

## The GBT Diffuse Ionized Gas Survey: Tracing the Diffuse Ionized Gas around the Giant HII Region W43

MATTEO LUISI,<sup>1,2</sup> L. D. ANDERSON,<sup>1,2,3</sup> BIN LIU,<sup>4</sup> DANA S. BALSER,<sup>5</sup> T. M. BANIA,<sup>6</sup> TREY V. WENGER,<sup>7</sup> AND L. M. HAFFNER<sup>8</sup>

<sup>1</sup>*Department of Physics and Astronomy, West Virginia University, Morgantown WV 26506, USA*

<sup>2</sup>*Center for Gravitational Waves and Cosmology, West Virginia University, Chestnut Ridge Research Building, Morgantown WV 26505, USA*

<sup>3</sup>*Adjunct Astronomer at the Green Bank Observatory, P.O. Box 2, Green Bank WV 24944, USA*

<sup>4</sup>*CAS Key Laboratory of FAST, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, People's Republic of China*

<sup>5</sup>*National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville VA 22903-2475, USA*

<sup>6</sup>*Institute for Astrophysical Research, Department of Astronomy, Boston University, 725 Commonwealth Ave., Boston MA 02215, USA*

<sup>7</sup>*Dominion Radio Astrophysical Observatory, Herzberg Astronomy and Astrophysics Research Centre, National Research Council Canada, P.O. Box 248, Penticton, BC V2A 6J9, Canada*

<sup>8</sup>*Department of Physical Sciences, Embry-Riddle Aeronautical University, Daytona Beach FL 32114, USA*

### ABSTRACT

The Green Bank Telescope (GBT) Diffuse Ionized Gas Survey (GDIGS) is a fully-sampled radio recombination line (RRL) survey of the inner Galaxy at C-band (4 – 8 GHz). We average together  $\sim 15$  H $\alpha$  RRLs within the receiver bandpass to improve the spectral signal-to-noise ratio. The average beam size for the RRL observations at these frequencies is  $\sim 2'$ . We grid these data to have spatial and velocity spacings of  $30''$  and  $0.5 \text{ km s}^{-1}$ , respectively. Here we discuss the first RRL data from GDIGS: a six square-degree-area surrounding the Galactic HII region complex W43. We attempt to create a map devoid of emission from discrete HII regions and detect RRL emission from the diffuse ionized gas (DIG) across nearly the entire mapped area. We estimate the intensity of the DIG emission by a simple empirical model, taking only the HII region locations, angular sizes, and RRL intensities into account. The DIG emission is predominantly found at two distinct velocities:  $\sim 40 \text{ km s}^{-1}$  and  $\sim 100 \text{ km s}^{-1}$ . While the  $100 \text{ km s}^{-1}$  component is associated with W43 at a distance of  $\sim 6 \text{ kpc}$ , the origin of the  $40 \text{ km s}^{-1}$  component is less clear. Since the distribution of the  $40 \text{ km s}^{-1}$  emission cannot be adequately explained by ionizing sources at the same velocity, we hypothesize that the plasma at the two velocity components is interacting, placing the  $40 \text{ km s}^{-1}$  DIG at a similar distance as the  $100 \text{ km s}^{-1}$  emission. We find a correlation between dust temperature and integrated RRL intensity, suggesting that the same radiation field that heats the dust also maintains the ionization of the DIG.

*Keywords:* HII regions — ISM: abundances — ISM: bubbles — photon-dominated region (PDR) — radio lines: ISM

### 1. INTRODUCTION

The warm ionized medium (WIM), is a low-density ( $\sim 0.1 \text{ cm}^{-3}$ ), diffuse ionized gas component of the interstellar medium (ISM), first proposed by Hoyle & Ellis (1963). It makes up  $\sim 20\%$  of the total Milky Way gas mass and  $>90\%$  of its ionized gas mass (Reynolds 1991). With temperatures ranging from 6000 to 10,000 K (Haffner et al. 2009) and a hydrogen ionization ratio  $n(\text{H}^+)/n(\text{H}^0) \geq 13$  at the upper end of this temperature range (Reynolds et al. 1998), it provides a major source of pressure at the Galactic midplane (Boulares & Cox 1990). The “Diffuse Ionized Gas” (DIG) is often quoted in the same context as the WIM, with the WIM some-

times referring to the lower-density gas that cannot be traced to individual source regions.

Most WIM/DIG studies to date have been conducted by observing H $\alpha$  emission. Although H $\alpha$  is very bright compared to other ionized gas tracers, it suffers from extinction, limiting the distance to which the DIG in the inner Galaxy ( $90^\circ > \ell > -90^\circ$ ) can be studied. Photometric H $\alpha$  surveys of the DIG include the Southern H-Alpha Sky Survey Atlas (SHASSA; Gaustad et al. 2001) and the Virginia Tech Spectral-Line Survey (VTSS; Dennison et al. 1998). With its coarse spatial and spectral resolution, the all-sky Wisconsin H-Alpha Mapper (WHAM; Reynolds et al. 1998) provides a clear view

of the global properties of the DIG, but is inappropriate for studying the inner Galaxy mid-plane where the majority of massive stars are located. Using the Near Infrared Camera and Multi-Object Spectrometer (NICMOS) instrument aboard the Hubble Space Telescope, [Wang et al. \(2010\)](#) report on a high-resolution photometric Paschen- $\alpha$  survey of the Galactic Center. Due to the longer wavelength, their data suffer less from extinction than H $\alpha$  observations. The lack of velocity information, however, provides a challenge in disentangling DIG emission along the line of sight.

Radio observations give us an opportunity to investigate the distribution of the DIG throughout the Galactic disk. Despite its low density,  $\sim$ 80–90% of the total Galactic free-free emission is believed to come from the DIG and thus the DIG is a major source of radio continuum and radio recombination line (RRL) emission ([Reynolds 1984](#)). There have been numerous studies of the DIG using low frequency radio observations (e.g., [Roshi & Anantharamaiah 2000](#)), with a spatial resolution of  $\sim 2^\circ$ . [Liu et al. \(2013\)](#) reports on an ongoing beam-sampled 1.4 GHz RRL survey using the 305-m Arecibo Radio Telescope called “SIGGMA.” These data extend from  $65^\circ > \ell > 32^\circ$ ,  $|b| < 1.5^\circ$ . The spatial resolution is  $\sim 3'$ , and the spectral resolution is  $\sim 5 \text{ km s}^{-1}$ . Because of the survey strategy, some large-scale DIG emission is filtered out in SIGGMA. The other extant RRL survey with sub-degree resolution is the 1.4 GHz survey of [Alves et al. \(2015\)](#). Due to limited sensitivity, however, the data were binned to  $20 \text{ km s}^{-1}$  spectral resolution at  $15'$  spatial resolution.

While the most likely source of ionizing photons is high-mass stars (e.g., [Domgörgen & Mathis 1994](#)), the processes of how the DIG maintains its ionization are still not fully understood. For example, the DIG is found to be in a lower ionization state compared to the plasma in H II regions ([Madsen et al. 2006](#)). Furthermore, the radiation field emitted by O-stars changes as the radiation transitions from the H II region through the photodissociation region (PDR; e.g., [Hoopes & Walterbos 2003](#)). Using Monte Carlo photoionization simulations, [Wood & Mathis \(2004\)](#) argue that helium-ionizing photons in density bounded H II regions are suppressed as radiation escapes from the region, while the hydrogen-ionizing continuum hardens. The suppression of helium-ionizing photons should lead to a decrease in the  $N(\text{He}^+)/N(\text{H}^+)$  ionic abundance ratio with distance from the region, an effect observed for a sample of Galactic H II regions ([Luisi et al. 2016, 2019](#)). It is not clear, however, where this suppression takes place, and what mechanisms are responsible.

It further remains unclear how the radiation from O-stars in H II regions is able to propagate across the kiloparsec size-scales required given the observed DIG distribution. The radiation must first escape through the H II region PDRs, and H II regions themselves should be surrounded by “halos” of diffuse ionized gas. These halos have been observed around individual Galactic H II regions ([Deharveng et al. 2009; Anderson et al. 2015](#)), but not in a systematic study. Although the halos are sometimes thought of as being separate from the DIG (e.g., [McKee & Williams 1997](#)), here we do not draw a distinction.

H II regions are known to leak a significant fraction of their ionizing photons into the ISM, but the exact percentage is only known very roughly. [Oey & Kennicutt \(1997\)](#) find that up to 50% of the ionizing radiation escapes Large Magellanic Cloud (LMC) nebulae. [Zurita et al. \(2000, 2002\)](#) simulate escape fractions from H II regions in the galaxy NGC 157 to match the observed H $\alpha$  distribution, and find good agreement when 30% of Lyman continuum photons escape from each H II region and propagate through the ISM isotropically. These studies have so far only focused on the most luminous regions. Recent investigations of less luminous H II regions find a lower percentage of leaking photons than that of previous works on more luminous regions:  $15 \pm 5\%$  for NGC 7538 ([Luisi et al. 2016](#)) and  $25 \pm 10\%$  for RCW 120 ([Anderson et al. 2015](#)). These results may indicate that giant H II regions have a much larger effect in maintaining the DIG than compact H II regions and may be the primary source of ionizing photons in the ISM, as previously suggested by, e.g., [Zurita et al. \(2000\)](#).

Distinguishing the DIG from H II region emission is a major challenge in analyzing RRL emission from the Galactic plane. In the H II Region Discovery Survey (HRDS), [Bania et al. \(2010\)](#) observed discrete H II regions and find that nearly 30% of all targets have two or more separate RRL components at different velocities. Using subsequent observations, [Anderson et al. \(2015\)](#) show that one of these components is from the DIG and the other is from the compact H II region that was targeted. In a follow-up study, [Luisi et al. \(2017\)](#) analyze the DIG-only components from the HRDS from  $40^\circ > \ell > 18^\circ$  and  $|b| < 1^\circ$  and find that the intensity of diffuse Spitzer GLIMPSE  $8.0 \mu\text{m}$  emission is correlated with both the locations of discrete H II regions and the intensity of the diffuse RRL emission. This suggests that the soft UV photons responsible for creating the infrared (IR) emission through excitation of polycyclic aromatic hydrocarbons (PAHs) may have a similar origin as the more energetic photons causing the RRL emission.

The Green Bank Telescope (GBT) Diffuse Ionized Gas Survey (GDIGS) is an ongoing RRL survey of the Milky Way disk that is investigating the distribution and properties of the DIG. GDIGS will encompass  $\sim 46 \text{ deg}^2$  from  $32.3^\circ > \ell > -5^\circ$  and  $|b| < 0.5^\circ$  once finished. In addition, GDIGS will make several explorations above and below the Galactic plane around active star formation regions. The observing frequencies of 4–8 GHz make GDIGS essentially extinction-free, with higher spatial and spectral resolution compared to previous large-scale RRL surveys.

Here we show and analyze our GDIGS data of the H II region complex W43, one of the most active zones of star formation in the inner Galaxy (e.g., Nguyen Luong et al. 2011). W43 is located at a parallax distance of  $5.49^{+0.39}_{-0.34} \text{ kpc}$  (Zhang et al. 2014), close to the near end of the Galactic bar and the inner Scutum arm, a region of complex gas dynamics (e.g., Benjamin et al. 2005; Wegg et al. 2015). In fact, Nguyen Luong et al. (2011) suggest that the W43 complex may have formed due to its particular location via converging gas flows. In a GBT study of the inner-Galaxy DIG, Luisi et al. (2017) show that the diffuse gas near W43 is predominantly found at two distinct velocities:  $\sim 40 \text{ km s}^{-1}$  and  $\sim 100 \text{ km s}^{-1}$ . It is still uncertain whether these two velocity components originate from two distinct distances or whether they may be caused by noncircular streaming motions near the end of the bar. With the data provided by GDIGS, we can probe the distribution of the DIG near W43 in more detail than previous observations (see §3). We discuss the kinematics of the W43 complex in §4 and constrain the  $N(\text{He}^+)/N(\text{H}^+)$  ionic abundance ratio of the DIG in §5. We analyze the dust temperature distribution in §6 and conclude in §7.

## 2. OBSERVATIONS

GDIGS is a radio recombination line survey at C-band (4–8 GHz) using the GBT. The average half-power beam width at these frequencies is  $\sim 2'$ . We make total-power spectral observations in “On-the-Fly” (OTF) mode. The telescope is slewed at  $54'' \text{ s}^{-1}$ , and data are sampled every 0.38 s, or  $20''$ . Rows are spaced every  $40''$ , which is the Nyquist rate at the highest frequency. We observe a reference position  $\sim 3^\circ$  off the Galactic mid-plane every 16 rows, or  $\sim 20$  minutes, such that the time between any on-source scan and a reference scan is less than 10 minutes. To ensure that the reference positions are devoid of RRL emission, we perform a total-power pointed observation on each reference position prior to the mapping observations.

The pointed observations use the same setup as the OTF maps, with On- and Off-source integration times

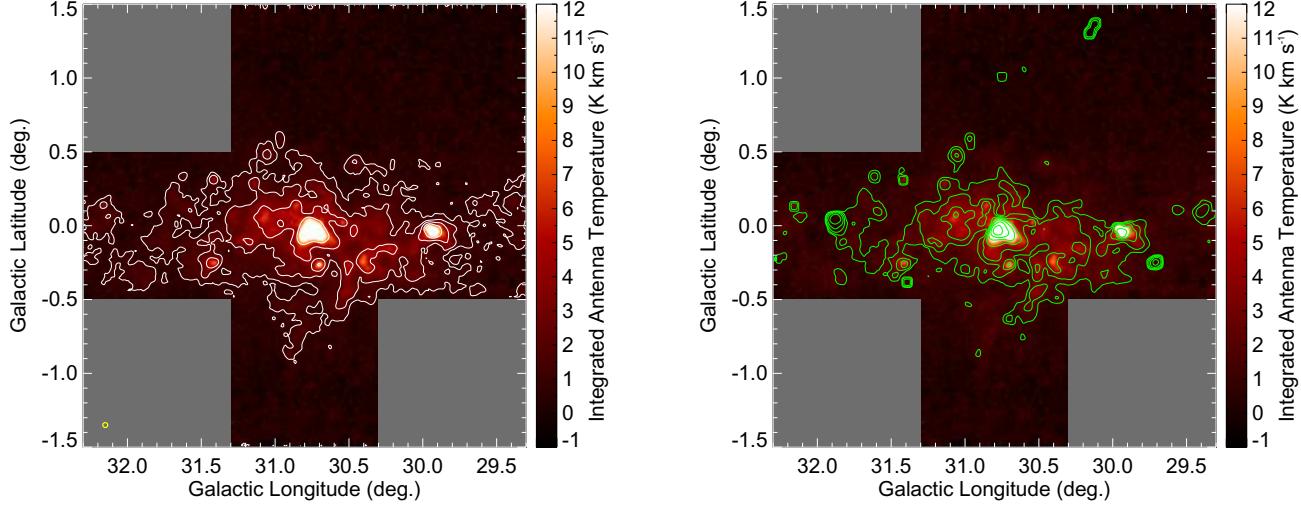
of 6 minutes per scan. The Off-source scans for these observations track the same azimuth and zenith angle path as the On-source scans, and are offset 6.5 minutes in RA such that they follow the same path on the sky.

We configure the VEGAS backend to simultaneously tune to 64 different spectral windows at two orthogonal polarizations, each within the 4–8 GHz receiver bandpass. Each spectral window spans 23.44 MHz with a spectral resolution of 2.9 kHz, or a velocity resolution of  $0.1\text{--}0.2 \text{ km s}^{-1}$  over the receiver bandpass. Of these 64 tunings, 22 are  $\text{H}\alpha$  lines from  $n = 95$  to 117, excluding  $\text{H}113\alpha$ , which is compromised by the nearby  $\text{H}142\beta$  line. In our configuration, the performance of the C-band receiver is poor below 4.3 GHz and above 7.5 GHz, which typically reduces the number of good  $\text{H}\alpha$  lines to 15 (namely the  $\text{H}97\alpha$  through  $\text{H}111\alpha$  transitions).

The advantage of observing a large number of  $\text{H}\alpha$  transitions simultaneously is that we can average all lines at a given position to make one sensitive spectrum (Balser 2006), a technique that is well-understood (Anderson et al. 2011; Liu et al. 2013; Alves et al. 2015; Luisi et al. 2018). After removing transient RFI, we subtract a polynomial baseline for each transition and shift the spectra so that they are aligned in velocity (Balser 2006). We re-grid the 15 good  $\text{H}\alpha$  lines to a velocity resolution of  $0.5 \text{ km s}^{-1}$  and a spatial resolution of  $30''$  using the *gbtgridder*<sup>1</sup> with a Gaussian kernel. We then average the individual maps using a weighting factor of  $t_{\text{intg}} T_{\text{sys}}^{-2}$  where  $t_{\text{intg}}$  is the integration time and  $T_{\text{sys}}$  is the system temperature. We finally perform a 0th-order baseline correction on the averaged maps by subtracting the median value of each single spectrum. To account for possible hydrogen and helium RRLs, we only use the known line-free portion of the spectrum (at hydrogen and helium velocities outside the  $10\text{--}130 \text{ km s}^{-1}$  range) in calculating the median.

We initially calibrate the intensity scale of our data using noise diodes fired during data acquisition. While the derived flux densities of the averaged maps agree with the values for the continuum flux calibrator 3C286 (see Ott et al. 1994) to within 4% regardless of elevation and weather conditions, the flux densities of the individual spectral windows may show deviations up to  $\sim 10\%$ . We therefore apply an intensity correction based on periodic mapping observations of the primary flux calibrator 3C286 to all individual maps prior to averaging. This reduces the deviation in intensity of the individual maps to  $< 3\%$ . We also periodically verify the calibration scale by observing

<sup>1</sup> <https://github.com/GreenBankObservatory/gbtgridder>



**Figure 1.** Left: Moment 0 map of the W43 complex, integrated from 10 to 130 km s<sup>-1</sup>. The average GBT beam size is shown as the yellow circle in the bottom left corner. Contours are at 1, 2, 4, and 8 K km s<sup>-1</sup>. The rms noise per pixel of the moment 0 map is 0.28 K km s<sup>-1</sup>. The emission at  $\ell \sim 30.8^\circ$  is due to W43 itself (W43-Main), while G29 (W43-South) is at  $\ell \sim 29.9^\circ$ . Essentially all visible emission is due to hydrogen. Right: Same, but the green contours are VGPS 1.4 GHz continuum emission and are logarithmically spaced between 5 and 320 K. Bright VGPS regions not associated with strong RRL emission are typically supernova remnants (e.g., Green 2014).

the RRL emission from W43 using pointed observations with the same setup. In our map, the root mean square (rms) noise per pixel within a single 0.5 km s<sup>-1</sup> velocity channel is 16.0 mK, which corresponds to 4.8 mK channel<sup>-1</sup> or 2.4 mJy beam<sup>-1</sup> channel<sup>-1</sup>. When smoothed to a spectral resolution of 5 km s<sup>-1</sup>, the rms noise is 0.78 mJy beam<sup>-1</sup>.

The GDIGS data discussed here encompass a total of six square degrees surrounding the H II region W43. Our maps include three square degrees within the Galactic plane from  $32.3^\circ > \ell > 29.3^\circ$ ,  $|b| < 0.5^\circ$ , as well as a two square degree exploration above the plane ( $31.3^\circ > \ell > 29.3^\circ$ ,  $0.5^\circ < b < 1.5^\circ$ ) and one square degree below the plane ( $31.3^\circ > \ell > 30.3^\circ$ ,  $-1.5^\circ < b < -0.5^\circ$ ).

### 3. SPATIAL DISTRIBUTION OF THE DIG NEAR W43

The W43 complex contains two large regions of active star formation: W43-Main at  $\ell \sim 30.8^\circ$  and W43-South at  $\ell \sim 29.9^\circ$ . W43-Main (hereon simply referred to as “W43”) is one of the brightest radio sources in the sky, with a GBT peak intensity of  $\sim 2$  K ( $\sim 1$  Jy) at X-band (Balser et al. 2011) and a total IR luminosity of  $\sim 6.5 \times 10^6 L_\odot$  (see Blum et al. 1999). In W43’s core lies a giant H II region ionized by a cluster of Wolf-Rayet and OB stars. W43-South, here referred to as “G29”, harbors the compact H II region G29.96–0.02, which has an IR luminosity of  $\sim 1.4 \times 10^6 L_\odot$  (Cesaroni et al. 1994). There are numerous other H II regions in the vicinity of the W43 complex. In the *WISE* Catalog of Galactic

H II regions (V2.2; Anderson et al. 2014), 298 H II regions and H II region candidates lie within the six square degrees mapped. Of these 298 regions, 115 are within 30' of W43 and 76 are within 30' of G29.

Due to the large luminosities and the large number of discrete H II regions near W43 and G29, DIG emission should be strong toward these directions. As part of a study of the compact H II region NGC 7528, Luisi et al. (2016) discovered diffuse hydrogen RRL emission near W43 with GBT X-band antenna temperatures  $> 10$  mK up to several hundred pc away from the ionizing star cluster. In a follow-up study of the large-scale distribution of the DIG, Luisi et al. (2017) confirmed that there is strong large-scale diffuse RRL emission across several square degrees near  $\sim 30^\circ$ . Due to the coarse resolution, however, the spatial correlation between H II regions and the DIG could not be probed in detail. Below we show our recent GDIGS observations of the W43 complex and examine the relationship between RRL emission from discrete H II regions and diffuse RRL emission.

#### 3.1. GDIGS mapping data

There are numerous bright RRL sources in the GDIGS datacube. As expected, the sources with the largest hydrogen RRL intensities are W43 and G29. We show a moment 0 map of the surveyed area in Figure 1. In order to maximize our signal-to-noise (S/N) ratio, the moment 0 map is integrated over the 10–130 km s<sup>-1</sup> local standard of rest (LSR) velocity range, within which we expect nearly all hydrogen RRL emission (see §4). We

find that 41 of the 298 H II regions in the map have an integrated hydrogen emission intensity  $> 10 \text{ K km s}^{-1}$  and four have integrated emission  $> 100 \text{ K km s}^{-1}$ .

We find an angularly large low-intensity structure below the Galactic plane that is clearly visible in the moment 0 map (Figure 1, left panel). This structure extends downward from W43 to  $b \sim -1^\circ$  where the intensity drops below the detection limit. We do not detect a counterpart to this structure above the plane. The structure may be associated with the “worm-ionized medium” (e.g., Heiles et al. 1996), low-density ionized gas within filamentary structures extending outwards from the Galactic plane. The worms may act as pathways for Lyman continuum photons to escape from the mid-plane. The structure in Figure 1 coincides spatially and in velocity with the worm “GW30.5–2.5” found by Heiles et al. (1996). Based on their S-band continuum data, GW30.5–2.5 extends from W43 to  $b \sim -4^\circ$ , and from  $\ell \sim 30^\circ$  to  $32^\circ$ . It is likely that here we detect only the denser part of GW30.5–2.5. The fact that there are no discrete H II regions near  $\ell \approx 30.9^\circ$ ,  $b \approx -1.0^\circ$  (see §3.2) suggests that the majority of emission within this structure is truly due to the DIG.

From our GDIGS maps, we can constrain the emission measure (EM),

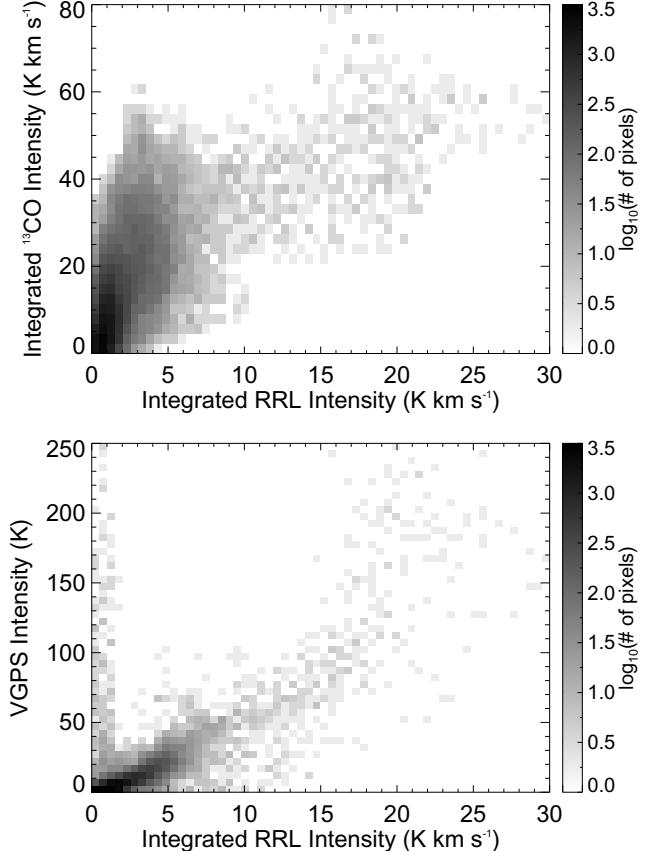
$$\frac{\text{EM}}{\text{pc cm}^{-6}} \equiv \int_{\text{los}} \left( \frac{n_e}{\text{cm}^{-3}} \right)^2 d \left( \frac{s}{\text{pc}} \right), \quad (1)$$

where  $n_e$  is the electron number density and  $s$  is the path length, integrated along the line of sight. For H $\alpha$  lines and assuming an observing frequency of 5912.5 MHz,

$$\frac{\text{EM}}{\text{pc cm}^{-6}} \approx 1.01 \times 10^{-2} \left( \frac{T_L}{\text{K}} \right) \left( \frac{\Delta V}{\text{km s}^{-1}} \right) \left( \frac{T_e}{\text{K}} \right)^{3/2}, \quad (2)$$

where  $T_L$  is the hydrogen RRL intensity,  $\Delta V$  is the hydrogen RRL full width at half maximum (FWHM), and  $T_e$  is the electron temperature (see Condon & Ransom 2016, Anderson et al., in prep.). Using Equation 2 and assuming  $T_e = 8000 \text{ K}$ , we find a maximum EM of  $\sim 2 \times 10^5 \text{ pc cm}^{-6}$  toward W43. The rms (per-pixel) sensitivity of our data is  $\text{EM} \approx 2000 \text{ pc cm}^{-6}$ .

Comparison with other gas tracers shows that there is a strong correlation between RRL intensity and  $^{13}\text{CO}$  emission from the Galactic Ring Survey (GRS; Jackson et al. 2006) tracing the denser molecular gas. We show the pixel-by-pixel correlation after regridding the GRS data to the resolution of the GDIGS RRL map in the top panel of Figure 2. We calculate Pearson’s  $r$  as a measure of the correlation and find  $r = 0.73$ . The large value of  $r$  is expected since the W43 complex is one of the most active regions of ongoing star formation.



**Figure 2.** Top: Density plot of the pixel-by-pixel correlation between the integrated intensity GDIGS RRL map and GRS  $^{13}\text{CO}$  data. The grayscale represents the number of pixels within each grid point. The GRS data are gridded to the lower resolution of the GDIGS data. There is a strong correlation between RRL intensity and  $^{13}\text{CO}$  emission ( $r = 0.73$ ). Bottom: Correlation between the GDIGS RRL map and 1.4 GHz VGPS data ( $r = 0.67$ ). The data points with large VGPS intensity but weak RRL emission are mostly due to supernova remnants and active galactic nuclei. Since the GBT beam is larger than the pixel size, the intensities of pixels within one beam area are not truly independent of each other.

There is also a correlation ( $r = 0.67$ ) between RRL intensity and 1.4 GHz radio continuum emission from the Very Large Array (VLA) Galactic Plane Survey (VGPS; Stil et al. 2006, see bottom panel of Figure 2), a 1'-resolution, 1.4 GHz VLA survey with additional short-spacing data from the 100 m Effelsberg telescope (Reich & Reich 1986; Reich et al. 1990).

Numerous sources are bright in 1.4 GHz continuum but not in RRLs; these regions are predominantly supernova remnants (Green 2014) and active galactic nuclei (see Figure 1, right panel). The continuum intensity of these sources is dominated by synchrotron emission. The change in slope at integrated RRL intensities above

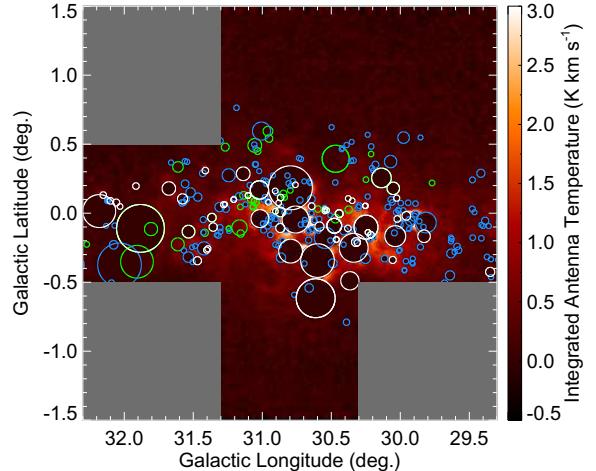
$\sim 18 \text{ K km s}^{-1}$  in the bottom panel of Figure 2 may indicate increased non-thermal emission in the regions of strongest RRL emission.

The correlation between RRL intensity and  $\text{H}\alpha$  emission from the WHAM Sky Survey (Haffner et al. 2003, not shown in the figure) is poor ( $r = 0.24$ ), presumably because the spatial resolution of WHAM is only  $\sim 1^\circ$  and because the W43 complex is known to be optically obscured (e.g., Green et al. 2019).

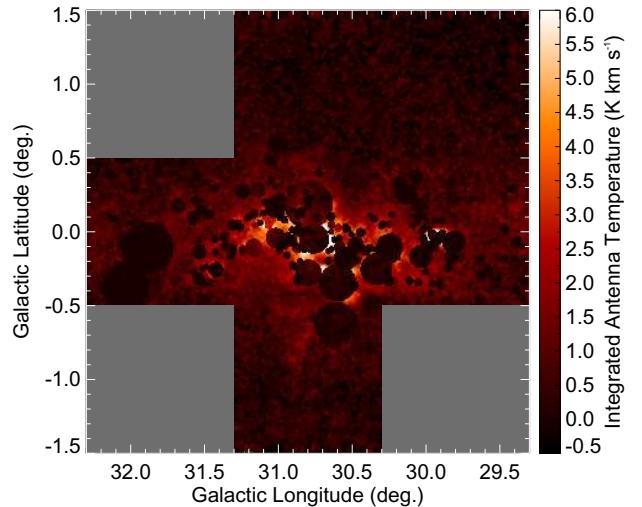
### 3.2. DIG-only data

In order to study the diffuse gas emission unaffected by emission from discrete H II regions, it is necessary to disentangle the RRL emission based on its origin. We know the locations and angular sizes of all H II regions and H II region candidates from the *WISE* catalog and can blank emission from within these regions. The *WISE* region sizes are determined by their IR morphology:  $22 \mu\text{m}$  emission from hot dust grains within the ionized H II region volume surrounded by  $12 \mu\text{m}$  emission from PAHs tracing the H II region PDR. Since PAHs cannot survive in a strong radiation field, the *WISE* region radii determined by the  $12 \mu\text{m}$  emission usually overestimate the size of the fully ionized H II regions. Without knowledge of the H II region velocities, however, we are forced to blank the GDIGS data across all spectral channels. This possibly results in the loss of diffuse emission, since diffuse gas may lie along the same line of sight as the H II region. Fortunately, the *WISE* catalog includes velocity information for 99 confirmed H II regions within the mapped range (partially compiled by the HRDS; see Bania et al. 2012). For these regions, we only exclude data within  $\pm 1.5 \text{ FWHM}$  of the H II region velocity, whereas we blank all velocity channels for the remaining 198 H II region candidates.

As an example, we show the W43 complex integrated over the  $90 - 110 \text{ km s}^{-1}$  velocity range with H II regions from the *WISE* catalog displayed and blanked (Figure 3). We also show the H II region-subtracted moment 0 map, or “DIG-only” map, in Figure 4. We detect diffuse RRL emission across essentially the entire survey range, although it is faint at  $|b| > 1^\circ$ . The total integrated hydrogen intensity between  $10$  and  $130 \text{ km s}^{-1}$  is  $2,772 \text{ K km s}^{-1}$  for the DIG-only data, compared to  $5,096 \text{ K km s}^{-1}$  for the unmasked datacube. Even if our measurements of some H II region sizes and velocities are inaccurate, the vast majority of the emission shown in Figure 4 cannot be explained by H II region emission alone. We find the strongest DIG emission near W43 and in the zone between W43 and G29, suggesting that at these locations the ISM is strongly affected by radiation leaking from H II regions. We note that



**Figure 3.** Diffuse RRL emission, integrated from  $90 \text{ km s}^{-1}$  to  $110 \text{ km s}^{-1}$ . Emission from known H II regions within these velocities (white circles) was cut from the map. We also removed emission from H II region candidates (blue circles). The green circles indicate H II regions outside the  $90 \text{ km s}^{-1}$  to  $110 \text{ km s}^{-1}$  velocity range.



**Figure 4.** DIG-only moment 0 map. Velocity channels within  $\pm 1.5 \times \text{FWHM}$  of the H II region velocity were blanked for each known H II region. For H II region candidates without known velocities, we removed emission from all velocity channels. There is extended DIG emission between W43 and G29.

the brightest pixels of the DIG-only map are located directly outside W43 and may trace RRL emission from the H II region itself, possibly because the radius of W43 is underestimated in the *WISE* catalog.

We derive the average scale height,  $h$ , of the DIG-only RRL data at all longitudes where our maps extend to  $|b| = 1.5^\circ$ , and find  $h = 30 \pm 4'$ . Assuming a distance

of 6 kpc (see §4), this corresponds to  $h = 52 \pm 7$  pc. This number is much lower than that derived from the combination of pulsar and H $\alpha$  data ( $250 < h < 500$  pc; Berkhuijsen et al. 2006), suggesting that we are sensitive to only the denser DIG components near discrete H II regions.

### 3.3. Modeling the DIG Emission

If the ionization of the DIG is maintained by O-type stars within H II regions, it should be possible to estimate the spatial distribution of the DIG from the H II region distribution. The *WISE* catalog provides an ideal tool for this analysis as it is essentially complete in the inner Galaxy and contains the locations and angular sizes of all H II region candidates.

Photoionization models (e.g., CLOUDY; Ferland et al. 2017) can, in principle, predict the ionizing radiation field around a single H II region given enough information on that region. Galactic H II regions, however, vary substantially in morphology and in their physical parameters (e.g., density), and for the majority of H II regions in our survey these parameters are unknown. It is therefore unfeasible to model each H II region based on its individual properties. Instead, we aim to estimate the large-scale distribution and intensity of the DIG from a purely empirical model with as few input parameters as necessary. We restrict our model inputs to only the location and size of each H II region, and its average GDIGS RRL intensity.

In a study of eight Galactic H II regions, Luisi et al. (2019) discovered that the slope of the hydrogen RRL intensity with distance from the H II region is roughly inversely proportional to the H II region size traced by its PDR. This is in agreement with previous studies (e.g., Zurita et al. 2000), who found that large H II region complexes are believed to affect the ISM out to larger distances compared to compact H II regions. For H II regions in a constant density environment, this discovery is supported by simple geometry. The RRL emission intensity is proportional to the number density squared and therefore this relationship does not hold for variations in density, since high-density ionized gas will generate disproportionately large RRL intensities compared to low-density gas. For our model, however, we assume that the angular size of the H II region constrains the slope of RRL emission with distance from that region.

The functional form of the hydrogen RRL emission intensity with distance from an H II region is not well-understood. Previous work showed that the hydrogen RRL emission decreases roughly as a power-law beyond the H II region PDR (Luisi et al. 2016) for the two stud-

ied H II regions: NGC 7538 and W43. The difference between the power-law model and an exponential function, however, was marginal. There is no statistically significant difference between these two models in the RRL emission surrounding the eight Galactic H II regions observed in Luisi et al. (2019).

In our model we assume that all DIG emission is due to photons leaking from H II regions and that the diffuse hydrogen RRL emission decreases with distance from an H II region using two different functional forms: an exponential (see §3.3.1) and a power law (see §3.3.2). The model treats each region as perfectly circular and isotropic, with a radius equal to the radius given in the *WISE* catalog. We further assume that the RRL emission intensity at the boundary of each H II region is proportional to the average integrated RRL intensity within the region. We derive the average integrated intensities from our GDIGS moment 0 map, using the H II region sizes defined by the *WISE* catalog.

We compare each model with the GDIGS DIG-only map and find the best-fit model by minimizing the sum of the squares of the residuals,  $\sum \epsilon^2$ , using the GDIGS moment 0 map as the reference. We also compute Pearson's correlation coefficient,  $r$ , as a proxy for the quality of the model.

#### 3.3.1. Exponential model

For the exponential model, we assume

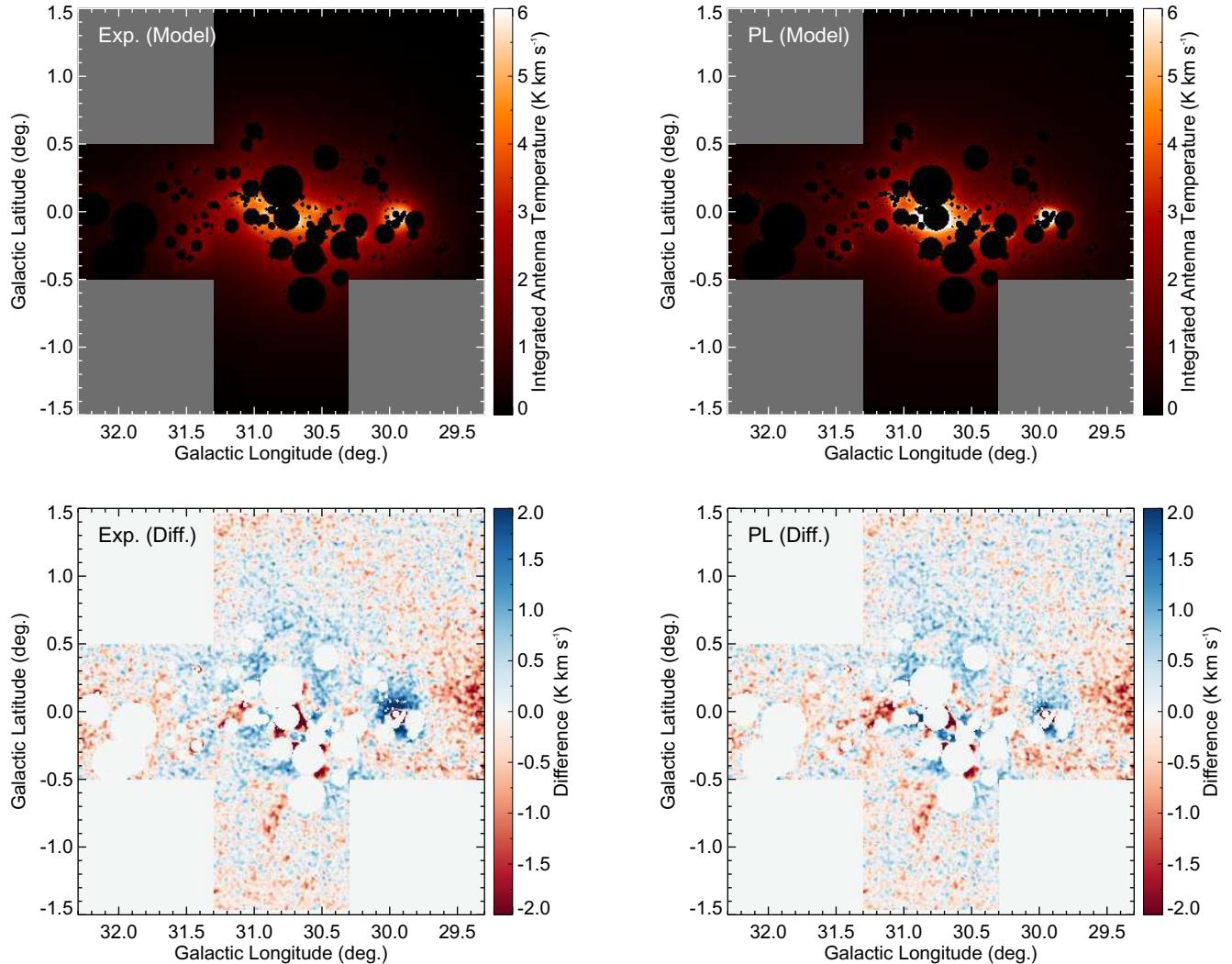
$$T_{A, \text{model}}(x, y) = k \sum_{i=1}^N T_{A, i} \exp\left(-m \frac{r_i}{r_{\text{PDR}, i}}\right), \quad (3)$$

where  $T_{A, \text{model}}$  is the antenna temperature of the model for each individual pixel in the map,  $k$  and  $m$  are the fit parameters,  $N = 298$  is the total number of H II regions in the map,  $T_{A, i}$  is the GDIGS moment 0 map intensity averaged over H II region  $i$ ,  $r_i$  is the distance from the H II region, and  $r_{\text{PDR}, i}$  is the radius of the region as given in the *WISE* catalog. The parameter  $k$  is always calculated such that the total observed flux is reproduced, i.e. the total integrated intensity of the model map must be equal to the total integrated intensity of the moment 0 map;  $k$  is therefore not truly a free parameter. We give our best-fit values for  $k$ ,  $m$ , Pearson's  $r$ , and the sum of the squares of the residuals,  $\sum \epsilon^2$ , in Table 1. We show the model in the left panel of Figure 5.

We assume that the major uncertainty contributions for the model are the H II region sizes defined by the *WISE* catalog and the averaged integrated intensities from our moment 0 map; we assume a 20% uncertainty for both. We perform a Monte Carlo uncertainty estimation of the model output parameters  $k$  and  $m$  by resampling the input parameters  $T_{A, i}$  and  $r_{\text{PDR}, i}$  from a Gaus-

**Table 1.** H II Region Model Parameters

Velocity range ( $\text{km s}^{-1}$ )	Functional form	$k$	$m$	Pearson's $r$	$\sum \epsilon^2$ ( $\text{K}^2 \text{km}^2 \text{s}^{-2}$ )
10–130	Exponential	$0.133 \pm 0.033$	$0.33 \pm 0.05$	0.859	20,627
10–130	Power law	$0.262 \pm 0.071$	$1.83 \pm 0.11$	0.880	17,452
10–65	Exponential	$0.165 \pm 0.041$	$0.32 \pm 0.05$	0.677	3,795
10–65	Power law	$0.307 \pm 0.083$	$1.81 \pm 0.11$	0.683	3,676
65–130	Exponential	$0.148 \pm 0.037$	$0.36 \pm 0.05$	0.879	11,456
65–130	Power law	$0.276 \pm 0.075$	$1.89 \pm 0.12$	0.905	8,917



**Figure 5.** Top left: model of diffuse hydrogen RRL emission, assuming an exponential decrease in RRL intensity with distance from each H II region. Bottom left: difference between the GDIGS moment 0 data and the model. Larger values indicate stronger emission in the model. Top right: same as top left, but assuming a power-law decrease in RRL emission with distance from each H II region. The power-law model is a better fit to the data (see §3.3). Bottom right: difference between the GDIGS moment 0 data and the power-law model.

sian probability distribution centered on their nominal values and a standard deviation of 20%. We recompute  $k$  and  $m$  ten times (since the model is computationally intensive, we are restricted in the total number of samples). The uncertainty values given in Table 1 are the standard deviations of the recomputed values of  $k$  and  $m$ .

An additional source of uncertainty is the rms noise in the GDIGS moment 0 map, as well as the fact that real H II regions are neither homogeneous nor spherically symmetric. By adding random noise to the GDIGS map and recomputing  $k$  and  $m$ , we find that the rms noise of the map has a negligible impact on the overall uncertainty (< 1%). Since we cannot model each H II region separately, the uncertainty associated with the H II region morphologies is unknown. Based on visual inspection of the difference maps between the model and the data, however, the uncertainty contribution must be significant. We therefore consider the uncertainties shown in Table 1 to be lower limits of the overall model uncertainty.

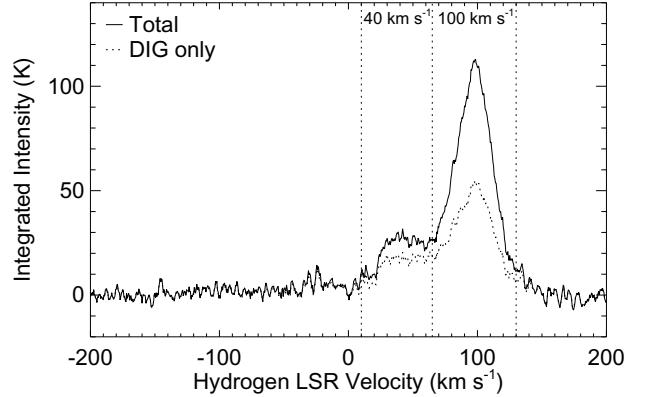
### 3.3.2. Power law model

Here we use a power-law model of the form

$$T_{A, \text{model}}(x, y) = k \sum_{i=1}^N T_{A, i} \left( \frac{r_i}{r_{\text{PDR}, i}} \right)^{-m} \quad (4)$$

to model the hydrogen RRL emission around W43. Again, we give our best-fit values in Table 1 and show the model map in the right panel of Figure 5.

The power law model is a significantly better fit to the data than the exponential model, both in terms of the squared residuals and Pearson's  $r$ . The diffuse emission at  $\ell > 31.5^\circ$  is marginally brighter in the moment 0 RRL data than either model map, as well as emission at  $\ell < 29.8^\circ$ . This excess emission may indicate that there is a truly diffuse component associated with the Galactic plane that we are not fitting in the model. With the exception of the DIG feature below the plane near  $\ell \approx 30.9^\circ$ , both models estimate the low-intensity DIG at  $|b| > 0.5^\circ$  quite well. While the exponential model tends to underestimate the RRL intensity near W43 and G29, however, the power law model better represents the bright DIG emission outside these luminous H II region complexes. This difference is particularly noticeable toward the Galactic north of G29. Overall, we find that the model provides a reasonable estimate of the diffuse RRL emission by only considering the aggregate effect of H II regions on the ISM.

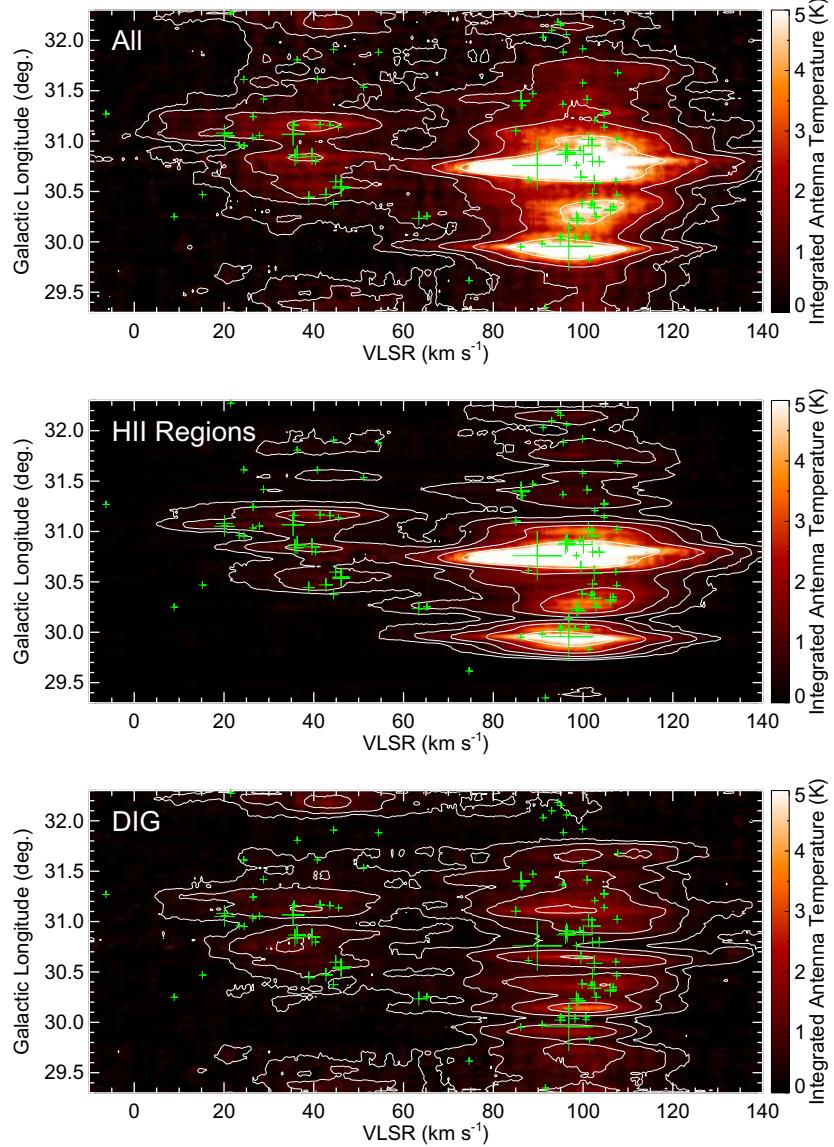


**Figure 6.** RRL spectrum of the mapped area. The emission is mostly found near two velocities:  $\sim 40 \text{ km s}^{-1}$  and  $\sim 100 \text{ km s}^{-1}$ . Emission from the DIG (dotted line) follows that of the overall ionized gas. The weak emission near a hydrogen LSR velocity of  $-25 \text{ km s}^{-1}$  is due to helium associated with the  $\sim 100 \text{ km s}^{-1}$  hydrogen RRL emission.

## 4. KINEMATICS OF THE DIG NEAR W43

Due to its location near the end of the Galactic bar, the gas dynamics near W43 are complex. The DIG emission at  $\sim 30^\circ$  is mainly found at two distinct velocities,  $\sim 40 \text{ km s}^{-1}$  and  $\sim 100 \text{ km s}^{-1}$ . Determining the distance to the DIG is hampered by not only the kinematic distance ambiguity (KDA; see Anderson et al. 2012a), but also possibly strong streaming motions at the bar/spiral arm interface (e.g., Beuther et al. 2012).

The hydrogen RRL emission at  $\sim 40 \text{ km s}^{-1}$  and  $\sim 100 \text{ km s}^{-1}$  is clearly visible in the GDIGS data. We show a total-power spectrum of the mapped area in Figure 6. In the top panel of Figure 7 we show the longitude–velocity ( $\ell, v$ )-diagram of the W43 complex, restricted to  $|b| < 0.5^\circ$ . The  $100 \text{ km s}^{-1}$  component is distributed across the entire longitude range, while the  $40 \text{ km s}^{-1}$  emission is mostly restricted to  $\ell > 30.3^\circ$ . The two components are separated in velocity space which may suggest that they arise from two separate gas clouds. It is surprising that there is essentially no  $40 \text{ km s}^{-1}$  emission near G29 ( $\ell \sim 29.9^\circ$ ). Some weak RRL emission is present between  $30 - 70 \text{ km s}^{-1}$  at  $\ell < 29.6^\circ$ , which is likely caused by radiation escaping from two nearby H II regions outside the mapped area: G029.165–0.035 and G029.110–0.023. The WISE catalog lists the coordinates of the H II regions as  $(\ell, b) = (29.165^\circ, -0.035^\circ)$  for G029.165–0.035 and  $(\ell, b) = (29.110^\circ, -0.023^\circ)$  for G029.110–0.023, and their hydrogen RRL velocities as  $61.7$  and  $52.4 \text{ km s}^{-1}$ , respectively.

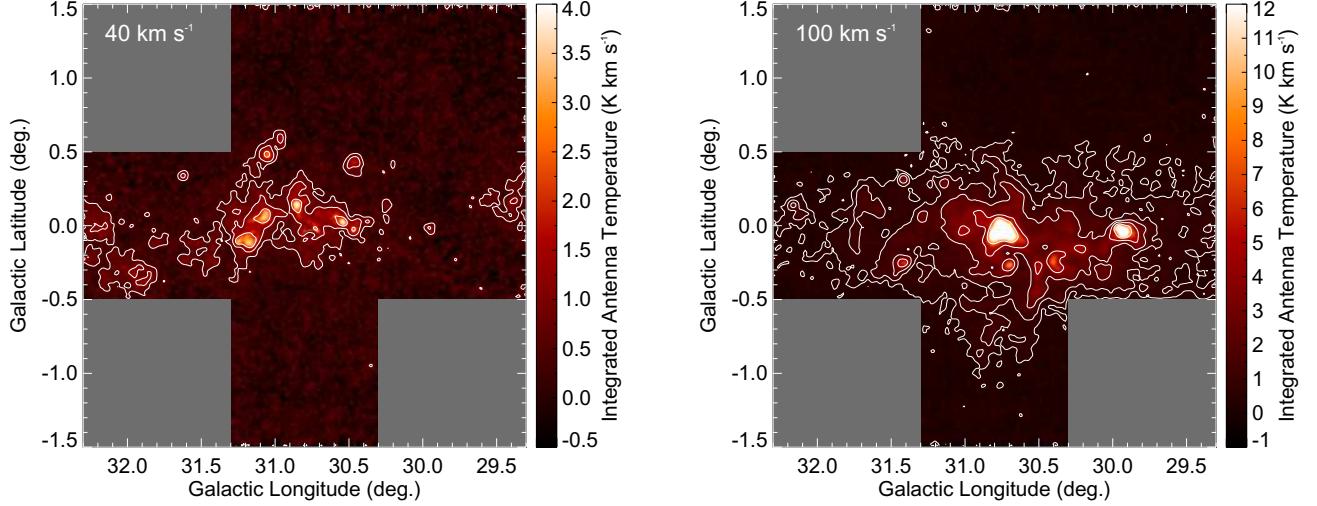


**Figure 7.** Top: longitude–velocity diagram of the W43 complex. The RRL emission clearly originates from two distinct velocities ( $\sim 40 \text{ km s}^{-1}$  and  $\sim 100 \text{ km s}^{-1}$ ), with less emission in between. The  $\sim 40 \text{ km s}^{-1}$  emission is restricted to  $\ell > 30.3^\circ$ , while the  $\sim 100 \text{ km s}^{-1}$  emission is extended across the entire mapped longitude range. Contours are at 0.25, 0.5, 1, 2, and 4 K. The green crosses indicate the location of H II regions; the size of each cross is proportional to the RRL intensity of the region. Middle: same, for H II region emission only. Bottom: same, for DIG-only map.

The DIG emission appears to follow the same trends as the overall ionized gas emission. We show the DIG-only ( $\ell, v$ )-diagram of the W43 complex in the bottom panel of Figure 7. We find that in the DIG-only datacube, the  $40 \text{ km s}^{-1}$  and  $100 \text{ km s}^{-1}$  components appear more strongly separated in velocity than for the H II region ( $\ell, v$ )-diagram shown in the middle panel of Figure 7. This does not necessarily imply, however, that much of the RRL emission between  $40$  and  $100 \text{ km s}^{-1}$  must be due to discrete H II regions. Due to their Gaussian velocity profile with typical RRL widths  $\sim 25 \text{ km s}^{-1}$ , emission from bright H II regions is detectable over a wide

velocity range. By removing the H II region component, we also subtract fainter emission at velocities distant from the H II region velocities. As a result, the lack of RRL emission in the  $40 - 100 \text{ km s}^{-1}$  range in the DIG-only data may simply highlight the removal of discrete H II regions.

The integrated emission of the DIG-only data associated with the  $40 \text{ km s}^{-1}$  component (defined as within the  $10 - 65 \text{ km s}^{-1}$  range) is  $787 \text{ K km s}^{-1}$ , while the integrated  $100 \text{ km s}^{-1}$  emission (within the  $65 - 130 \text{ km s}^{-1}$  range) is  $1,994 \text{ K km s}^{-1}$ . For the original data including discrete H II regions, the numbers are  $1,177 \text{ K km s}^{-1}$



**Figure 8.** Left: Moment 0 map of the W43 complex for the  $40 \text{ km s}^{-1}$  velocity component (integrated from  $10\text{--}65 \text{ km s}^{-1}$ ). Contours are at  $0.5, 1, 2, 4$ , and  $8 \text{ K km s}^{-1}$ . The faint emission at  $\ell < 29.6^\circ$  is likely due to two nearby H II region outside the mapped area. Right: Same, but for the  $100 \text{ km s}^{-1}$  velocity component (integrated from  $65\text{--}130 \text{ km s}^{-1}$ ).

( $40 \text{ km s}^{-1}$  component) and  $3,933 \text{ K km s}^{-1}$  ( $100 \text{ km s}^{-1}$  component), respectively. The fraction of DIG emission to total emission is larger for the  $40 \text{ km s}^{-1}$  component (67%) than the  $100 \text{ km s}^{-1}$  component (51%). There is no statistically significant difference between the scale heights of the DIG-only 40 and  $100 \text{ km s}^{-1}$  velocity components ( $30 \pm 15'$  and  $29 \pm 6'$ , respectively).

#### 4.1. The origin of the $40 \text{ km s}^{-1}$ velocity component

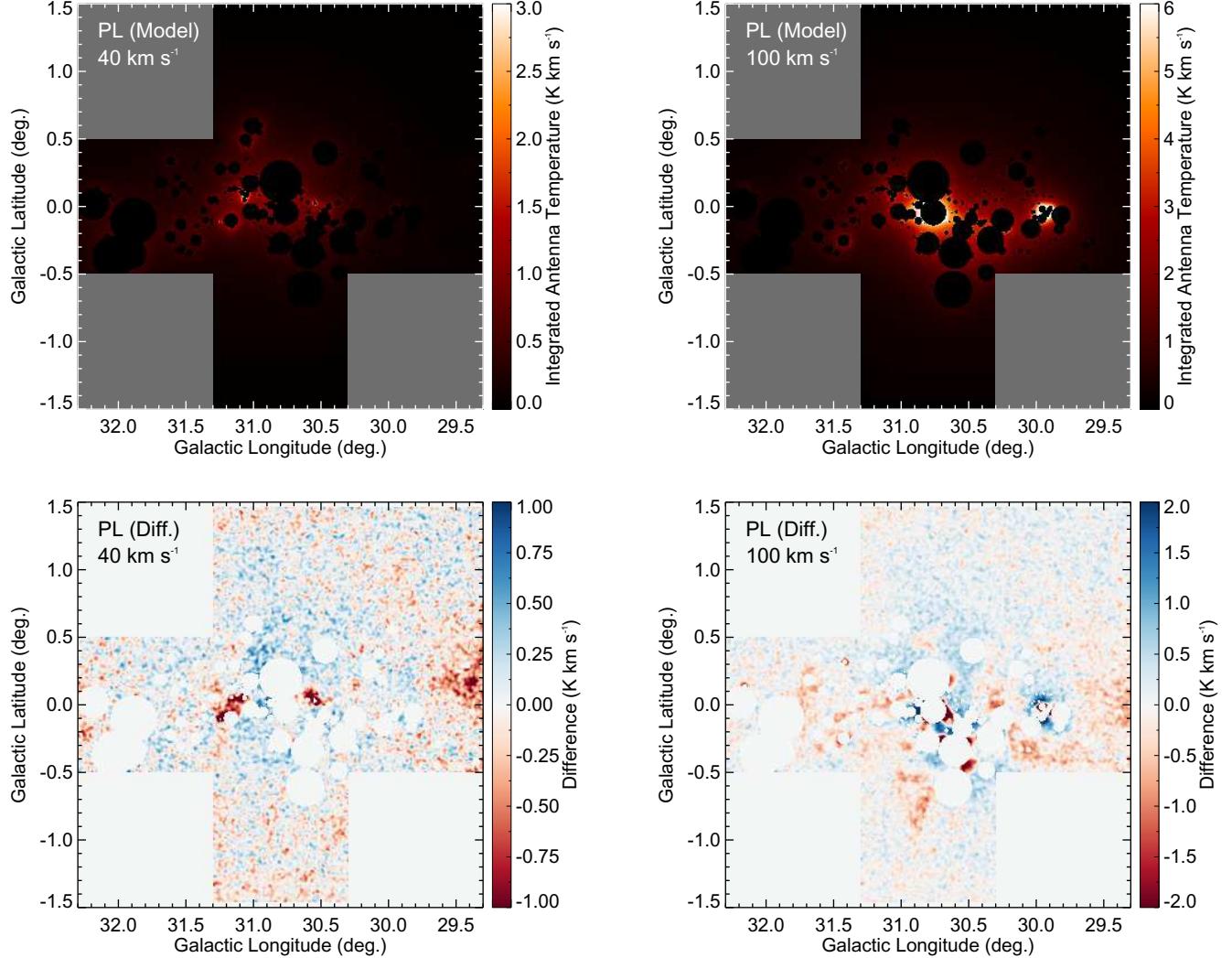
While the  $100 \text{ km s}^{-1}$  DIG component is likely associated with W43 at a distance of  $\sim 6 \text{ kpc}$ , the origin of the  $40 \text{ km s}^{-1}$  component is less clear. Luisi et al. (2017) argue that the most likely kinematic distance for the 40 component is  $\sim 12 \text{ kpc}$ . Based on their pointed RRL data, however, they cannot rule out that the observed velocities near W43 are caused by interacting gas clouds at a single distance. Assuming a combination of unperturbed, purely circular gas motion around the Galactic center and gas streaming motions along the bar, Luisi et al. (2017) derive an expected perturbed velocity component of  $57 \pm 4 \text{ km s}^{-1}$  for the W43 complex, which is near the observed  $40 \text{ km s}^{-1}$  emission. An ionized gas component centered near  $40 \text{ km s}^{-1}$  is also found in  $\text{H}\alpha$  emission from the WHAM survey. Based on its vertical morphology, Krishnarao et al. (2017) associate this component with the nearby Sagittarius-Carina spiral arm at a distance of  $\sim 2 \text{ kpc}$ .

We show moment 0 maps of the 40 and  $100 \text{ km s}^{-1}$  components in Figure 8. The strongest  $40 \text{ km s}^{-1}$  emission is found as an arc-like feature surrounding W43 toward the Galactic north, extending over a range of  $\sim 1^\circ$  in longitude. Fainter  $40 \text{ km s}^{-1}$  emission, possibly

associated with nearby H II regions, is visible at  $\ell \sim 32^\circ$ . The  $40 \text{ km s}^{-1}$  emission component found at  $\ell < 29.6^\circ$  is likely caused by H II regions outside the mapped area (see above).

The spatial confinement of the strong  $40 \text{ km s}^{-1}$  emission at the end of the Galactic bar is suggestive of a non-circular origin due to streaming motions around the bar. This would place the observed  $40 \text{ km s}^{-1}$  DIG at relatively small physical distances ( $\sim 100 \text{ pc}$ ) from the W43 complex. As a result, we would expect W43 and nearby H II regions at similar velocities ( $\sim 90 \text{ km s}^{-1}$ ) to contribute to the ionization of the diffuse  $40 \text{ km s}^{-1}$  component.

We test this hypothesis by running our model of expected DIG emission (§3.3) separately for the 40 and  $100 \text{ km s}^{-1}$  velocity components. If much of the  $40 \text{ km s}^{-1}$  DIG maintains its ionization from sources at larger velocities, we would expect a poor correlation between the model output and the GDIGS data when only considering the DIG and H II regions near  $40 \text{ km s}^{-1}$  (within the  $10\text{--}65 \text{ km s}^{-1}$  range). We show our model output and the difference maps for the two velocity components in Figure 9 (power-law model only) and give our best-fit model parameters in Table 1 (power-law and exponential model). We find that the correlation between the  $100 \text{ km s}^{-1}$  data and the model is indeed much better than that of the  $40 \text{ km s}^{-1}$  component (Pearson's  $r$  of 0.905 vs. 0.683). We note that the DIG emission found at  $\ell < 29.6^\circ$  may negatively impact our model quality for the  $40 \text{ km s}^{-1}$  component. When we blank all pixels with increased DIG emission at  $\ell < 29.6^\circ$  and re-run the model, the correlation is only marginally improved



**Figure 9.** Top left: model of diffuse hydrogen RRL emission, for the  $40 \text{ km s}^{-1}$  velocity component only, assuming a power-law decrease in RRL emission with distance from each H II region. Bottom left: difference map between the GDIGS moment 0 data and the model. Larger values indicate stronger emission in the model. Top right: same as top left, but for the  $100 \text{ km s}^{-1}$  component. The model quality is higher for the  $100 \text{ km s}^{-1}$  emission, indicating that the  $40 \text{ km s}^{-1}$  DIG may partially maintain its ionization from H II regions at larger velocities such as W43. This may indicate a small physical separation between the  $40 \text{ km s}^{-1}$  DIG and W43. Bottom right: difference map between the GDIGS moment 0 data and the power-law model.

( $r = 0.710$ ). This suggests that the DIG near  $40 \text{ km s}^{-1}$  maintains a significant portion of its ionization from H II regions outside the  $10\text{--}65 \text{ km s}^{-1}$  range, possibly from W43 itself.

If a significant fraction of the DIG near  $40 \text{ km s}^{-1}$  maintains its ionization from W43 and other physically nearby H II regions, it would place the  $40 \text{ km s}^{-1}$  emission at a distance of  $\sim 6 \text{ kpc}$  as well. The absence of emission between the two velocity components in  $\ell, v$ -space is unexpected, however. The lack of RRL detection between  $55$  and  $75 \text{ km s}^{-1}$  may be explained by a clumpy nature of the  $40 \text{ km s}^{-1}$  DIG, connected to the

$100 \text{ km s}^{-1}$  component by low-density plasma below our detection threshold.

## 5. IONIC ABUNDANCES

As radiation escapes from an H II region and into the surrounding ISM, the spectrum is affected by absorption and emission processes. Photoionization simulations suggest that helium-ionizing photons become suppressed as the radiation propagates through the H II region while the hydrogen-ionizing continuum is hardened (Wood & Mathis 2004). The radiation hardening may be due to the much larger ionization cross-section of helium compared to hydrogen, resulting in

the absorption of a large fraction of helium-ionizing photons within the ionization front of the H II region (Osterbrock 1989). For O-stars with effective temperatures  $< 35,000$  K, Weber et al. (2019) suggest that nearly the entire helium-ionizing radiation is absorbed within the H II region.

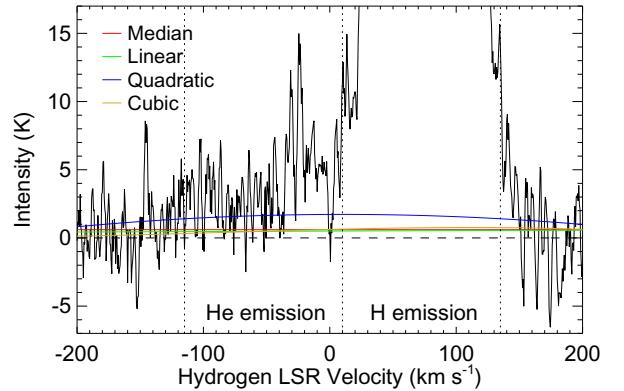
The radiation field hardness can be constrained by the  $y^+ = N(\text{He}^+)/N(\text{H}^+)$  ionic abundance ratio using RRL observations. Helium is ionized by more energetic radiation ( $h\nu \geq 24.6$  eV) than hydrogen ( $h\nu \geq 13.6$  eV), and therefore a larger value of  $y^+$  indicates a harder radiation field. Here,

$$y^+ = \frac{T_L(\text{He}^+) \Delta V(\text{He}^+)}{T_L(\text{H}^+) \Delta V(\text{H}^+)}, \quad (5)$$

where  $T_L(\text{He}^+)$  and  $T_L(\text{H}^+)$  are the RRL intensities of helium and hydrogen, respectively, and  $\Delta V(\text{He}^+)$  and  $\Delta V(\text{H}^+)$  are their FWHM line widths (Peimbert et al. 1992). In a previous study we found that for six out of eight observed H II regions,  $y^+$  decreases with distance from the region, suggesting that helium-ionizing photons are indeed suppressed (Luisi et al. 2019). The two regions that did not follow this trend were S206 and M17, which should have essentially no neutral helium within the H II region due to very hard radiation (Balser 2006). A larger sample size, however, is required to determine whether our results are applicable to the entire Galactic H II region population.

A major challenge in determining  $y^+$  using our GDIGS data is distinguishing between hydrogen emission and helium emission. The helium line is shifted by  $\sim -122 \text{ km s}^{-1}$  from that of hydrogen and, given a spectrum with sufficient S/N ratio, the helium peak can typically be easily separated from the hydrogen peak for a source with a well-defined velocity. The observed hydrogen RRL emission near W43, however, spans a velocity range from  $V_{\text{LSR}} \approx 0 \text{ km s}^{-1}$  to  $V_{\text{LSR}} \approx 140 \text{ km s}^{-1}$  (see Figure 7). At hydrogen velocities between  $\sim 0 \text{ km s}^{-1}$  and  $\sim 20 \text{ km s}^{-1}$  both hydrogen and helium emission may therefore be present. In addition, we cannot rule out hydrogen emission at  $V_{\text{LSR}} < 0 \text{ km s}^{-1}$ , possibly originating from the Outer Galaxy. Since we are unable to disentangle the low-velocity hydrogen emission from helium emission, we simply define all RRL emission below a velocity of  $10 \text{ km s}^{-1}$  as due to helium, and emission  $> 10 \text{ km s}^{-1}$  as due to hydrogen. We find that the exact choice of the cutoff velocity has only a minor impact on the derived  $y^+$  values.

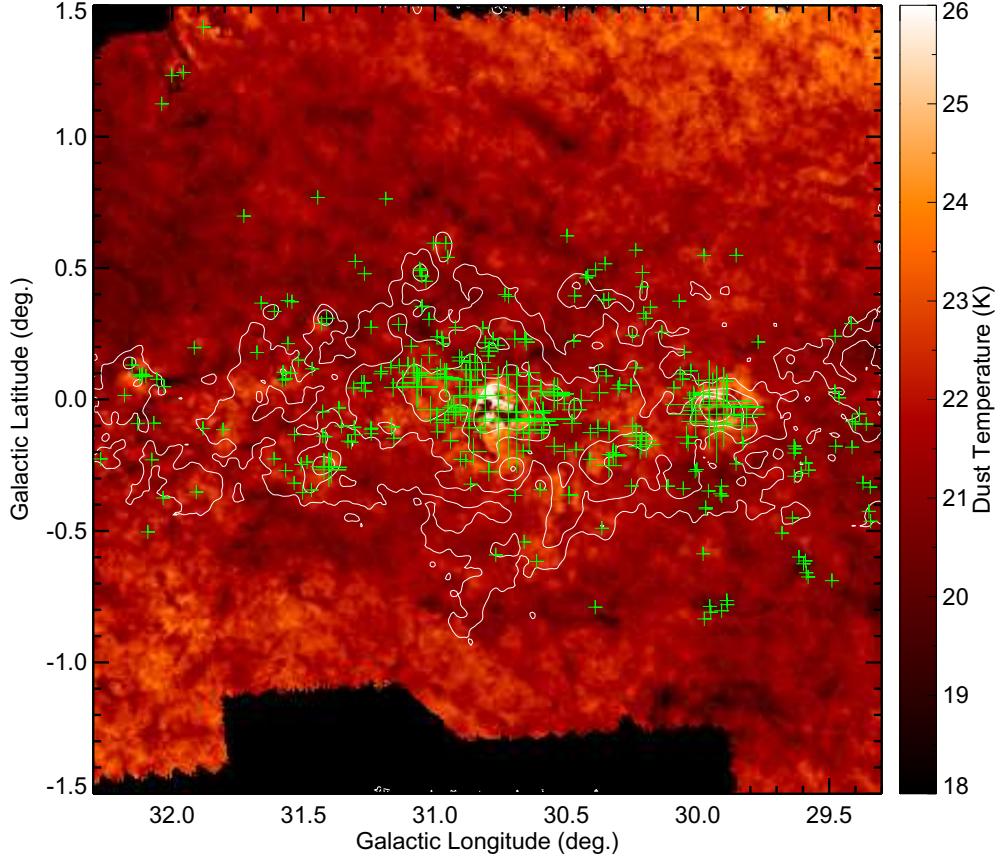
We derive  $T_L(\text{H}^+) \Delta V(\text{H}^+)$  for our averaged H $\alpha$  map by integrating the RRL emission over the entire map spatially, and between  $10 \text{ km s}^{-1}$  and  $132 \text{ km s}^{-1}$  in velocity space. We likewise calculate  $T_L(\text{He}^+) \Delta V(\text{He}^+)$



**Figure 10.** RRL spectrum of the mapped area, zoomed in to highlight the effect of fitting baselines of different orders. We fitted the baselines to the same emission-free regions for all pixels and across the entire velocity range of the data ( $-350$  to  $+350 \text{ km s}^{-1}$ ). The baseline models are shown in color. The dotted vertical lines mark the velocity ranges of expected helium and hydrogen emission. Although the difference between baseline models appears small, the aggregate effect is substantial when deriving  $y^+$ . Note that the baseline models are not directly fitted to the spectrum shown here, but rather for each individual pixel in the datacube; here we show only the combined effect on the overall spectrum.

by integrating the emission over the same helium velocities. We discard data at any other velocity to minimize confusion over whether these data should be classified as hydrogen or helium emission. We also discard data at  $\ell < 30.3^\circ$ ,  $b > 0.8^\circ$  where the baseline is poor. We find  $y^+ = 0.078$  for the entire map, a value similar to that found for the H II region population from the HRDS ( $y^+ = 0.068 \pm 0.023$ ; see Wenger et al. 2013).

We repeat the above analysis for the DIG-only datacube (see §3.2) and find  $y^+ = 0.087$  for the DIG, larger than our  $y^+$  found for the entire map. This value is in rough agreement with the expected elemental helium abundance in the ISM ( $\sim 8.5\%$ , see Balser 2006, and references therein). We note that the derived value of  $y^+$  for the DIG-only data is strongly affected by baseline instability, which we believe is the major uncertainty contribution. In order to quantify the effect of baseline fluctuations, we subtract a 1st to 3rd-order polynomial baseline from our data. The different baseline models change the derived  $y^+$  values by up to  $\pm 0.028$ , although the higher-order baselines likely overfit the data. We show the magnitude of different baseline models in Figure 10. We therefore conservatively estimate a  $\pm 35\%$  uncertainty on all values of  $y^+$ . Due to the large uncertainty, the difference in  $y^+$  between our original averaged H $\alpha$  map and the DIG-only data is not statistically significant.



**Figure 11.** Dust temperature map derived from *IRAS* and *Herschel* data and smoothed to a resolution of  $30''$ . The contours are of integrated hydrogen RRL intensity and contour levels are at 1, 2, 4, and  $8 \text{ K km s}^{-1}$ . The green crosses indicate the location of H II regions; the size of each cross is proportional to the RRL intensity of the region.

## 6. DUST TEMPERATURE

Radiation escaping from H II regions heats the ambient medium. Although dust grains account for only  $\sim 1\%$  of the mass, they are an important coolant of the ISM since they absorb high-energy photons and re-emit thermally at IR wavelengths. Variations in the dust temperature surrounding H II regions are common. Outside the PDR of the bubble H II region RCW 120 dust temperature enhancements were found to be correlated with “holes” in the PDR, along which radiation can escape into the ISM (Anderson et al. 2012b; Anderson et al. 2015). In a study of the Galactic H II region NGC 7538 small dust temperature enhancements to the north and east of the region coincide with extended radio emission, suggesting that photons from the H II region leak preferably along these directions (Lusci et al. 2016).

We derive the dust temperature map of the W43 complex shown in Figure 11 using  $100 \mu\text{m}$  data from the Infrared Astronomical Satellite (*IRAS*; Neugebauer et al. 1984), and  $250$ ,  $350$ , and  $500 \mu\text{m}$  data obtained with the SPIRE instrument on the *Herschel* Space Observa-

tory (Griffin et al. 2010) as part of the Hi-GAL Galactic Plane Survey (Molinari et al. 2010). We assume that the dust grain flux is optically thin and follows that of a gray-body,

$$F_\nu \propto \kappa_\nu B_\nu(T_d) N_{\text{dust}}, \quad (6)$$

where  $F_\nu$  is the flux density per beam,  $\kappa_\nu$  is the dust opacity in  $\text{cm}^2 \text{g}^{-1}$ ,  $B_\nu(T_d)$  is the Planck function for dust temperature  $T_d$  at frequency  $\nu$  (given in  $\text{Jy sr}^{-1}$ ), and  $N_{\text{dust}}$  is the dust column density in  $\text{cm}^{-2}$  (Anderson et al. 2012b). We re-grid the IR maps to the resolution of the GDIGS data ( $30''$ ) and fit the gray-body model pixel-by-pixel, assuming that the functional form for the dust opacity is  $\kappa_\nu = 0.1 (\nu/10^{12} \text{ Hz})^\beta$  (Beckwith et al. 1990), with a dust emissivity index of  $\beta = 2$ . We also fit the column density of gas and dust simultaneously with the dust temperatures, assuming

$$N_{\text{H}} = R \frac{F_\nu}{2.8 m_{\text{H}} \kappa_\nu B_\nu(T_d) \Omega}, \quad (7)$$

where  $R = 100$  is the gas-to-dust mass ratio,  $m_{\text{H}}$  is the mass of a hydrogen atom in g (with the factor of 2.8

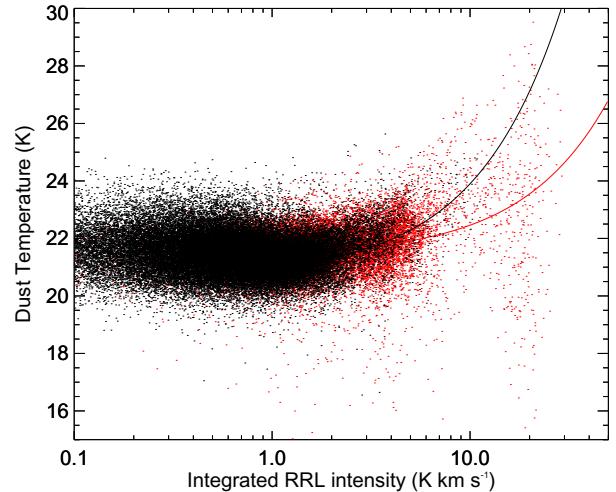
accounting for heavier elements), and  $\Omega$  is the beam size in sr (Anderson et al. 2012b).

The average dust temperature in the field of view is  $\sim 21$  K, with no significant large-scale trends (see Figure 11). The two most luminous H II regions on the map, W43 and G29, both have generally higher dust temperatures than average, but also show cold patches. These cold areas correspond to larger column densities, indicating cold dust clumps foreground to the H II regions (see Bally et al. 2010). Apart from W43 and G29, there are a number of H II regions with derived dust temperatures lower than average and large column densities in combination with above-average  $T_d$  values from their surroundings (e.g., the H II region complex G032.110+00.090 at  $\ell \approx 32.1^\circ$ ,  $b \approx 0.1^\circ$  and G031.401–00.259 at  $\ell \approx 31.4^\circ$ ,  $b \approx -0.3^\circ$ ). This suggests that these regions are still embedded in molecular clouds, but that a fraction of their radiation escapes into the ISM and heats the diffuse gas. Many H II regions in Figure 11 have the same dust temperature as their surroundings. Most of these regions have integrated RRL intensities  $< 1 \text{ K km s}^{-1}$ , which may indicate that their luminosities are insufficient to have a measurable effect on the derived dust temperature.

We show the pixel-by-pixel correlation of integrated RRL intensity and dust temperature in Figure 12. We fit the correlation between RRL emission and  $T_d$  separately for directions coincident with discrete H II regions and the diffuse ISM, and find that there is a correlation between dust temperature and RRL intensity at integrated intensities greater than  $\sim 1 \text{ K km s}^{-1}$ . This may suggest that the same radiation field that heats the dust also maintains the ionization of the DIG. The slope of the correlation is steeper for diffuse locations. The low dust temperatures ( $T_d < 20$  K) at RRL intensities  $> 10 \text{ K km s}^{-1}$  are mostly due to the cold dust clumps assumed to be foreground to W43 and G29. Since large differences in dust heating mechanisms within H II regions and the diffuse gas are unlikely, these cold foreground clumps may be the principal cause for the distinct slopes, since they artificially introduce low dust temperatures (from the foreground clumps) associated with large RRL intensities (from the background star-forming regions). There is no statistically significant correlation between RRL intensity and dust temperature at lower RRL intensities.

## 7. SUMMARY

Using data from the GDIGS survey, an ongoing RRL survey at C-band (4–8 GHz) with the GBT, we perform a high-resolution spectroscopic study of the diffuse ionized gas in the Galactic plane. Here we present our



**Figure 12.** Pixel-by-pixel correlation between integrated hydrogen RRL intensity and dust temperature. The red dots mark emission coincident with discrete H II regions, while the black dots mark diffuse gas emission. The red and black lines are linear fits of the form  $T_d = a + b I_{\text{RRL}}$  to the data, where  $T_d$  is the dust temperature in K and  $I_{\text{RRL}}$  is the integrated RRL intensity in  $\text{K km s}^{-1}$ . We only fit data points with  $I_{\text{RRL}} > 1 \text{ K km s}^{-1}$ . Fit parameters are  $a = 21.37 \pm 0.01$ ,  $b = 0.109 \pm 0.003$  for the correlation between H II region RRL emission and  $T_d$ , and  $a = 20.68 \pm 0.02$ ,  $b = 0.322 \pm 0.009$  for the diffuse gas. Note that the GBT beam size is larger than the pixel size and therefore the intensities of pixels within one beam area are not truly independent of each other.

GDIGS data of W43, one of the most active star formation regions in the inner Galaxy. There are roughly 300 H II regions and H II region candidates within  $\sim 1.5^\circ$  of W43, making the area an ideal target to study the impact of H II regions on the surrounding ISM.

Using our fully-sampled RRL observations from the GDIGS survey and data from the *WISE* Catalog of Galactic H II Regions (Anderson et al. 2014), we generate DIG-only maps to analyze the properties of the diffuse gas unaffected by emission from discrete H II regions. We find that the emission from the DIG is strongest near W43 and G29, but there is also significant diffuse RRL emission at large distances from the Galactic plane far from any known active star-forming regions. One such emission region is  $\sim 1^\circ$  to the Galactic south of W43, which may be associated with the “worm-ionized medium” postulated by Heiles et al. (1996). This structure may indicate that there exist low-density pathways perpendicular to the Galactic plane, allowing Lyman continuum photons to travel the distances necessary to account for the observed DIG scale height of  $\sim 1$  kpc.

The intensity of the DIG emission can be estimated by a simple empirical model based on the locations, angular sizes, and RRL intensities of the H II regions in the field of view. As a first-order approximation, the model assumes that all DIG emission is due to photons leaking from H II regions. We find that the correlation between model emission and our GDIGS RRL data is improved by assuming a power-law decrease in hydrogen RRL emission beyond the H II region PDRs rather than an exponential decrease.

Most of the RRL emission near W43 and G29 is found at two distinct velocities: 40 and 100 km s<sup>-1</sup>. The distance determination for the ionized gas responsible for these two velocity components is hampered by their location in the inner Galaxy and possibly by strong non-circular motions near the end of the bar where W43 is located. Parallax measurements place W43 at a distance of  $\sim 5.5$  kpc (Zhang et al. 2014). Since the 100 km s<sup>-1</sup> gas component is spatially associated with W43 we argue that it is at roughly the same distance. The correlation between the observed and the modeled 40 km s<sup>-1</sup> RRL emission is poor, suggesting that the 40 km s<sup>-1</sup> component may partially maintain its ionization from H II regions at larger velocities. We therefore hypothesize that the two velocity components are interacting, placing the 40 km s<sup>-1</sup> emission at a distance of  $\sim 6$  kpc as well.

We use the helium-to-hydrogen ionic abundance ratio,  $y^+$ , as a tracer for the hardness of the radiation field, however, we do not find a statistically significant difference in  $y^+$  between our original averaged H $\alpha$  map

and the DIG-only data. The correlation between dust temperature and RRL intensity may suggest that the same radiation field that heats the dust also maintains the ionization of the DIG.

We thank the anonymous referee for useful comments and suggestions, which improved the quality of this manuscript. The Green Bank Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. *Herschel* is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA. We thank West Virginia University for its financial support of GBT operations, which enabled some of the observations for this project. This work is supported by NSF grants AST1812639 to LDA and AST1714688 to TMB. BL thanks West Virginia University for hosting him as a postdoctoral researcher during completion of this work. BL was partially supported by the National Natural Science Foundation of China (Grant #11503036).

*Facility:* Green Bank Telescope. *Software:* GBTIDL (Marganian et al. 2006), gbtgriddler (<https://github.com/GreenBankObservatory/gbtgriddler>).

## REFERENCES

Alves, M. I. R., Calabretta, M., Davies, R. D., et al. 2015, MNRAS, 450, 2025, doi: [10.1093/mnras/stv751](https://doi.org/10.1093/mnras/stv751)

Anderson, L. D., Bania, T. M., Balser, D. S., & Rood, R. T. 2011, ApJS, 194, 32, doi: [10.1088/0067-0049/194/2/32](https://doi.org/10.1088/0067-0049/194/2/32)

Anderson, L. D., Bania, T. M., Balser, D. S., & Rood, R. T. 2012a, ApJ, 754, 62, doi: [10.1088/0004-637X/754/1/62](https://doi.org/10.1088/0004-637X/754/1/62)

Anderson, L. D., Bania, T. M., Balser, D. S., et al. 2014, ApJS, 212, 1, doi: [10.1088/0067-0049/212/1/1](https://doi.org/10.1088/0067-0049/212/1/1)

Anderson, L. D., Hough, L. A., Wenger, T. V., Bania, T. M., & Balser, D. S. 2015, ApJ, 810, 42, doi: [10.1088/0004-637X/810/1/42](https://doi.org/10.1088/0004-637X/810/1/42)

Anderson, L. D., Zavagno, A., Deharveng, L., et al. 2012b, A&A, 542, 27, doi: [10.1051/0004-6361/201117283](https://doi.org/10.1051/0004-6361/201117283)

Anderson, L. D., Deharveng, L., Zavagno, A., et al. 2015, ApJ, 800, 101, doi: [10.1088/0004-637X/800/2/101](https://doi.org/10.1088/0004-637X/800/2/101)

Bally, J., Anderson, L. D., Battersby, C., et al. 2010, A&A, 518, L90, doi: [10.1051/0004-6361/201014596](https://doi.org/10.1051/0004-6361/201014596)

Balser, D. S. 2006, AJ, 132, 2326, doi: [10.1086/508515](https://doi.org/10.1086/508515)

Balser, D. S., Rood, R. T., Bania, T. M., & Anderson, L. D. 2011, ApJ, 738, 27, doi: [10.1088/0004-637X/738/1/27](https://doi.org/10.1088/0004-637X/738/1/27)

Bania, T. M., Anderson, L. D., & Balser, D. S. 2012, ApJ, 759, 96, doi: [10.1088/0004-637X/759/2/96](https://doi.org/10.1088/0004-637X/759/2/96)

Bania, T. M., Anderson, L. D., Balser, D. S., & Rood, R. T. 2010, ApJL, 718, L106, doi: [10.1088/2041-8205/718/2/L106](https://doi.org/10.1088/2041-8205/718/2/L106)

Beckwith, S. V. W., Sargent, A. I., Chini, R. S., & Guesten, R. 1990, AJ, 99, 924, doi: [10.1086/115385](https://doi.org/10.1086/115385)

Benjamin, R. A., Churchwell, E., Babler, B. L., et al. 2005, ApJL, 630, L149, doi: [10.1086/491785](https://doi.org/10.1086/491785)

Berkhuijsen, E. M., Mitra, D., & Mueller, P. 2006, Astronomische Nachrichten, 327, 82, doi: [10.1002/asna.200510488](https://doi.org/10.1002/asna.200510488)

Beuther, H., Tackenberg, J., Linz, H., et al. 2012, ApJ, 747, 43, doi: [10.1088/0004-637X/747/1/43](https://doi.org/10.1088/0004-637X/747/1/43)

Blum, R. D., Damineli, A., & Conti, P. S. 1999, AJ, 117, 1392, doi: [10.1086/300791](https://doi.org/10.1086/300791)

Boulares, A., & Cox, D. P. 1990, ApJ, 365, 544, doi: [10.1086/169509](https://doi.org/10.1086/169509)

Cesaroni, R., Churchwell, E., Hofner, P., Walmsley, C. M., & Kurtz, S. 1994, A&A, 288, 903

Condon, J. J., & Ransom, S. M. 2016, Essential Radio Astronomy (Princeton University Press)

Deharveng, L., Zavagno, A., Schuller, F., et al. 2009, A&A, 496, 177, doi: [10.1051/0004-6361/200811337](https://doi.org/10.1051/0004-6361/200811337)

Dennison, B., Simonetti, J. H., & Topasna, G. A. 1998, PASA, 15, 147, doi: [10.1071/AS98147](https://doi.org/10.1071/AS98147)

Domgörgen, H., & Mathis, J. S. 1994, ApJ, 428, 647, doi: [10.1086/174275](https://doi.org/10.1086/174275)

Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, RMxAA, 53, 385. <https://arxiv.org/abs/1705.10877>

Gaustad, J. E., McCullough, P. R., Rosing, W., & Van Buren, D. 2001, PASP, 113, 1326, doi: [10.1086/323969](https://doi.org/10.1086/323969)

Green, D. A. 2014, Bulletin of the Astronomical Society of India, 42, 47

Green, G. M., Schlafly, E. F., Zucker, C., Speagle, J. S., & Finkbeiner, D. P. 2019, arXiv e-prints, arXiv:1905.02734. <https://arxiv.org/abs/1905.02734>

Griffin, M. J., Abergel, A., Abreu, A., et al. 2010, A&A, 518, L3, doi: [10.1051/0004-6361/201014519](https://doi.org/10.1051/0004-6361/201014519)

Haffner, L. M., Reynolds, R. J., Tufte, S. L., et al. 2003, ApJS, 149, 405, doi: [10.1086/378850](https://doi.org/10.1086/378850)

Haffner, L. M., Dettmar, R.-J., Beckman, J. E., et al. 2009, RvMP, 81, 969, doi: [10.1103/RevModPhys.81.969](https://doi.org/10.1103/RevModPhys.81.969)

Heiles, C., Reach, W. T., & Koo, B.-C. 1996, ApJ, 466, 191, doi: [10.1086/177503](https://doi.org/10.1086/177503)

Hoopes, C. G., & Walterbos, R. A. M. 2003, ApJ, 586, 902

Hoyle, F., & Ellis, G. R. A. 1963, AJP, 16, 1, doi: [10.1071/PH630001](https://doi.org/10.1071/PH630001)

Jackson, J. M., Rathborne, J. M., Shah, R. Y., et al. 2006, ApJS, 163, 145, doi: [10.1086/500091](https://doi.org/10.1086/500091)

Krishnarao, D., Haffner, L. M., Benjamin, R. A., Hill, A. S., & Barger, K. A. 2017, ApJ, 838, 43, doi: [10.3847/1538-4357/aa63e6](https://doi.org/10.3847/1538-4357/aa63e6)

Liu, B., McIntyre, T., Terzian, Y., et al. 2013, AJ, 146, 80, doi: [10.1088/0004-6256/146/4/80](https://doi.org/10.1088/0004-6256/146/4/80)

Luisi, M., Anderson, L. D., Balser, D. S., Bania, T. M., & Wenger, T. V. 2016, ApJ, 824, 125, doi: [10.3847/0004-637X/824/2/125](https://doi.org/10.3847/0004-637X/824/2/125)

Luisi, M., Anderson, L. D., Balser, D. S., Wenger, T. V., & Bania, T. M. 2017, ApJ, 849, 117, doi: [10.3847/1538-4357/aa8fd2](https://doi.org/10.3847/1538-4357/aa8fd2)

Luisi, M., Anderson, L. D., Bania, T. M., et al. 2018, PASP, 130, 084101, doi: [10.1088/1538-3873/aac8e9](https://doi.org/10.1088/1538-3873/aac8e9)

Luisi, M., Anderson, L. D., Liu, B., Anish Roshi, D., & Churchwell, E. 2019, ApJS, 241, 2, doi: [10.3847/1538-4365/aaf6a5](https://doi.org/10.3847/1538-4365/aaf6a5)

Madsen, G. J., Reynolds, R. J., & Haffner, L. M. 2006, ApJ, 652, 401, doi: [10.1086/508441](https://doi.org/10.1086/508441)

Marganian, P., Garwood, R. W., Braatz, J. A., Radziwill, N. M., & Maddalena, R. J. 2006, in Astronomical Society of the Pacific Conference Series, Vol. 351, Astronomical Data Analysis Software and Systems XV, ed. C. Gabriel, C. Arviset, D. Ponz, & S. Enrique, 512. <http://adsabs.harvard.edu/abs/2006ASPC..351..512M>

McKee, C. F., & Williams, J. P. 1997, ApJ, 476, 144

Molinari, S., Swinyard, B., Bally, J., et al. 2010, A&A, 518, L100, doi: [10.1051/0004-6361/201014659](https://doi.org/10.1051/0004-6361/201014659)

Neugebauer, G., Habing, H. J., van Duinen, R., et al. 1984, ApJL, 278, L1, doi: [10.1086/184209](https://doi.org/10.1086/184209)

Nguyen Luong, Q., Motte, F., Schuller, F., et al. 2011, A&A, 529, A41

Oey, M. S., & Kennicutt, R. C. J. 1997, MNRAS, 291, 827

Osterbrock, D. 1989, Astrophysics of gaseous nebulae and active galactic nuclei (University Science Books)

Ott, M., Witzel, A., Quirrenbach, A., et al. 1994, A&A, 284, 331

Peimbert, M., Rodriguez, L. F., Bania, T. M., Rood, R. T., & Wilson, T. L. 1992, ApJ, 395, 484, doi: [10.1086/171668](https://doi.org/10.1086/171668)

Reich, P., & Reich, W. 1986, A&AS, 63, 205

Reich, W., Reich, P., & Fuerst, E. 1990, A&AS, 83, 539

Reynolds, R. J. 1984, ApJ, 282, 191, doi: [10.1086/162190](https://doi.org/10.1086/162190)

Reynolds, R. J. 1991, in IAU Symposium, Vol. 144, .., 67–76

Reynolds, R. J., Tufte, S. L., Haffner, L. M., Jaehnig, K., & Percival, J. W. 1998, PASA, 15, 14, doi: [10.1071/AS98014](https://doi.org/10.1071/AS98014)

Roshi, D. A., & Anantharamaiah, K. R. 2000, ApJ, 535, 231, doi: [10.1086/308813](https://doi.org/10.1086/308813)

Stil, J. M., Taylor, A. R., Dickey, J. M., et al. 2006, AJ, 132, 1158, doi: [10.1086/505940](https://doi.org/10.1086/505940)

Wang, Q. D., Dong, H., Cotera, A., et al. 2010, MNRAS, 402, 895, doi: [10.1111/j.1365-2966.2009.15973.x](https://doi.org/10.1111/j.1365-2966.2009.15973.x)

Weber, J. A., Pauldrach, A. W. A., & Hoffmann, T. L. 2019, A&A, 622, A115, doi: [10.1051/0004-6361/201832649](https://doi.org/10.1051/0004-6361/201832649)

Wegg, C., Gerhard, O., & Portail, M. 2015, MNRAS, 450, 4050, doi: [10.1093/mnras/stv745](https://doi.org/10.1093/mnras/stv745)

Wenger, T. V., Bania, T. M., Balser, D. S., & Anderson, L. D. 2013, ApJ, 764, 34, doi: [10.1088/0004-637X/764/1/34](https://doi.org/10.1088/0004-637X/764/1/34)

Wood, K., & Mathis, J. S. 2004, MNRAS, 353, 1126, doi: [10.1111/j.1365-2966.2004.0784](https://doi.org/10.1111/j.1365-2966.2004.0784)

Zhang, B., Moscadelli, L., Sato, M., et al. 2014, ApJ, 781, 89, doi: [10.1088/0004-637X/781/2/89](https://doi.org/10.1088/0004-637X/781/2/89)

Zurita, A., Rozas, M., & Beckman, J. E. 2000, A&A, 363, 9