure A2. Column density $f_{\text{low}}^{k_{v, \text{lim}}} N_{\text{HI}}^*$ as a function of $k_{v, \text{lim}}$ computed from the denoised data of the DF field from DHIGLS.

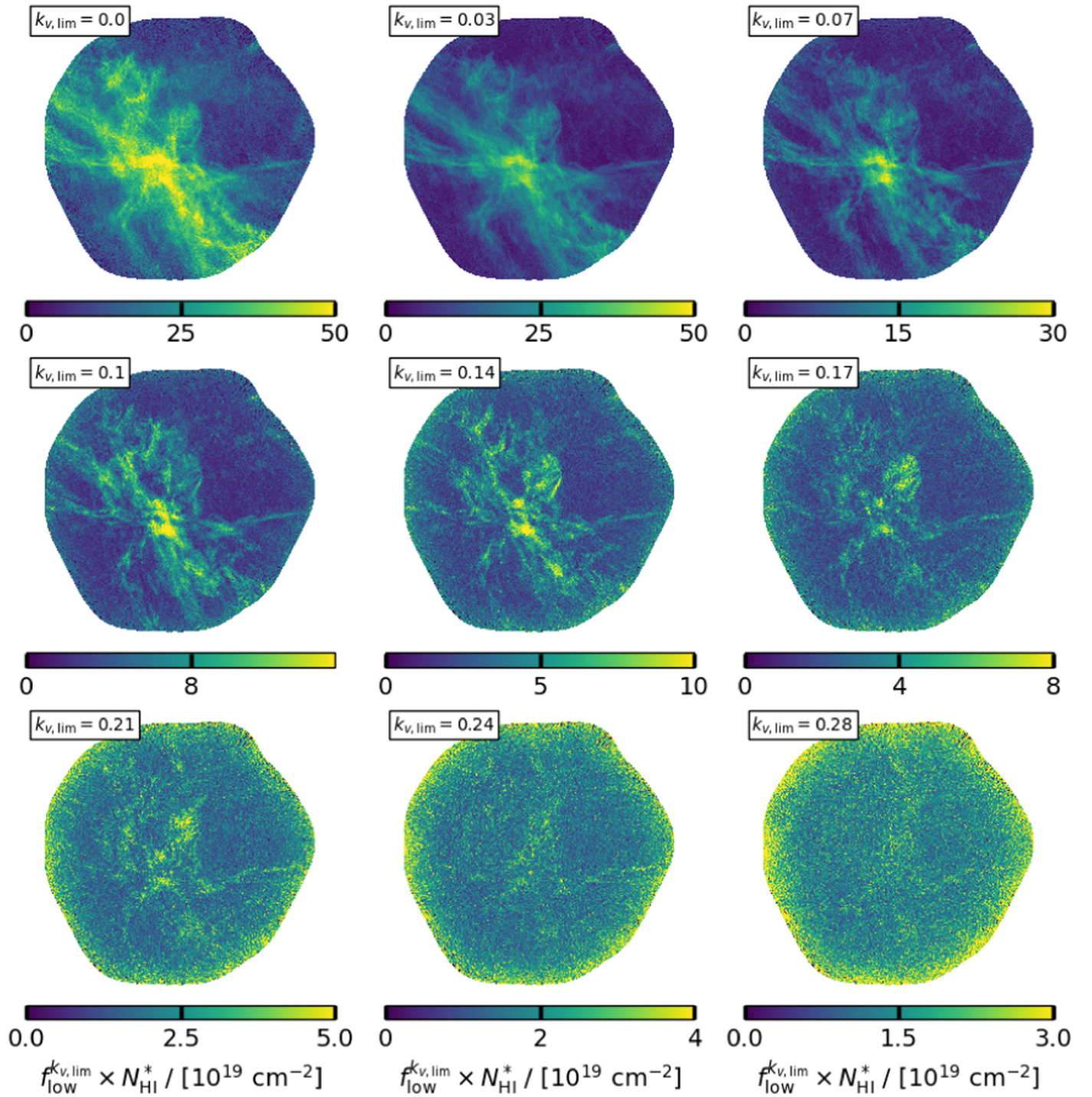


Figure A3. Column density $f_{\text{low}}^{k_{v,\text{lim}}} N_{\text{HI}}^*$ as a function of $k_{v,\text{lim}}$ computed from the original data (no denoising) of the DF field from DHIGLS.

Appendix B HI4PI

In Figure B1 we present the same two quantities but for nine $k_{v,\text{lim}}$ values sampled in the range $0.01 < k_v (\text{km s}^{-1})^{-1} < 0.27$.

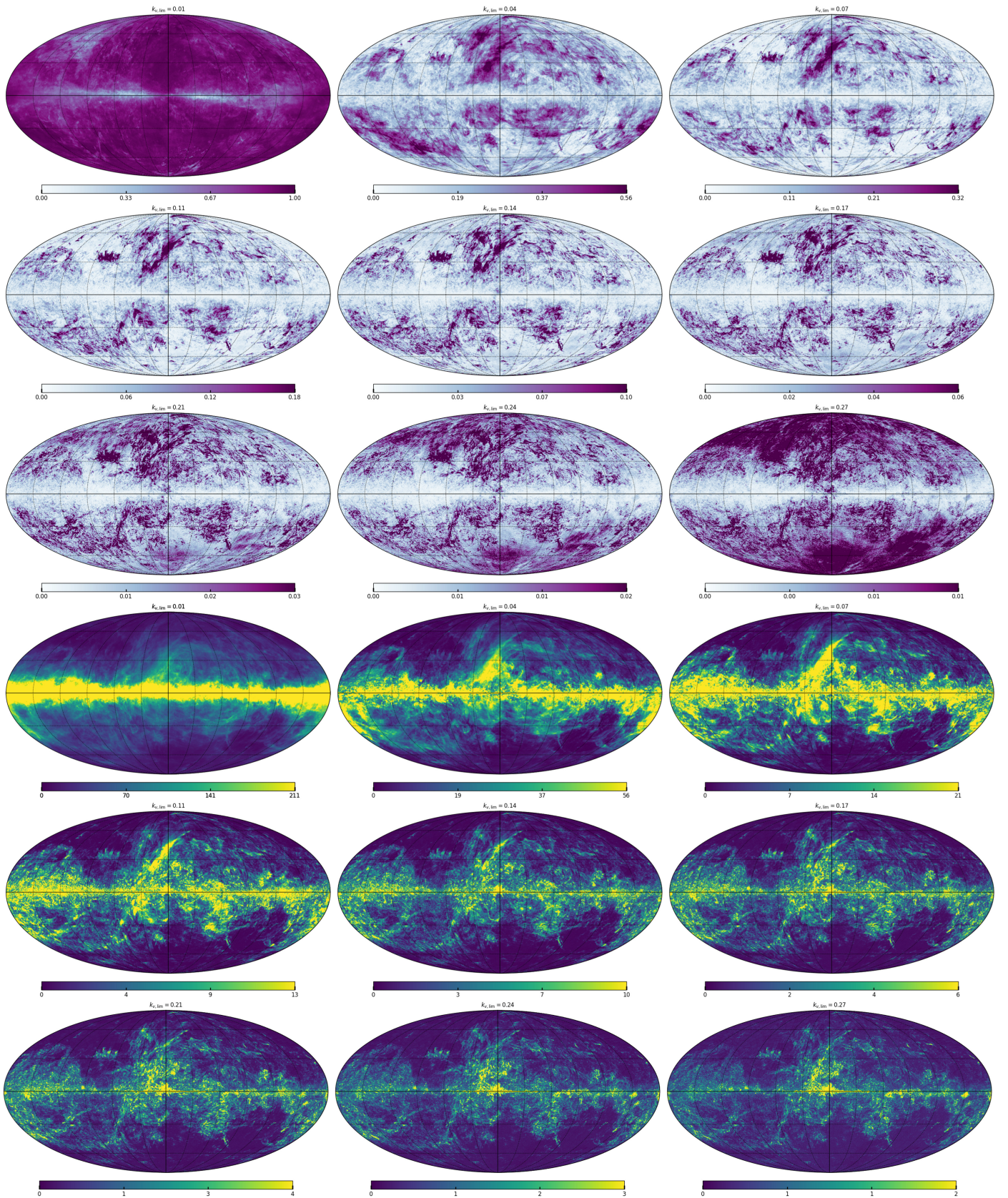









Figure B1. Mollweide projection centered on the Galactic center of the lower limit on the cold gas mass fraction $f_{\text{low, HI4PI}}^{k_v, \text{lim}}$ (top), and the lower limit on the cold gas column density $J_{\text{low, HI4PI}}^{k_v, \text{lim}} N_{\text{HI}}^*$ in units of 10^{19} cm^{-2} (bottom) of the entire HI4PI data set in the velocity range $-90 < v < 90 \text{ km s}^{-1}$.

ORCID iDs

Antoine Marchal  <https://orcid.org/0000-0002-5501-232X>
 Peter G. Martin  <https://orcid.org/0000-0002-5236-3896>
 Marc-Antoine Miville-Deschênes  <https://orcid.org/0000-0002-7351-6062>
 Naomi M. McClure-Griffiths  <https://orcid.org/0000-0003-2730-957X>
 Callum Lynn  <https://orcid.org/0000-0001-6846-5347>
 Andrea Bracco  <https://orcid.org/0000-0003-0932-3140>
 Luka Vujeva  <https://orcid.org/0000-0001-7697-8361>

References

- Audit, E., & Hennebelle, P. 2005, *A&A*, **433**, 1
 Beuther, H., Bihr, S., Rugel, M., et al. 2016, *A&A*, **595**, A32
 Bialy, S., & Sternberg, A. 2019, *ApJ*, **881**, 160
 Blagrove, K., Martin, P. G., Joncas, G., et al. 2017, *ApJ*, **834**, 126
 Bracco, A., Bresnahan, D., Palmeirim, P., et al. 2020, *A&A*, **644**, A5
 Burkhardt, B., Lee, M.-Y., Murray, C. E., & Stanimirović, S. 2015, *ApJL*, **811**, L28
 Chepurmov, A., Burkhardt, B., Lazarian, A., & Stanimirović, S. 2015, *ApJ*, **810**, 33
 Chepurmov, A., & Lazarian, A. 2006, arXiv:astro-ph/0611465
 Chepurmov, A., Lazarian, A., Stanimirović, S., Heiles, C., & Peek, J. E. G. 2010, *ApJ*, **714**, 1398
 Clark, S. E., Peek, J. E. G., & Miville-Deschênes, M. A. 2019, *ApJ*, **874**, 171
 Clark, S. E., Peek, J. E. G., & Putman, M. E. 2014, *ApJ*, **789**, 82
 Das, K. K., Zucker, C., Speagle, J. S., et al. 2020, *MNRAS*, **498**, 5863
 de Blok, E., Fraternali, F., Heald, G., et al. 2015, in *Advancing Astrophysics with the Square Kilometre Array (AASKA14)* (Trieste: SISSA), 129
 Dénes, H., McClure-Griffiths, N. M., Dickey, J. M., Dawson, J. R., & Murray, C. E. 2018, *MNRAS*, **479**, 1465
 Dickey, J. M., McClure-Griffiths, N., Gibson, S. J., et al. 2013, *PASA*, **30**, e003
 Draine, B. T. 1978, *ApJS*, **36**, 595
 Field, G. B. 1965, *ApJ*, **142**, 531
 Field, G. B., Goldsmith, D. W., & Habing, H. J. 1969, *ApJL*, **155**, L149
 González, M., Audit, E., & Huynh, P. 2007, *A&A*, **464**, 429
 Habing, H. J. 1968, *BAN*, **19**, 421
 Hanbury Brown, R., Davies, R. D., & Hazard, C. 1960, *Obs*, **80**, 191
 Haud, U. 2000, *A&A*, **364**, 83
 Heesch, D. S. 1955, *ApJ*, **121**, 569
 Heiles, C., & Troland, T. H. 2003, *ApJ*, **586**, 1067
 Hennebelle, P., & Pérault, M. 1999, *A&A*, **351**, 309
 Hennebelle, P., & Pérault, M. 2000, *A&A*, **359**, 1124
 Hensley, B. S., Murray, C. E., & Dodici, M. 2022, *ApJ*, **929**, 23
 HI4PI Collaboration, Ben Bekhti, N., Flöer, L., et al. 2016, *A&A*, **594**, A116
 Hunter, J. D. 2007, *CSE*, **9**, 90
 Kalberla, P. M. W., Burton, W. B., Hartmann, D., et al. 2005, *A&A*, **440**, 775
 Kalberla, P. M. W., & Haud, U. 2015, *A&A*, **578**, A78
 Kalberla, P. M. W., & Haud, U. 2018, *A&A*, **619**, A58
 Kalberla, P. M. W., Kerp, J., & Haud, U. 2020, *A&A*, **639**, A26
 Kalberla, P. M. W., McClure-Griffiths, N. M., Pisano, D. J., et al. 2010, *A&A*, **521**, A17
 Kerp, J., Winkel, B., Ben Bekhti, N., Flöer, L., & Kalberla, P. M. W. 2011, *AN*, **332**, 637
 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
 Kim, S., Staveley-Smith, L., Dopita, M. A., et al. 1998, *ApJ*, **503**, 674
 Koch, E. W., Rosolowsky, E. W., Lockman, F. J., et al. 2018, *MNRAS*, **479**, 2505
 Lazarian, A. 2009, *SSRv*, **143**, 357
 Lazarian, A., & Pogosyan, D. 2000, *ApJ*, **537**, 720
 Lazarian, A., & Pogosyan, D. 2006, *ApJ*, **652**, 1348
 Lazarian, A., & Pogosyan, D. 2008, *ApJ*, **686**, 350
 Lee, M.-Y., Stanimirović, S., Douglas, K. A., et al. 2012, *ApJ*, **748**, 75
 Lee, M.-Y., Stanimirović, S., Murray, C. E., Heiles, C., & Miller, J. 2015, *ApJ*, **809**, 56
 Lei, M., & Clark, S. E. 2023, *ApJ*, **947**, 74
 Lindner, R. R., Vera-Ciro, C., Murray, C. E., et al. 2015, *AJ*, **149**, 138
 Liszt, H. 2001, *A&A*, **371**, 698
 Marchal, A., & Martin, P. G. 2023, *ApJ*, **942**, 70
 Marchal, A., Martin, P. G., & Gong, M. 2021, *ApJ*, **921**, 11
 Marchal, A., & Miville-Deschênes, M.-A. 2021, *ApJ*, **908**, 186
 Marchal, A., Miville-Deschênes, M.-A., Orioux, F., et al. 2019, *A&A*, **626**, A101
 Martin, P. G., Blagrove, K. P. M., Lockman, F. J., et al. 2015, *ApJ*, **809**, 153
 McClure-Griffiths, N. M., Dickey, J. M., Gaensler, B. M., Green, A. J., & Haverkorn, M. 2006, *ApJ*, **652**, 1339
 McClure-Griffiths, N. M., Pisano, D. J., Calabretta, M. R., et al. 2009, *ApJS*, **181**, 398
 McClure-Griffiths, N. M., Stanimirović, S., Murray, C., et al. 2015, in *Advancing Astrophysics with the Square Kilometre Array (AASKA14)* (Trieste: SISSA), 130
 McClure-Griffiths, N. M., Stanimirović, S., & Rybarczyk, D. R. 2023, *ARA&A*, 61
 Meyerdieks, H., Heithausen, A., & Reif, K. 1991, *A&A*, **245**, 247
 Mohan, R., Dwarakanath, K. S., & Srinivasan, G. 2004a, *JApA*, **25**, 143
 Mohan, R., Dwarakanath, K. S., & Srinivasan, G. 2004b, *JApA*, **25**, 185
 Murray, C. E., Lindner, R. R., Stanimirović, S., et al. 2014, *ApJL*, **781**, L41
 Murray, C. E., Peek, J. E. G., & Kim, C.-G. 2020, *ApJ*, **899**, 15
 Murray, C. E., Stanimirović, S., Goss, W. M., et al. 2015, *ApJ*, **804**, 89
 Murray, C. E., Stanimirović, S., Goss, W. M., et al. 2018, *ApJS*, **238**, 14
 Murray, C. E., Stanimirović, S., Heiles, C., et al. 2021, *ApJS*, **256**, 37
 Nguyen, H., Dawson, J. R., Lee, M.-Y., et al. 2019, *ApJ*, **880**, 141
 Nguyen, H., Dawson, J. R., Miville-Deschênes, M. A., et al. 2018, *ApJ*, **862**, 49
 Nidever, D. L., Majewski, S. R., & Butler Burton, W. 2008, *ApJ*, **679**, 432
 Peek, J. E. G., Babler, B. L., Zheng, Y., et al. 2018, *ApJS*, **234**, 2
 Peek, J. E. G., & Clark, S. E. 2019, *ApJL*, **886**, L13
 Peek, J. E. G., Heiles, C., Douglas, K. A., et al. 2011, *ApJS*, **194**, 20
 Pingel, N. M., Dempsey, J., McClure-Griffiths, N. M., et al. 2022, *PASA*, **39**, e005
 Pingel, N. M., Lee, M.-Y., Burkhardt, B., & Stanimirović, S. 2018, *ApJ*, **856**, 136
 Pisano, D., Walter, F., & Stanimirović, S. 2018, in *ASP Conf. Ser. 517, Science with a Next Generation Very Large Array*, ed. E. Murphy (San Francisco, CA: ASP), 471
 Riegel, K. W., & Crutcher, R. M. 1972, *A&A*, **18**, 55
 Riener, M., Kainulainen, J., Henshaw, J. D., et al. 2019, *A&A*, **628**, A78
 Roy, N., Kanekar, N., Braun, R., & Chengalur, J. N. 2013, *MNRAS*, **436**, 2352
 Rybarczyk, D. R., Stanimirović, S., Gong, M., et al. 2022, *ApJ*, **928**, 79
 Saury, E., Miville-Deschênes, M. A., Hennebelle, P., Audit, E., & Schmidt, W. 2014, *A&A*, **567**, A16
 Seifried, D., Beuther, H., Walch, S., et al. 2022, *MNRAS*, **512**, 4765
 Stanimirović, S., Murray, C. E., Lee, M.-Y., Heiles, C., & Miller, J. 2014, *ApJ*, **793**, 132
 Stil, J. M., Taylor, A. R., Dickey, J. M., et al. 2006, *AJ*, **132**, 1158
 Syed, J., Wang, Y., Beuther, H., et al. 2020, *A&A*, **642**, A68
 Taank, M., Marchal, A., Martin, P. G., & Vujeva, L. 2022, *ApJ*, **937**, 81
 Taylor, A. R., Gibson, S. J., Peracaula, M., et al. 2003, *AJ*, **125**, 3145
 Tunner, H. 1958, *PMag*, **3**, 370
 van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, *CSE*, **13**, 22
 Vujeva, L., Marchal, A., Martin, P. G., & Taank, M. 2023, *ApJ*, **951**, 120
 Walter, F., Brinks, E., de Blok, W. J. G., et al. 2008, *AJ*, **136**, 2563
 Wang, Y., Beuther, H., Rugel, M. R., et al. 2020a, *A&A*, **634**, A83
 Wang, Y., Bihr, S., Beuther, H., et al. 2020b, *A&A*, **634**, A139
 West, J. L., Landecker, T. L., Gaensler, B. M., Jaffe, T., & Hill, A. S. 2021, *ApJ*, **923**, 58
 Winkel, B., Kerp, J., Flöer, L., et al. 2016, *A&A*, **585**, A41
 Wolfire, M. G., Hollenbach, D., McKee, C. F., Tielens, A. G. G. M., & Bakes, E. L. O. 1995, *ApJ*, **443**, 152
 Wolfire, M. G., McKee, C. F., Hollenbach, D., & Tielens, A. G. G. M. 2003, *ApJ*, **587**, 278