# The W51 GMC in 3D: The DGMF says that SF is done in W51 B but booming in W51 A. Also, ULIRGs are missing ionized gas?

Adam Ginsburg[1], John Bally[2], Cara Battersby[3], Jeremy Darling[2], Allison Youngblood[2], Mayra Santos[4], Hector Arce[4], Erik Rosolowsky[5],

Adam.G.Ginsburg@gmail.com

[2] *CASA, University of Colorado, 389-UCB, Boulder, CO 80309*

[1] *European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-85748 Garching bei München, Germany*

[3] *CfA*

[4] *Yale*

[5] *Edmonton*

## ABSTRACT

We present new 2 cm and 6 cm maps of $H_2CO$, radio recombination lines, and the radio continuum in the W51 star forming complex acquired with Arecibo and the Green Bank Telescope at $\sim 50''$ resolution. We perform a detailed analysis of the data and data processing with an emphasis on ensuring the reliability of the data. The data set has been made public at `http://dx.doi.org/10.7910/DVN/26818`.

We present an analysis of the three-dimensional structure of the W51 region, determining the relative line-of-sight positions of molecular and ionized gas. We show evidence that the W51 C supernova remnant is interacting with the W51 B molecular filament. We argue that the W51 A and B star-forming regions are physically associated and interacting, supporting the idea that the W51 A cluster formation was triggered by cloud-cloud collision.

We measure gas densities using the $H_2CO$ densitometer, finding largely consistent mass-weighted volume densities $n \gtrsim 10^4$ cm$^{-3}$ throughout the W51 GMC, with higher density $n \gtrsim 10^5$ cm$^{-3}$ gas associated with the proto-clusters. We present continuous measurements of the dense gas mass fraction (DGMF) over the range $10^4$ cm$^{-3}$ $<$ $n(H_2) < 10^6$ cm$^{-3}$. We measure a high DGMF in W51 Main, indicating that nearly all of its mass should form stars, and a very low DGMF in W51 B, suggesting that its star-forming lifetime has come to an end and it is being disrupted by stellar feedback.

We did not detect *any* $H_2CO$ emission throughout the W51 GMC; all gas dense enough to emit under normal conditions is in front of bright continuum sources. This nondetection implies that the $H_2CO$ emission detected in other galaxies, e.g. Arp 220, comes from high-density gas that is not directly affiliated with forming massive stars, i.e. the entire ISM of these galaxies is very dense.

## 1. Introduction

To do. To-do items are coded in red. Compiled on 2014/08/03 at 18:05:14.

Massive star clusters are among the most visually outstanding features in the night sky. In other galaxies, they are useful probes of the star formation history and can be individually identified and measured (Bastian 2008). Locally, they are the essential laboratories in which we can study the formation of massive stars (Davies 2012).

In order to utilize these clusters as laboratories, we need to understand their formation in detail. Clusters are often assumed and measured to be coeval (e.g. Kudryavtseva et al. 2012), but uncertainties remain (Beccari et al. 2010). In the most massive clusters, there are predictions that multiple generations or an extended generation of stars should form prior to gas expulsion because the gas will remain gravitationally bound (Bressert et al. 2012). Feedback from and within young massive clusters is an active field of numerical study (Rogers & Pittard 2013; Dale et al. 2013, 2012; Dale & Bonnell 2008; Dale et al. 2005; Parker & Dale 2013; Myers et al. 2014; Krumholz et al. 2014) While only 5-35% of all stars form in bound clusters (Kruijssen 2012), these clusters form the basis of our understanding of stars and stellar evolution [citation needed], and understanding their formation is therefore crucial.

The results of cluster formation may be decided before the first stars are formed. The starless initial conditions of massive clusters have not yet been definitively observed (Ginsburg et al. 2012) though there are viable candidates such as G0.253+0.016 (Longmore et al. 2012). The initial conditions for star formation on any scale are clearly turbulent. However, there is no evidence whether these initial conditions differ in any qualitative way from turbulence in local, low-mass star-forming regions.

This paragraph may need expansion or removal. See Section 6.2 Star formation appears to occur only above a set volume density threshold in molecular gas, specifically $n(\mathrm{H_2}) \gtrsim 10^4$ cm$^{-3}$ (Lada et al. 2010, , who advocate a column density threshold corresponding to this density). This threshold is contentious (Burkert & Hartmann 2012; Clark & Glover 2013), but it can only be evaluated when measurements of density and column density are simultaneously available.

### 1.1. W51

The W51 cloud complex (Figure 1), containing the two W51 massive proto-cluster candidates from (Ginsburg et al. 2012), is located at $\ell \sim 49, b \sim -0.3$, very near the Galactic midplane[1] at a distance of 5.1 kpc (Sato et al. 2010). It is a well-known and thoroughly studied collection of clouds massing $M > 10^6 M_\odot$ (Carpenter & Sanders 1998; Bieging et al. 2010; Kang et al. 2010;

[1]The midplane at $d = 5.1$ kpc is offset approximately -0.22 to -0.33 degrees from $b = 0$ depending on our solar system's height above the midplane, $Z_\odot = 20$ pc or 30 pc, respectively (Reed 2006; Joshi 2007).

Parsons et al. 2012). The radio-bright regions are generally known as W51 A to the east, W51 B to the west, and W51 C for the southern component, known to trace a supernova remnant (Koo et al. 1995; Brogan et al. 2000, 2013).

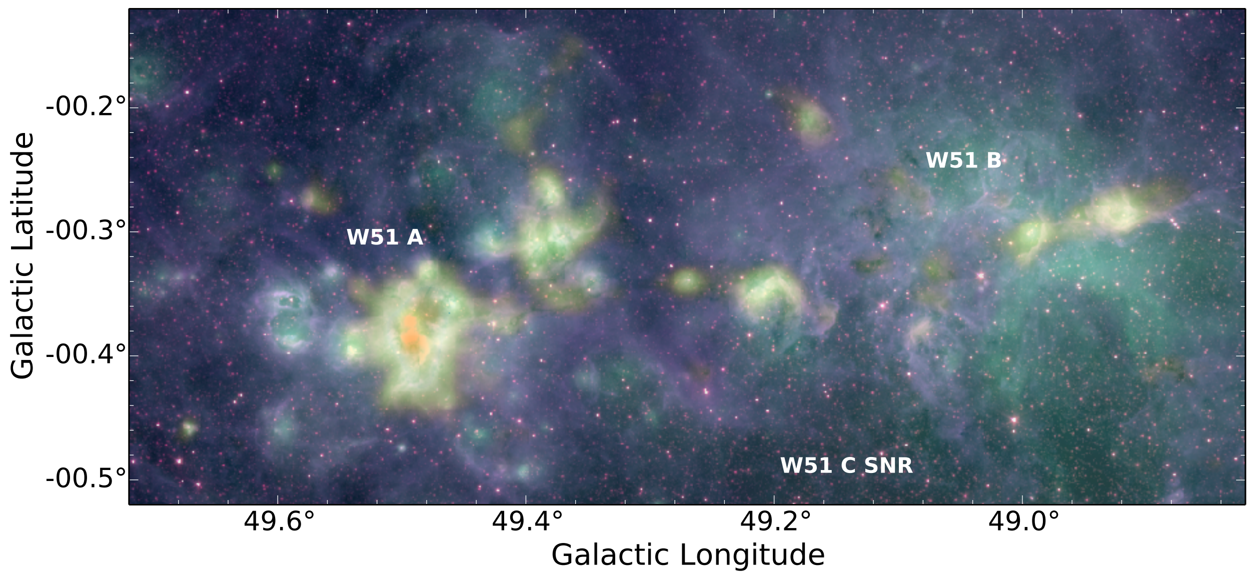



Fig. 1.— A color composite of the W51 region with major regions, W51 A, B, and C, labeled. The blue, green, and red colors are WISE bands 1, 3, and 4 (3.4, 12, and 22 $\mu$m) respectively. The yellow-orange semitransparent layer is from the Bolocam 1.1 mm Galactic Plane Survey data (Aguirre et al. 2011; Ginsburg et al. 2013a). Finally, the faint whitish haze filling in most of the image is from a 90 cm VLA image by Brogan et al. (2013), which primarily traces the W51 C supernova remnant.

### 1.2. Formaldehyde

Formaldehyde ($H_2CO$) has been recognized as a useful probe of physical conditions in the molecular interstellar medium for decades (Mangum & Wootten 1993). The centimeter lines, $H_2CO$ $1_{10} - 1_{11}$ (6.2 cm, 4.82966 GHz) and $2_{11} - 2_{12}$ (2.1 cm, 14.48848 GHz), have a peculiar excitation process in which collisions overpopulate the lower of the two $K_c$ rotational states. Because the $1_{10} - 1_{11}$ and $2_{11} - 2_{12}$ level pairs populate differently depending on the volume density of the colliding partner (generally $H_2$), their ratio is sensitive to the local gas volume density.

$H_2CO$ $1_{10} - 1_{11}$ has been observed in the W51 main region with the VLA (Martin-Pintado et al. 1985a) and Westerbork (Arnal & Goss 1985), and this data was used to gain some early constraints on the geometry of the region (e.g. Carpenter & Sanders 1998). Henkel et al. (1980) presented observations of the $1_{10} - 1_{11}$ and $2_{11} - 2_{12}$ lines, and Martin-Pintado et al. (1985b) presented single-dish mapping observations of the $H_2CO$ $2_{11} - 2_{12}$ line toward the W51 Main region, but both treated the region as a single-density structure.

## 2. Observations & Data Reduction

### 2.1. GBT 2 cm

The W51 survey was performed in September 2011 and 2012. The GBT data were taken as part of program AGBT10B/019; the raw data are available from the NRAO archive (`https://archive.nrao.edu/`). The data presented in this paper include sessions 10, 11, 14, 16, 17, 20, 21, and 22; the other sessions from this project include maps of outer galaxy regions and a single-pointing survey of Galactic plane sources that will be presented in another paper.

Data were taken in on-the-fly mode with the GBT Ku-band dual-beam system. Cross-hatched north-south and east-west maps were created in Galactic coordinates.

The data was reduced using custom-made scripts based off of both GBTIDL's mapping routines by Glen Langston (`https://safe.nrao.edu/wiki/bin/view/Kbandfpa/KfpaPipelineHowTo`) and Phil Perillat's AOIDL routines. The code is available at `https://github.com/keflavich/sdpy`. The data reduction code and workflow is included in a corresponding git repository: `https://github.com/keflavich/w51_singledish_h2co_maps`.

The individual spectral were calibrated using a noise calibration diode as usual. The first and last scans of each observation were used as off positions, and the background level to be subtracted off of the continuum was determined by linearly interpolating between these scans.

The GBT data infrequently exhibited major data artifacts; a key phase of the reduction process was excising bad observations, which were usually isolated to a single component of the backend.

The Green Bank data have a main beam efficiency $\eta_{MB} = 0.886$, or a gain of 1.98 K/Jy assuming a 51 $''$ beam (see Mangum et al. 2013, for additional discussion). The GBT data were also corrected for atmospheric opacity using Ron Maddalena's `getForecastValues`[2] with a typical zenith optical depth $\tau_z \approx 0.02$, so this correction was never more than $\sim 5\%$.

Typical noise levels were $\sim 10 - 20$ mK per 1 km s$^{-1}$ channel; the levels vary across the map. See Section 3.1 for details.

### 2.2. Arecibo 6 cm

The Arecibo data were taken as part of project A2705 over the course of 4 nights, September 10, 11, 12, and 15 2012. On the first night, September 10 2012, a significant fraction of the data

[2]`http://www.gb.nrao.edu/~rmaddale/Weather/`

was lost due to an internal instrument error within the Mock spectrometer, which resulted in a loss of the high-resolution component of the $H_2CO$ data for that night. As a result, we have focused our study on the lower-resolution ($\sim 1$ km s$^{-1}$) data.

The fields were observed with east-west maps using the C-band receiver. No crosshatching was performed with Arecibo.

The Arecibo data reduction process for W51 presented unique challenges: at C-band, the entire region surveyed contains continuum emission, so no truly suitable 'off' position was found within the survey data. Similarly, $H_2CO$ is ubiquitous across the region, so it was necessary to 'mask out' the absorption lines when building an off position. This was done by interpolating across the line-containing region with a polynomial fit. The fits were inspected interactively and tuned to avoid over-predicting the background.

The Arecibo data were corrected to main beam brightness temperature $T_{MB}$ using a main-beam efficiency as a function of zenith angle in degrees ($za$):

$$\eta_{MB}(za) = 0.491544 + 0.00580397za - 0.000341992za^2$$

This is a fit to 5 years worth of calibration data acquired at Arecibo and assembled by Phil Perrilat.

Typical noise levels were $\sim 50$ mK per 1 km s$^{-1}$ channel. See Section 3.1 for details.

### 2.3. Mapmaking

The maps were made by computing an output grid in Galactic coordinates with $15''$ pixels and adding each spectrum to the appropriate pixel[3]. In order to avoid empty pixels and maximize the signal-to-noise, the spectra were added to the grid with a weight set from a Gaussian with $FWHM = 20''$, which effectively smooths the output images from $FWHM \approx 50''$ to $\approx 54''$. See Mangum et al. (2007) for more detail on the on-the-fly mapping technique used here.

The Arecibo data were taken at a spectral resolution[4] of 0.68 km s$^{-1}$ and the GBT at 0.25 km s$^{-1}$ resolution. The spectra were regridded onto a velocity grid from $-50$ to 150 km s$^{-1}$ with 1 km s$^{-1}$ resolution. To achieve this, they were first smoothed by a Gaussian with $FWHM = 1$ km s$^{-1}$ then downsampled appropriately.

[3]We use the term 'pixel' to refer to a square data element projected on the sky with axes in Galactic coordinates. The term 'voxel', short for 'volumeicture element' (it's not a word) is used to indicate a cubic data element, with two axes in galactic coordinates and a third in frequency or velocity

[4]For most of the map area, data is available at much higher $\lesssim 0.2$ km s$^{-1}$ resolution, but the signal-to-noise at this resolution is relatively poor, no linewidths were observed to be that narrow, and most importantly, one Arecibo data set suffered from a malfunction that allowed data at 0.68 km s$^{-1}$ resolution to be taken, but not the corresponding 0.2 km s$^{-1}$ data.

## 3. The Data

The main products of the mapping data are PPV cubes in the two $H_2CO$ lines, the integrated continuum in the 2 and 6 cm bands, and optical depth data cubes. In this section we describe these data and the systematic errors associated with them.

### 3.1. PPV cubes

The PPV cubes were created with units of brightness temperature. The Arecibo cubes have contributions from 15-20 independent spectra in each pixel, though this hit rate varies in a systematic striped pattern parallel to the Galactic plane. The small overlap regions between different maps have a significantly higher number of samples; these regions constitute a small portion of the map. The resulting noise level is RMS $\sim 50-60$ mK except toward the H II regions, where it peaks at about 400 mK. The continuum is derived by averaging line-free channels; its signal-to-noise peaks at $\sim 900$.

The GBT data were mapped in an orthogonal grid pattern, so the hit coverage is more uniform on small scales, but because of the dual-beam Ku-band system, the overall noise levels are much more patchy. Additionally, the nights with better weather yielded a lower noise level. The noise ranges from $\sim 7$ mK in the W51 Main region to $\sim 20$ mK in the westmost region. As with the Arecibo data, the H II region adds noise, but the peak noise towards an H II region is only $\sim 20$ mK. This difference is because the diffuse H II region is fainter at 2 cm. The signal-to-noise ratio in the continuum peaks at $\sim 400$.

### 3.2. Continuum Images

Our observations comprise the highest resolution wide-area maps of the W51 region in the 6 cm and 2 cm continuum that preserve large angular scale structures. They can be used to provide the zero-spacing for future and extant VLA observations and are essential for the density measurements we describe later. We therefore carefully verify the quality and calibration of the continuum data.

While the noise is nominally quite low, there are significant systematic effects visible in the continuum maps. The continuum zero-point of each GBT map was determined by assuming that the first and last scan both observed zero continuum and that the sky background can be linearly interpolated between the start and end of the observations. In general, these are good assumptions, but they leave in systematic offsets of up to $\lesssim -0.15$ K in the maps, most likely because there is a $\sim 0.15$ K mostly uniform background.

The Arecibo data appear to have smaller systematic offsets, but they are more visually pronounced because there is much more diffuse emission at 6 cm. The continuum zero-point level in the Arecibo data was set to be the 10th percentile value of each scan, which is effectively the minimum value across each scan but with added robustness against noise-generated false minima. In the eastmost and westmost blocks, this strategy was very effective, as there are clearly areas in each scan that see no continuum. However, in the central block, this approach resulted in a vertical negative filament that almost certainly represents a local minimum that should be positive. This negative filament has values $\gtrsim -0.08$ K. Given the excellent agreement between the three independently observed blocks in the overlap regions between these blocks, it is clear that the continuum is reliable above $\gtrsim 0.5$ K, which is the entire regime in which it is a significant contributor to the total background emission (at lower levels, the CMB is dominant).

#### *3.2.1. Comparison between GBT and GPA data*

The Galactic Plane "A" survey (Langston et al. 2000) covered the Galactic plane at 14.35 GHz using the NRAO 300 ft telescope, with a reported FWHM beam size of $6.6'$. We compared our GBT continuum observations to theirs in order to determine whether a significant DC component is missing from our data. Because the GPA used 10 deg long scans in Galactic latitude, it should fully recover all diffuse Galactic Plane emission. In the released brightness temperature maps, brightness down to a scale of 1.5 deg is recovered. However, because the GPA data undersampled the sky (its $5'$ steps between scans were larger than the FWHM of the 14.35 GHz beam), point source fluxes in the GPA are underestimated by 19% and flux on small angular scales may be unreliable.

In order to perform the comparison, we first had to correct for an offset between the GPA and GBT data. We used the Image Registration toolkit (`http://image-registration.rtfd.org`) to measure the offset between the continuum images using a cross-correlation technique. The GBT and Arecibo data match to within $4''$, while the GPA data showed a $4'$ offset in longitude and $1'$ in latitude. We then resampled the GPA image onto the GBT grid using cubic spline interpolation, then smoothed both data sets to $9.5'$. There are image artifacts (particularly vertical streaking) in the GPA data that are diminished by this large smoothing kernel.

We compared the surface brightness in the GPA and GBT data, and found that the GPA data was $\sim 0.2$ K brighter than the GBT in the diffuse portion of the W51 Main region; the offset is not consistent with a purely multiplicative offset (Figure 3). The GBT observed the W51 Main peak to be moderately brighter, which is likely a result of the sparse sampling in the GPA. The morphological agreement between the maps is imperfect, perhaps in part because of the small area mapped in our GBT data, though there also appears to be vertical (along a line of constant longitude) stretching of the W51 main source in the unsmoothed GPA data that is not consistent with the GBT observations.

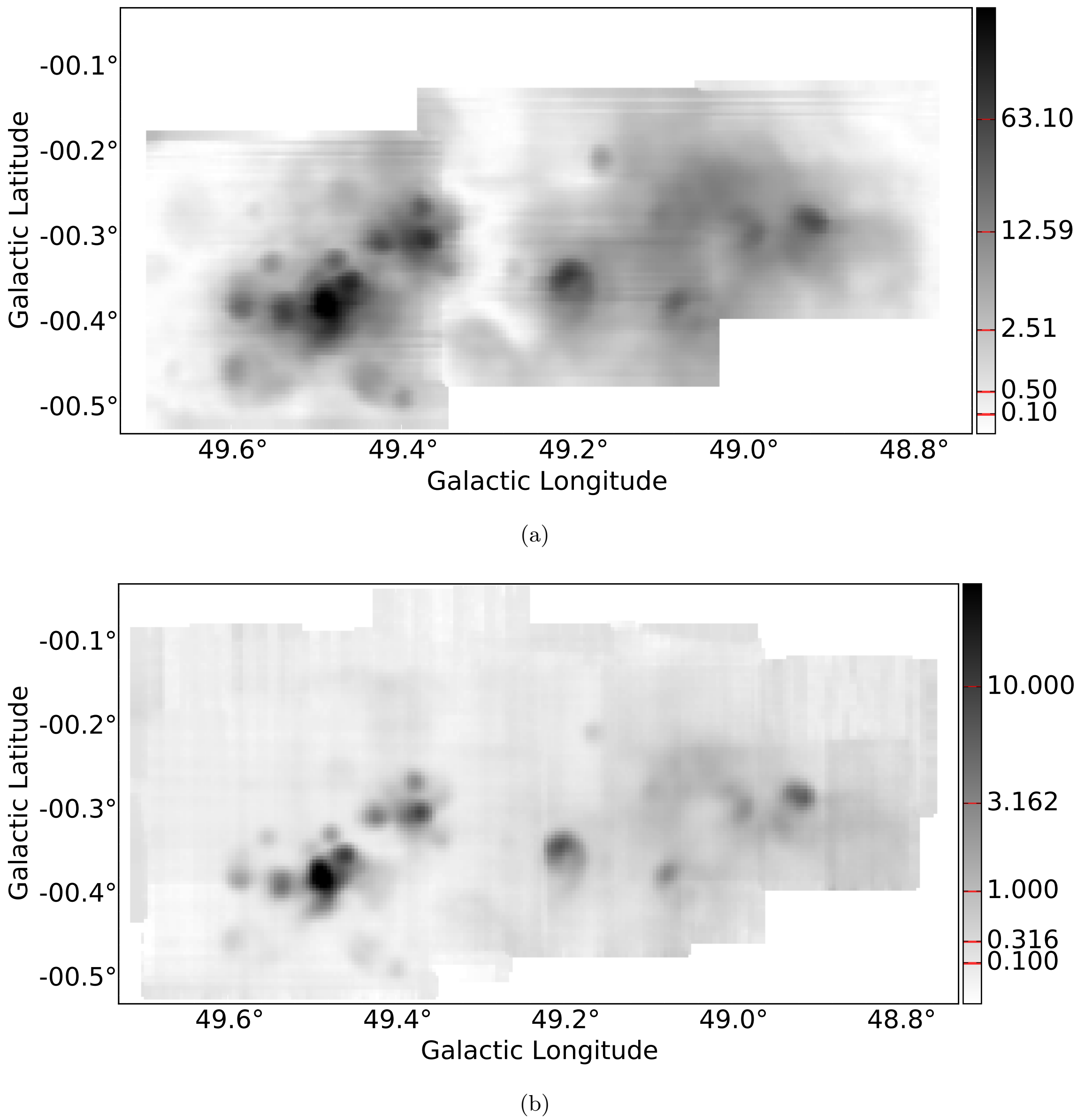



Fig. 2.— Continuum images of the 6 cm Arecibo data (left) and 2 cm GBT data (right)

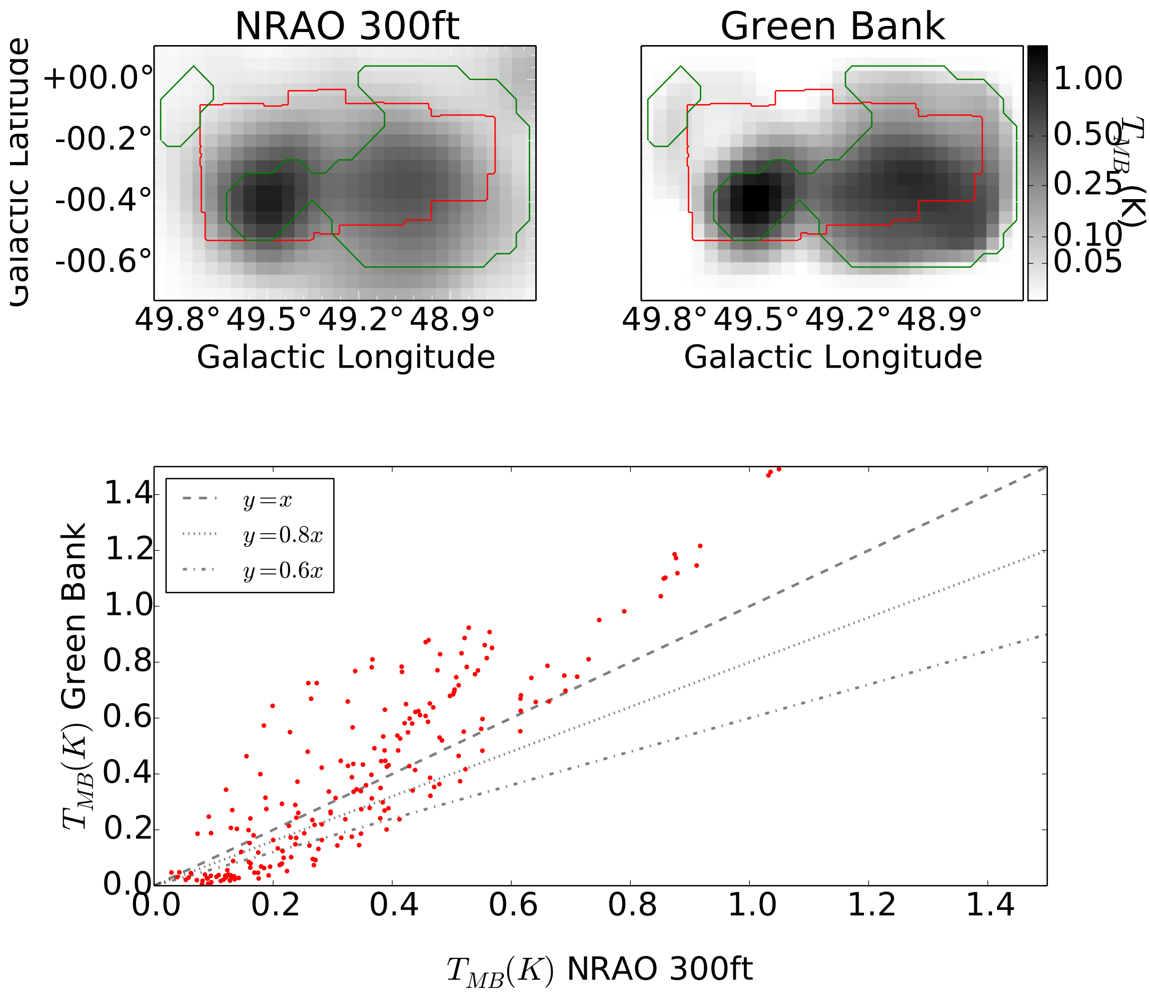



Fig. 3.— Comparison between the GBT and NRAO 300 ft (Langston et al. 2000) data. (top left) NRAO 300 ft 2 cm map (top right) Arecibo 6 cm map of the same region smoothed to about 8.9′. The colorbar applies to both figures, showing brightness temperature units in K. The red contours in both figures show the region observed by Green Bank; flux outside of those boundaries is extrapolated. The green contours show the region where $T_B$(GBT)> $T_B$(300 ft). (bottom) Plot of the GBT vs the 300 ft surface brightness measurements. The large red dots show the region within the red contours.

### *3.2.2. Comparison between Arecibo and Urumqi data*

We compare the 6 cm continuum to the Urumqi 25m data from Sun et al. (2007) and Sun et al. (2011a). Figure 4 shows the comparison of the Urumqi data and the Arecibo continuum data smoothed to 9.5′ resolution. The Arecibo data appears to be systematically brighter than the Urumqi data by about 40%. This systematic offset is worrisome for spectral index measurements, but should have little effect on the $H_2CO$ analysis. This offset appeared upon reanalysis in July 2014. It definitely, 100% certainly, did not exist the last time I ran the *exact* same analysis on the same data with the same code. I cannot track down this offset but it will probably keep me awake at nights for months. The increase from before is about 25%, which one will note is nowhere near $1/\eta_{MB} = 2$, so I really have no explanation.

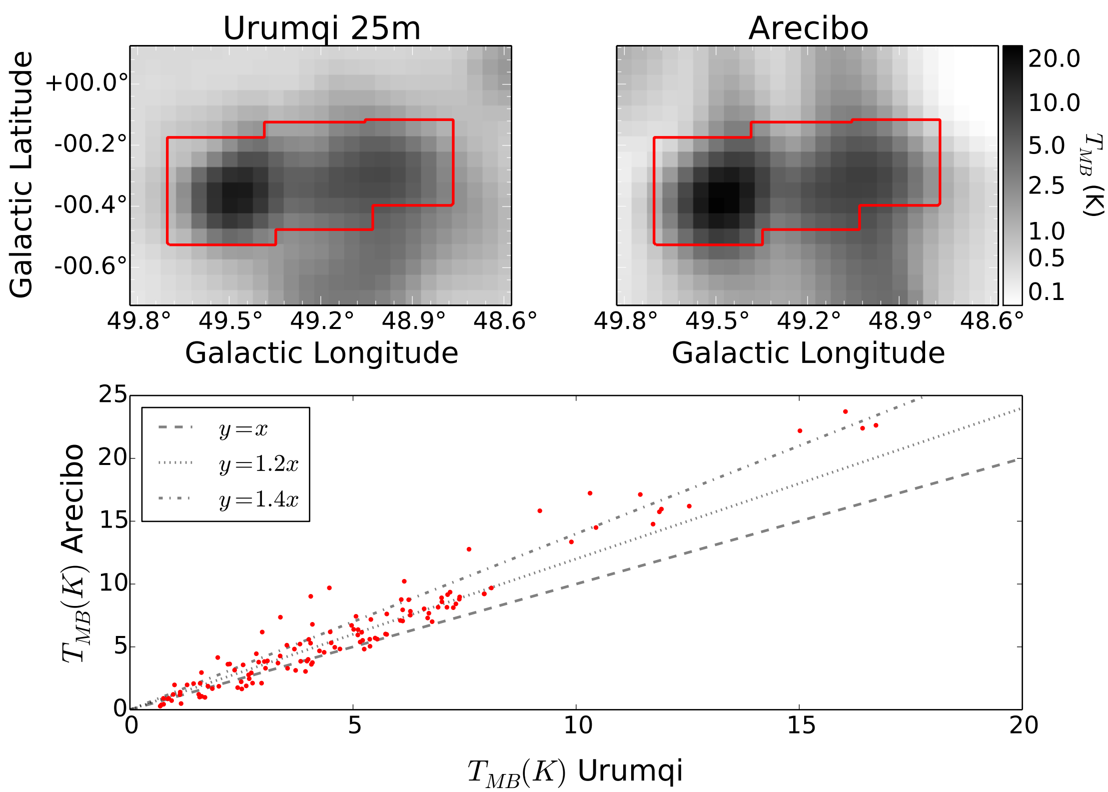



Fig. 4.— Comparison between the Arecibo and Urumqi 25m (Sun et al. 2011b) data. (top left) Urumqi 6 cm map of the W51 region. (top right) Arecibo 6 cm map of the same region smoothed to the 9.5′ resolution of the Urumqi data set. The colorbar applies to both figures, showing brightness temperature units in K. The red contours in both figures show the region observed by Arecibo; flux outside of those boundaries is extrapolated. (bottom) Plot of the Arecibo vs the Urumqi surface brightness measurements. The large red dots show the region within the red contours.

#### *3.2.3. Comparison of GBT and Arecibo data*

In order to compare the Green Bank and Arecibo continuum data, we converted the brightness temperature maps to Janskys assuming a beam FWHM of $50''$ for both surveys and central frequencies of 4.8 and 14.5 GHz for Arecibo and Green Bank respectively. Measured beam widths for both telescopes were $\sim 49 - 54''$, so the relative error from assuming the same beam size should be $\lesssim 10\%$. In this section, the target frequencies are referred to as $S_{5GHz}$ and $S_{15GHz}$ for brevity.

The data are well-correlated, with $S_{5GHz} \sim 1.4 S_{15GHz}$ ($S_{15GHz} \sim 0.7 S_{5GHz}$; Figure 5), consistent with a spectral index $\alpha_\nu = -0.3$ slightly steeper than usually observed for optically thin brehmsstrahlung and consistent with there being some contribution from synchrotron emission. The lower-brightness regions have a lower $S_{15GHz}/S_{5GHz}$, indicating that these regions are more affected by synchrotron. In Figure 6a, a great deal of structure in the $S_{15GHz}/S_{5GHz}$ ratio is evident in the vicinity of W51 Main: the ratio is higher towards the continuum peaks, indicating that the peaks have higher free-free optical depths, or lower relative contributions from synchrotron emission, than their envelopes.

We additionally compare the radio recombination lines observed simultaneously with the continuum and $H_2CO$. Hydrogen RRLs are often extremely well-correlated with the continuum and are therefore good indicators of the calibration quality.

In Figure 6, we show the ratios between the two frequencies in RRLs and continuum and the line-to-continuum ratios at both frequencies. The 'line' values are the integrated flux densities over the range 20 to 100 km s$^{-1}$, which includes all H$\alpha$ emission but no He$\alpha$.

The ratios between the $x$ and $y$ axis in each plot in Figure 5 are fitted using a total least squares approach with uniform errors for each data point. The line-to-continuum ratio is $L/C(H77\alpha) \sim 0.15$ and $L/C(H112\alpha) \sim 0.04$; in both cases there is little evidence for deviation from a linear relationship.

#### *3.2.4. Comparison of the RRL and continuum data*

Radio recombination lines are generally observed to be well-correlated with the corresponding radio continuum, particularly at low frequencies. At 5 and 15 GHz, the population level departure coefficients are close to 1, $b_n > 0.95$ (Wilson et al. 2009; Walmsley 1990).

While radio recombination lines are purely thermal in nature, the large-scale continuum may include a contribution from synchrotron emission. The morphological similarity between the 90 cm and 4 m (meter - i.e., 74 MHz) images presented by Brogan et al. (2013) and our 6 and 2 cm data hint that synchrotron emission could be significant. However, the high degree of correlation between the 2 and 6 cm described below suggest that synchrotron 'contamination' is minor at both wavelengths.

Figure 5 shows a comparison between the integrated RRL surface brightness and radio continuum at both 2 and 6 cm[5]. The figure shows the total least squares best-fit slopes to the data assuming uniform error, which yield a measurement of the line-to-continuum ratio.

We use the line-to-continuum ratio in both bands to measure the electron temperature using Equation 14.58 of Wilson et al. (2009), which assumes a plane-parallel, optically-thin emission region with lines formed in local thermodynamic equilibrium (the $*$ in $T_e^*$ is meant to indicate these three assumptions are made). The two lines yield consistent measurements, with mean $T_e^* \sim 7000 - 8000$ K; these measurements are consistent with smaller-scale measurements using the VLA with H92$\alpha$ (Mehringer 1994). There is only a little structure in the $T_e^*$ maps, with a hint of higher temperatures around G49.1-0.4, coincident with the W51C supernova remnant. Other structures are most likely due to the limited S/N.

Finally, we fit a single-component Gaussian to each pixel to produce velocity maps. These are discussed in Section 4.4.

### 3.3. Carbon and Helium RRLs

Helium RRLs were prevalent and reasonably well-correlated with the hydrogen RRLs, but we did not examine them in detail. He77$\alpha$ is detected at much higher signal-to-noise than than He107-112$\alpha$. There were no clear detections of C77$\alpha$ or C107-112$\alpha$, though there is a possible C77$\alpha$ signal at G49.366-0.304 with $v_{lsr} \approx 55$ km s$^{-1}$ and a possible detection toward W51 Main along the wing of the He77$\alpha$ line. The He77$\alpha$ line detections are associated with regions of high $H_n\alpha$ but not regions of different $T_e^*$.

---

[5]The H107, 108, 109, and 111$\alpha$ data were affected by missing (corrupted) data in one segment of the map. H107 and 108$\alpha$ were also affected by RFI. We therefore used the average of the H110 and H112$\alpha$ lines for the 6 cm line ananylsis.

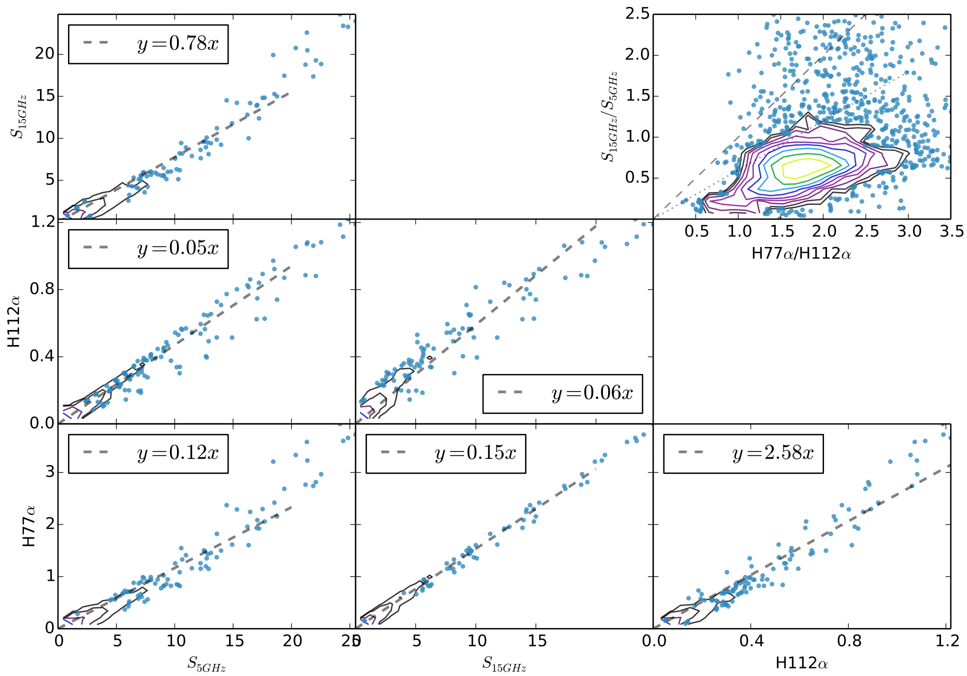



Fig. 5.— Plots of the 5 GHz and 15 GHz continuum and RRL flux densities against one another; all units are in Jy. The dashed lines show the total least squares best fit line with the slope shown in the legend. Wherever the density of points is too high to display, the points have been replaced with a contour plot showing the density of data points. The upper-right panel shows a comparison of the continuum ratio to the RRL ratio. The dashed line in the upper-right plot has slope 1, and the dotted line has slope 0.6.

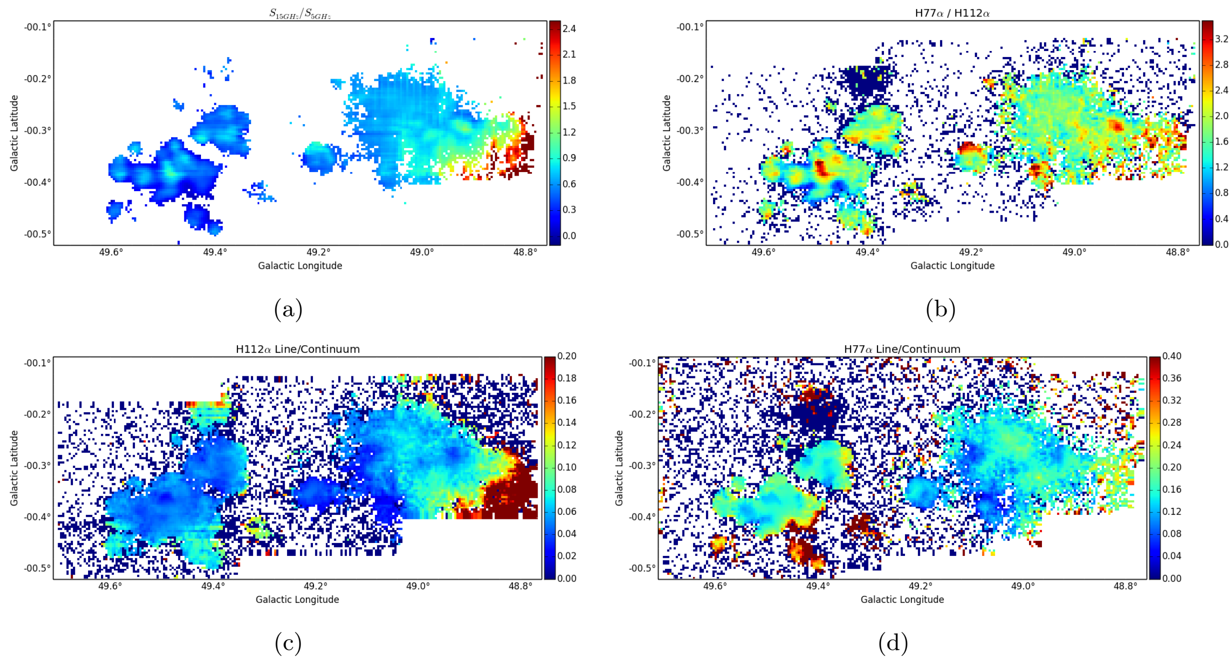



Fig. 6.— Ratio maps of the ionized gas in W51. (a) Continuum ratio $S_{15GHz}/S_{5GHz}$. For $\alpha = -0.1$, an optically thin free-free source, the ratio is 0.9, while for $\alpha = 2$, an optically thick source, the ratio is 9. (b) The ratio of the H77$\alpha$ peak to the H112$\alpha$ peak. (c) The line-to-continuum ratio H112$\alpha$ / $S_{5GHz}$ (d) The line-to-continuum ratio H77$\alpha$ / $S_{15GHz}$

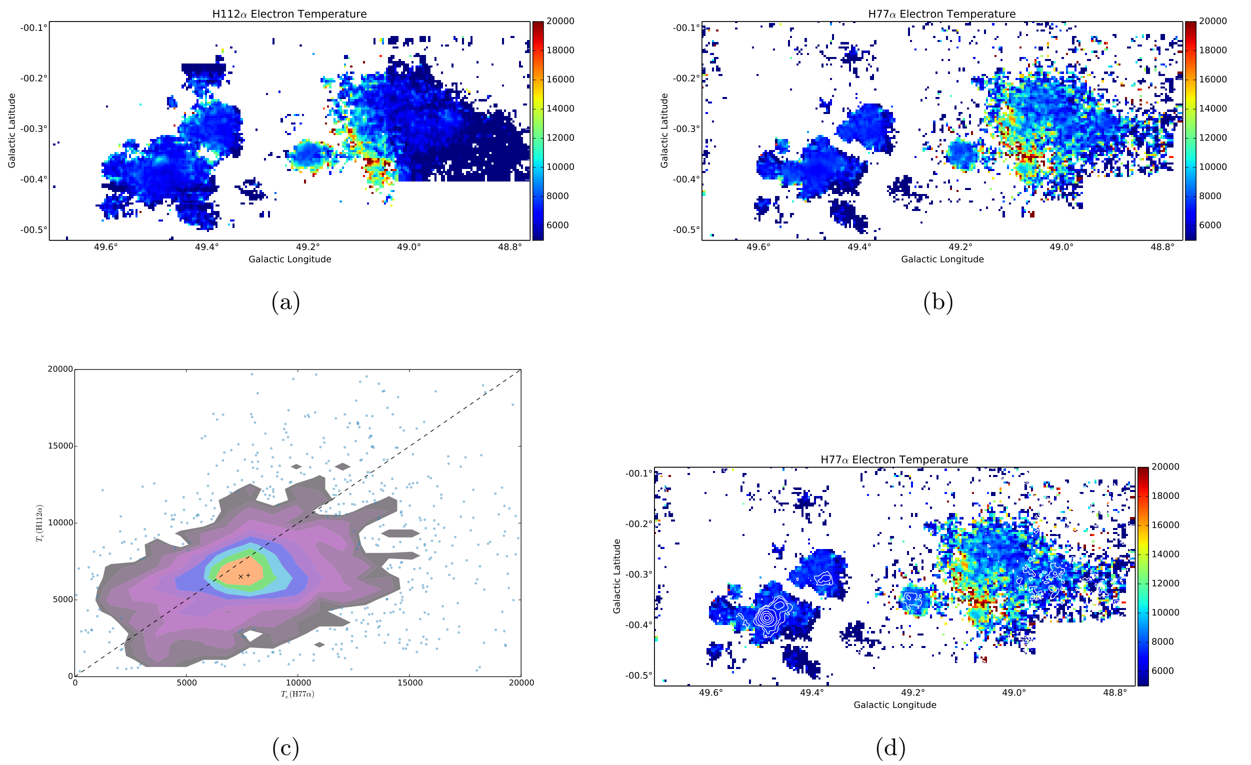



Fig. 7.— (a) The H112$\alpha$ electron temperature map showing $T_e^*$ in K. (b) The H77$\alpha$ electron temperature map showing $T_e^*$ in K. (c) The measured electron temperature in the 6 cm vs the 2 cm band at each spatial pixel with significant detected RRL emission. The contours show regions of increasing pixel density. The $x$ marks the median and the $+$ marks the mean over all valid pixels. (d) Same as (b), but with integrated He77$\alpha$ contours at levels [0.0125, 0.025, 0.05, 0.1, 0.15, 0.2] K km s$^{-1}$ overlaid. The contours on the right side ($\ell < 49$) most likely trace noise, since the noise in that region is higher.

### 3.4. Optical Depths

The data cubes were converted into "optical depth" data cubes by dividing the integrated $H_2CO$ absorption signature by the measured continuum level. We added a fixed background of 2.73 K to the reduced images to account for the CMB, which is absent from the images due to the background-subtraction process. We define an "observer's optical depth"

$$\tau_{obs} = -\ln\left[\frac{T_{mb}}{T_{bg}}\right] \tag{1}$$

as opposed to the 'true' optical depth, which is modeled in radiative transfer calculations

$$\tau = -\ln\left[\frac{T_{mb} - T_{ex}}{T_{bg} - T_{ex}}\right] \tag{2}$$

The approximation $\tau_{obs} = \tau$ is valid for $T_{ex} << T_{bg}$, which is true when an H II region is the backlight but generally not when the CMB is. Displaying the data on this scale makes regions of similar gas surface density appear the same, rather than being enhanced where there are backlights. The noise is correspondingly suppressed where backlighting sources are present.

$H_2CO$ absorption is ubiquitous across the map. In the Arecibo data, 8044 of 17800 spatial pixels have peak optical depths $> 5\sigma$, and 14547 have peaks $> 3\sigma$, so $H_2CO$ absorption is detected at $\sim 80\%$ of the observed positions.

The GBT $H_2CO$ $2_{11} - 2_{12}$ data have lower peak signal-to-noise because the continuum background is lower. Additionally, the $2_{11} - 2_{12}$ line is expected to trace denser gas and therefore be detected along fewer lines of sight. The $2_{11} - 2_{12}$ line is detected with a peak at $> 5\sigma$ in 3497 pixels (20%) and $> 3\sigma$ in 12254 pixels (69%). The high detection rate validates $H_2CO$ as an efficient dense-gas tracer.

There were no detections of $H_2CO$ $1_{10} - 1_{11}$ or $2_{11} - 2_{12}$ emission.

### 3.5. Ratio Cubes

In order to compare the $H_2CO$ spectra, we first created optical depth ratio cubes. The data were masked by selecting all voxels with signal-to-noise $> 2$ in *both* the $1_{10} - 1_{11}$ and $2_{11} - 2_{12}$ data cubes. To filter out spurious signals, of which many are expected given the 3.5 million voxels in the data cubes, we filtered the cube with a number-of-neighbors filter (i.e., a $3 \times 3$ kernel with all values of unity except the center pixel) and removed all pixels with fewer than 7 neighbors with detections. After filtering, 35039 voxels were left, or $\sim 1\%$ of the total (though this fraction is arbitrary, since the number of spectral pixels is determined by the gridding; i.e. we could have included a larger velocity range in the output cube). These are located in 5822 pixels, or 33% of the 17718 pixels valid in both data sets.

Histograms of the data are shown in Figure 8. They illustrate the non-Gaussian nature of the noise. The ratio distribution, in particular, appears to follow a power-law distribution with a cutoff at low ratio.

Note that there is hyperfine structure in the $1_{10} - 1_{11}$ lines, but the maximum offset is 1.15 km s$^{-1}$, and that hyperfine line is expected to be 1/6th of the total. For the purpose of these bulk comparisons on the large observed scales, it is safe to ignore the hyperfine structure.

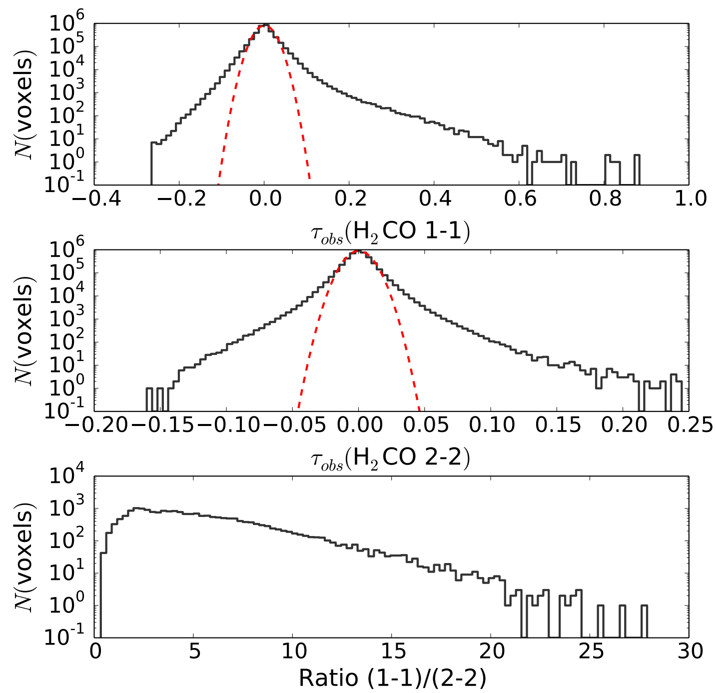



Fig. 8.— Histograms of the optical depth data cubes. In panels (a) and (b), a Gaussian distribution with width equal to the median absolute deviation is shown for reference; the noise is not well described by a single Gaussian distribution. In panel (c), the distribution of the ratio is shown.

### 3.6. Converting observed line ratios to volume densities

The $H_2CO$ line ratio can be transformed into a volume density of hydrogen $n(H_2)$ using large velocity gradient model grids. The column density of o-$H_2CO$, the ortho-to-para-ratio of $H_2$, and the gas temperature are the three main 'nuisance parameters' that can be marginalized over.

The o-$H_2CO$ column density is degenerate with the velocity gradient in LVG models. The $H_2$ column density is degenerate with this gradient *and* the abundance of o-$H_2CO$. Precise measurements of the $H_2CO$ abundance are not generally available, but typical values of $X_{\text{o-H}_2\text{CO}} = 10^{-10} - 10^{-8}$ relative to $H_2$ are generally assumed (Mangum & Wootten 1993; Ginsburg et al. 2011, 2013a; Ao et al. 2013) and found to be consistent with the observations. Nonetheless, little is known about local variations in o-$H_2CO$ abundance, except that it freezes out in cold, dense cores (Young et al. 2004).

The model grids were generated using RADEX LVG models (python wrapper `https://github.com/keflavich/pyradex/`; original code van der Tak et al. 2007). They assume a velocity gradient of 1 km $s^{-1}$ $pc^{-1}$. Ginsburg et al. (2011) and Ginsburg et al. (2013a) discussed the effect of a local gas density distribution on the molecular excitation, but due to the complexity involved in accounting for these effects, we ignore them here.

### 3.7. A note on nondetections

$H_2\,^{13}CO$ was not detected anywhere in the W51 complex in either the $1_{10} - 1_{11}$ or $2_{11} - 2_{12}$ lines. The peak signal-to-noise in the $1_{10} - 1_{11}$ cube was 180, so we report a $3 - \sigma$ upper limit on the $H_2CO/H_2\,^{13}CO$ ratio $R > 60$, which is consistent with solar values of the $^{12}C/^{13}C$ ratio.

## 4. Structure of the Molecular Cloud

The $H_2CO$ observations reveal two essential features of the W51 GMC: its density structure and its line-of-sight geometry.

### 4.1. Molecular Gas and $H_2CO$ modeling

Figure 9 shows the most important observed properties of the $H_2CO$ lines. The figures show the peak observed optical depth $\tau_{obs} = -\log(T_{MB}/T_{continuum})$ in each line along with the ratio of the $1_{10} - 1_{11}$ to the $2_{11} - 2_{12}$ optical depth. They are masked to show significant pixels determined by:

1. Selecting all voxels with $S/N > 2$ in both images or $S/N > 4$ in either and with at least 7 (of 26 possible) neighbors also having $S/N > 2$
2. Selecting all voxels with $>= 10$ neighbors having $S/N > 2$
3. Growing (dilating) the included mask region by 1 pixel in all directions
4. Selecting all voxels with $>= 5$ included neighbors
5. When used to mask 2D images, the selection is then collapsed such that any pixel containing at least one voxel along the line of sight is included

This approach effectively includes all significant pixels and all reliably detected regions within the data cube, though the number of neighbors used at each step and the selected growth size are somewhat arbitrary and could be modified with little effect.

Figure 9 contains two ratio maps. The first shows the observed optical depth ratio, while the second shows the 'true' optical depth ratio assuming an excitation temperature for each line, $T_{ex}(1_{10}-1_{11}) = 1.0$ K and $T_{ex}(2_{11}-2_{12}) = 1.5$ K. These excitation temperatures are representative of those expected for most of the modeled parameter space in which absorption is expected. Fitting of individual lines-of-sight confirm that good fits can be achieved using these assumed temperatures.

However, there are some clear outliers within the map: the clouds at G48.9-0.3 and G49.4-0.2 both show very low $1_{10}-1_{11}/2_{11}-2_{12}$ ratios over a broad area. As discussed in Sections 4.5.1 and 4.6, these two regions have H II regions in the foreground of the molecular gas. The ratios displayed in Figure 9 are therefore computed with an incorrect background assumed.

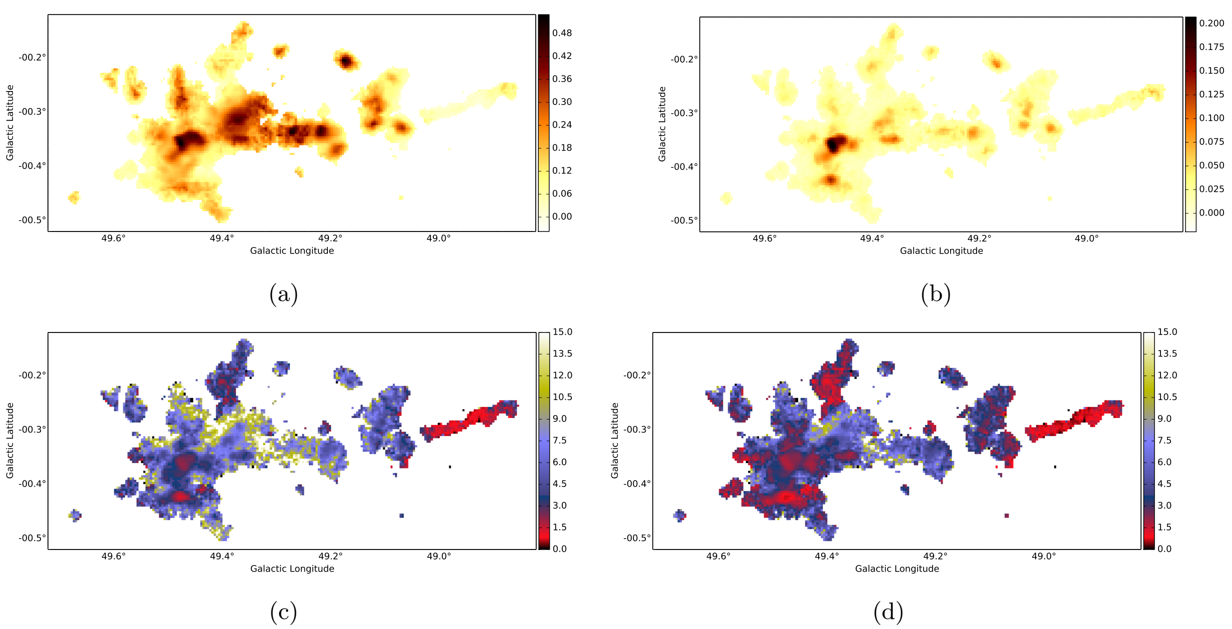



Fig. 9.— Plots of the peak *observed* optical depth $\tau_{obs} = -\log(T_{MB}/T_{continuum})$ in the (a) $1_{10}-1_{11}$ and (b) $2_{11}-2_{12}$ lines and (c) their ratio, $1_{10}-1_{11}/\ 2_{11}-2_{12}$. Figure (d) shows the 'true' optical depth ratio assuming $T_{ex}(1_{10}-1_{11}) = 1.0$ K and $T_{ex}(2_{11}-2_{12}) = 1.5$ K; these are reasonable and representative excitation temperatures but they are not fits to the data. The data are masked with a filter described in Section 4.1 and cover the range $75 > V_{LSR} > 40$ km s$^{-1}$; see Figures 10 and 11 for individual velocity components. In general, lower (redder) ratios in figures (c) and (d) indicate higher densities, however in the filament at 49.0-0.3, the low ratio is due to the geometry in which $T_{continuum}$ is in the *foreground* of the molecular gas.

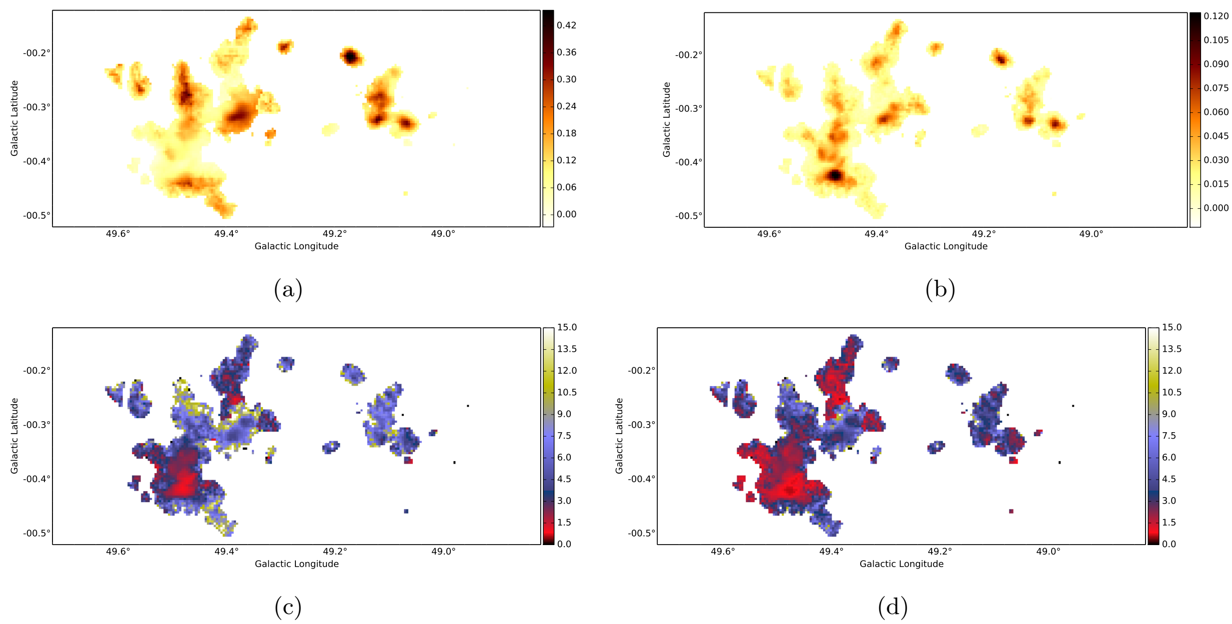



Fig. 10.— Same as Figure 9, but limited to $62 > V_{LSR} > 40$ km s$^{-1}$.

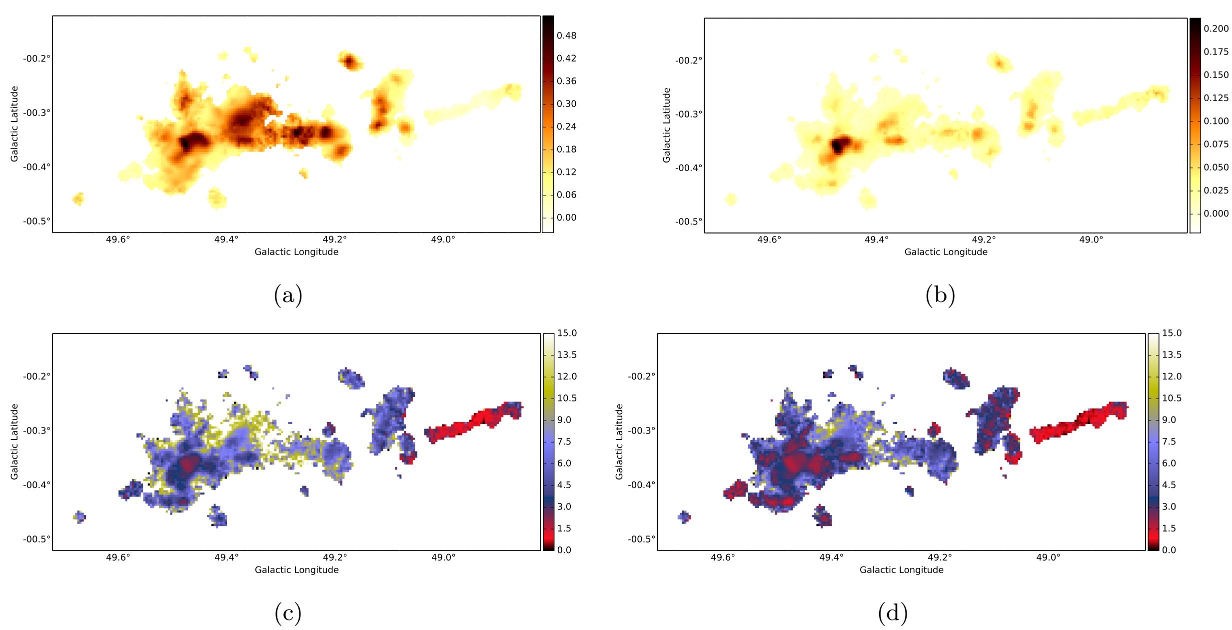



Fig. 11.— Same as Figure 9, but limited to $75 > V_{LSR} > 62$ km s$^{-1}$.

### 4.2. Density Maps

We computed the density using the $\chi^2$ minimization technique from Ginsburg et al. (2011). We measure $\chi^2$ over the full 4D parameter space (density, column density, temperature, and ortho-to-para ratio)

$$\chi^2 = \left(\frac{T_B(1_{10} - 1_{11}) - T_{model}(1_{10} - 1_{11})}{\sigma(T_B, 1_{10} - 1_{11})}\right)^2 + \left(\frac{T_B(2_{11} - 2_{12}) - T_{model}(2_{11} - 2_{12})}{\sigma(T_B, 2_{11} - 2_{12})}\right)^2 \quad (3)$$

We have therefore not enforced any constraints on the column density, temperature, or ortho-to-para ratio when fitting. The best-fit value of each of these parameters is taken to be the mean of those parameters over the range $\chi^2 - \chi^2_{min} < 1$.

The temperatures returned from the fitting process are, as expected, purely noise: the $H_2CO$ $1_{10} - 1_{11}/2_{11} - 2_{12}$ ratio provides virtually no constraint on the gas temperature and therefore leaving it as a free parameter has no effect on the fitted density. Similarly, the ortho-to-para ratio of $H_2$ is unconstrained in our data. In principle, the $H_2$ OPR has some effect on $H_2CO$ excitation, but in the regime we have modeled and obseved, no effect is apparent.

The o-$H_2CO$-column-weighted volume-density along each line of sight is shown in Figures 12 and 13. The former shows the weighted density over all voxels and the latter shows the weighted density over the two velocity ranges previously discussed. These projections include no information about the errors in the individual fits, which are available from Equation 3, but by weighting by column density, we have effectively selected the highest signal-to-noise points; the statistical errors are therefore negligible relative to the systematic in these maps.

The overall picture is of a central proto-cluster region with most of the gas mass at a density $n \sim 10^{5.5}$ cm$^{-3}$ within a diameter of $\sim 3$ pc, surrounded by a rich cloud in which most of the mass is at a density $\sim 10^4$ cm$^{-3}$ out to a diameter $d \sim 14$ pc.

### 4.3. Model fitting and geometry

Both $H_2CO$ lines are seen only in absorption. However, in some cases the absorption is against a continuum background, while in others the absorption may be only against the CMB.

We have fit the $H_2CO$ lines constrained by the LVG models (Section 3.6) to spectra averaged over regions with coherent molecular absorption signatures. We compared the $\chi^2$ values for fits with the observed continuum as the background to those with the background fixed to $T_{BG} = T_{CMB}$. We then selected the better of the two fits as representative of the physical conditions.

It is possible that there are multiple continuum emitters along the line of sight in many cases, with the absorbing molecular gas somewhere in the middle. While this possibility adds uncertainty to the measurements, there are some cases in which the dominant continuum can unambiguously be assigned a foreground or background position.

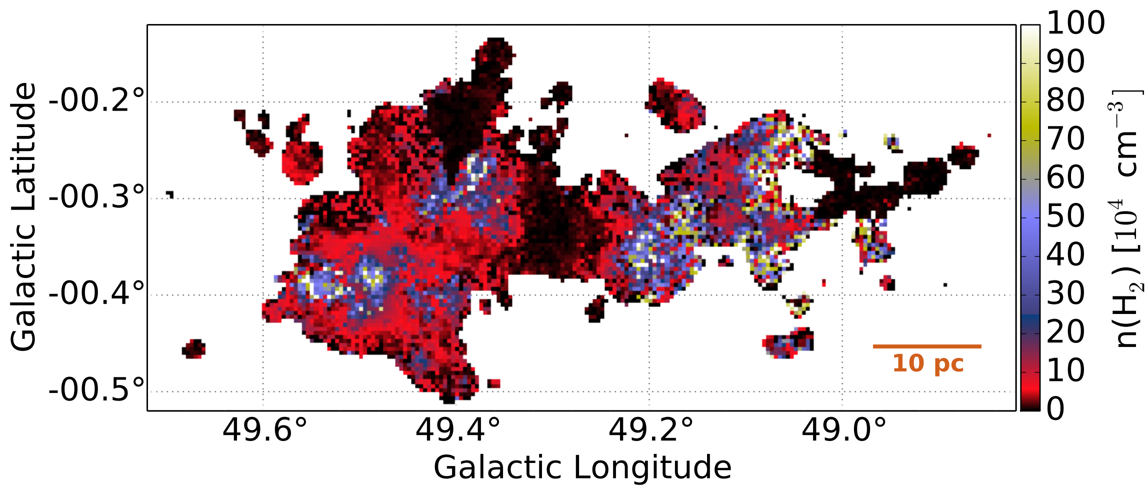



Fig. 12.— Map of the column-weighted volume density along the line of sight averaged over all velocities.

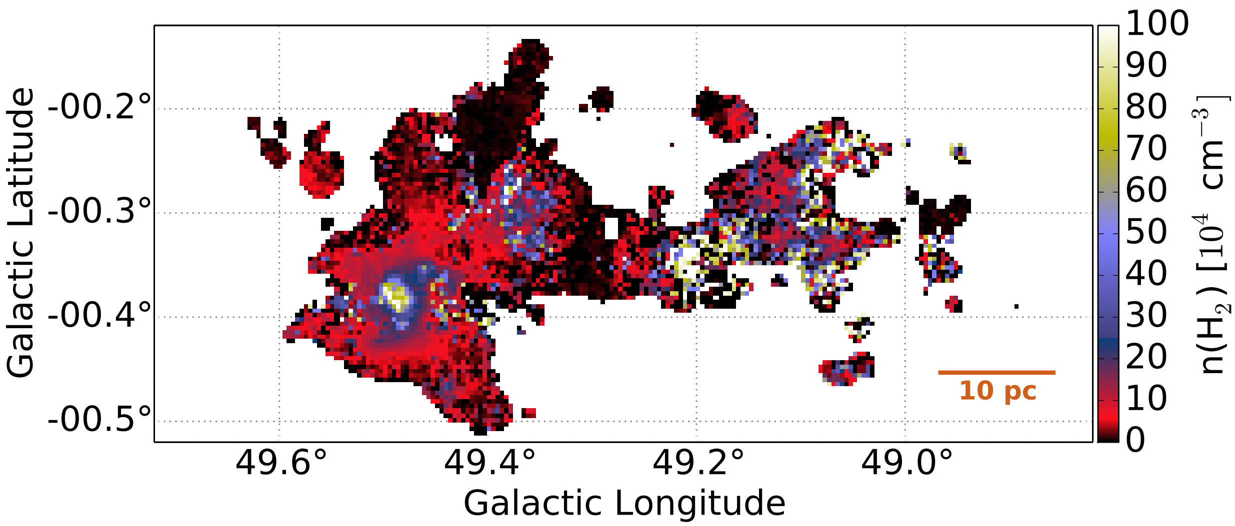
### 4.4. Kinematic Maps

Maps showing the overall kinematics of the region are shown in Figures 16 and 17. Figure 16 shows the velocity at peak absorption of the $H_2CO$ $1_{10}-1_{11}$ line and the fitted radio recombination line centroid velocity. Figure 17 shows the best simultaneous fit to the $H_2CO$ $1_{10}-1_{11}$ and $2_{11}-2_{12}$ absorption features over two different velocity ranges. The $1_{10}-1_{11}$ absorption velocity in Figure 16a approximately shows the velocity of the front-most molecular clouds along the line of sight at each position.

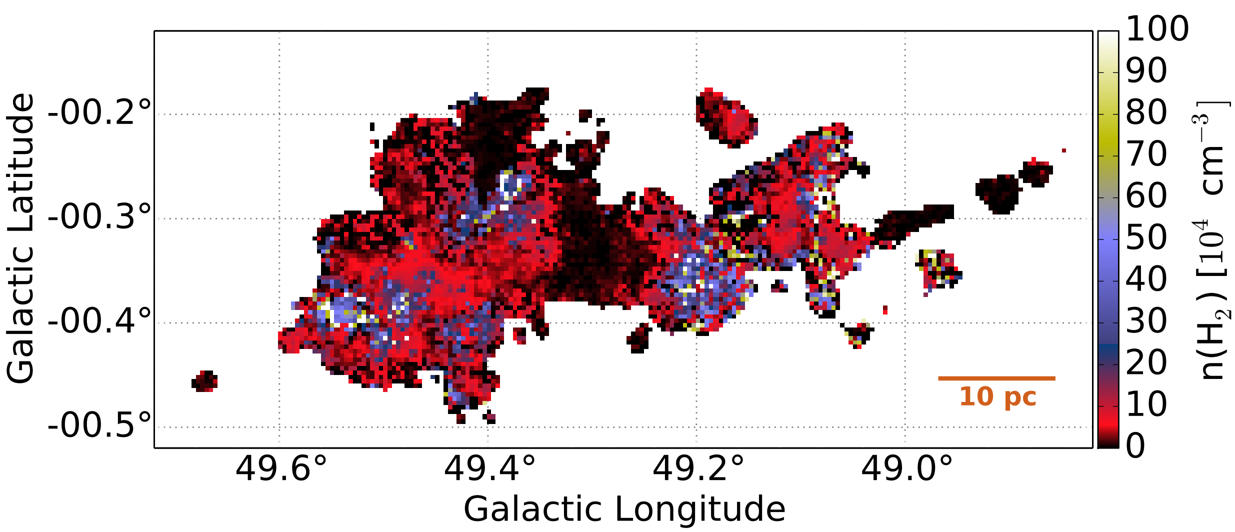
### 4.5. Geometry of Individual Regions

#### *4.5.1. The W51 B Filament*

The W51 B filament, labeled as the 66 km s$^{-1}$ cloud in Figure 14, exhibits bright CO emission ($T_A^* \sim 30-50$ K in the Parsons et al. (2012) CO 3-2 data) but has relatively weak $H_2CO$ absorption. The absorption models are inconsistent with the molecular gas being in front of the continuum emission, so Figure 14 shows the continuum sources in front of the cloud at lower $\ell$. Figure 18 shows an example model fit with the continuum assumed to be in front and in back, illustrating that the best-fit model parameters with continuum in the back do not reproduce the data. The relative positioning of the molecular gas behind the H II regions suggests that the molecular gas is also behind the W51 C supernova remnant.




(a)




(b)

Fig. 13.— Map of the column-weighted volume density along the line of sight, split into (a) the 40-62 km s$^{-1}$ component and (b) the 62 to 75 km s$^{-1}$ component.

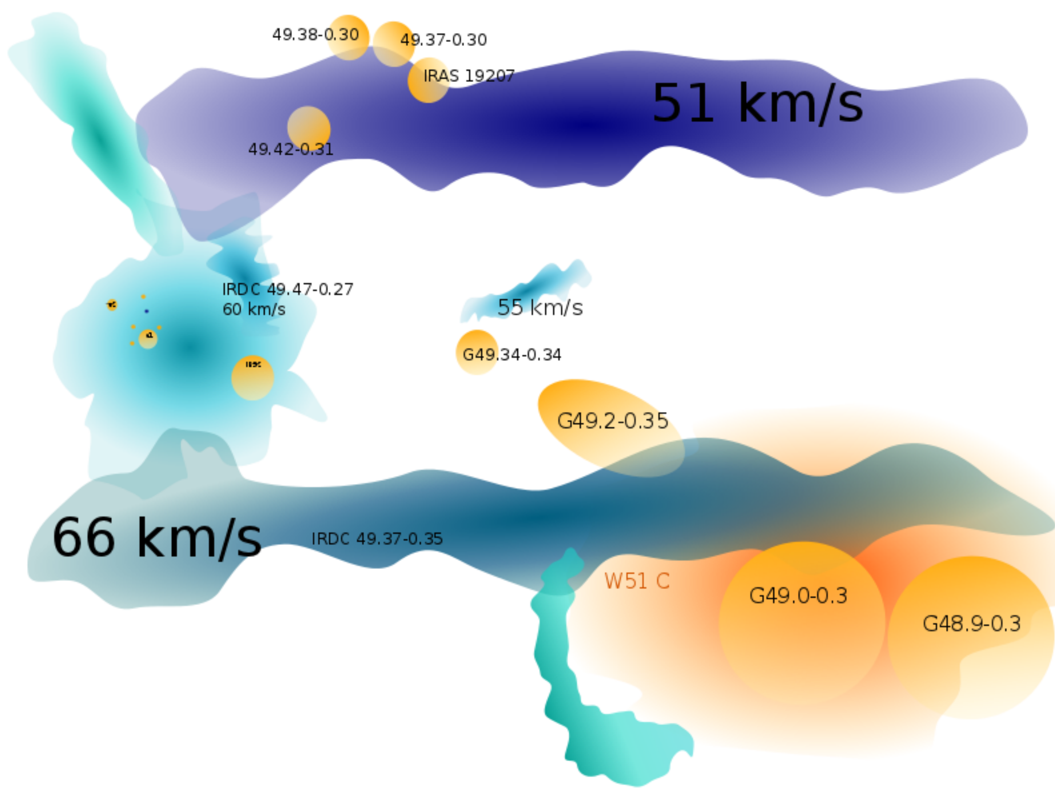



Fig. 14.— A sketched diagram of the W51 region as viewed from the Galactic north pole, with the observer looking up the page from the bottom (i.e., W51C is the front-most labeled object along our line-of-sight). There are a few significant differences between this and Figure 29 of Kang et al. (2010), particularly the relative geometry of the cloud and the H II regions in W51 B. We also show a good deal more detail, revealing that there are H II regions on both front and back of many clouds. This figure needs to be kept up to date with text changes. Ideally, I'd like to build an interactive version of this figure, but that may not happen in finite time.

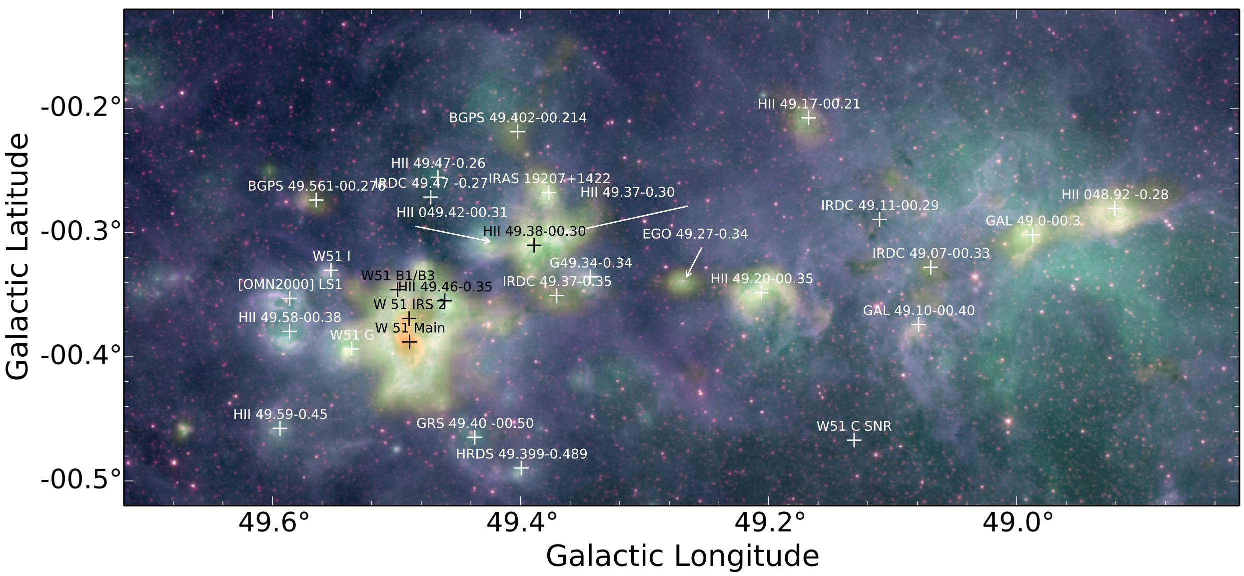


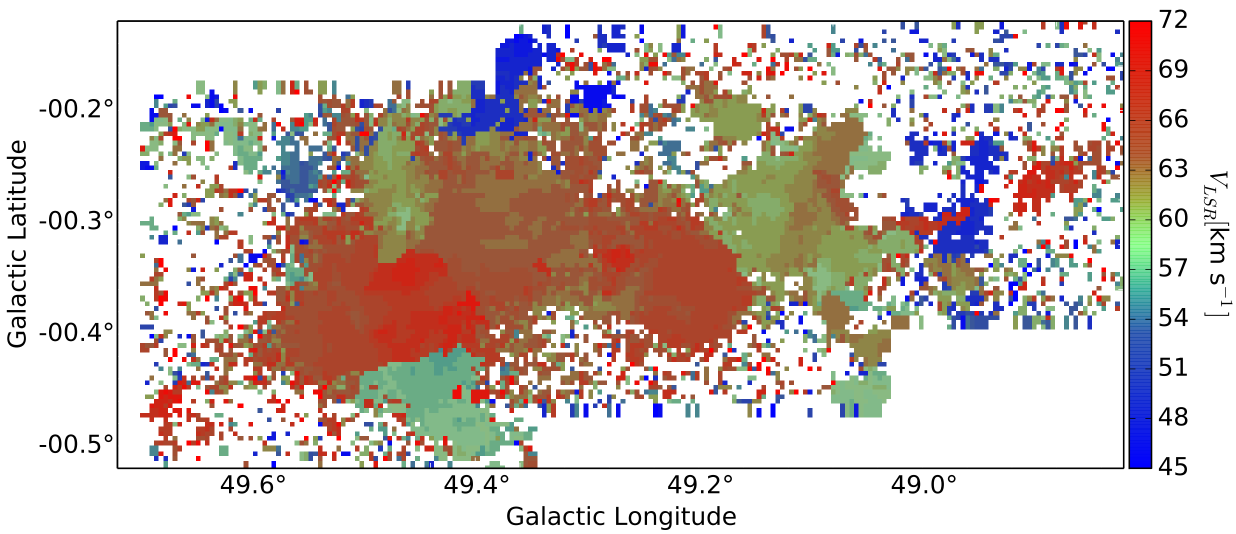

Fig. 15.— Labeled figures of the W51A and W51B regions, highlighting H II regions and infrared dark clouds. The colors are described in Figure 1. These labels can be compared to Figure 14 to associated labeled regions in the plane of the sky with their counterparts in the face-on view of our Galaxy.

### *4.5.2. The edge of W51 C*

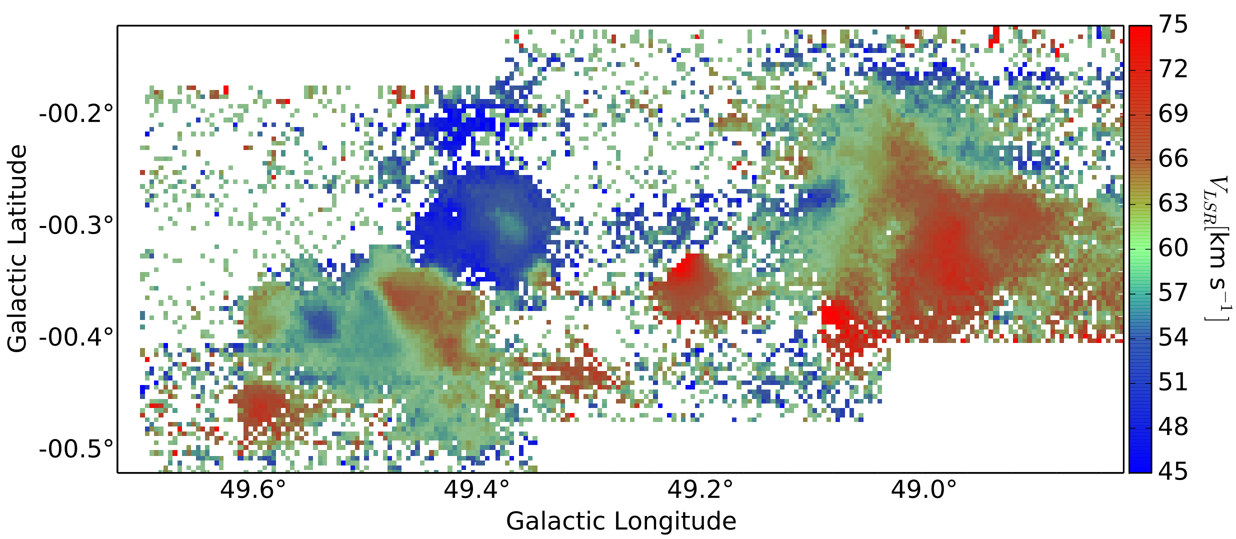
W51 C is a supernova remnant that spatially overlaps with the W51 B star forming region. Brogan et al. (2013) argue that the supernova remnant must be in front of the H II region G49.20-0.35 because the H II-region has not absorbed all of the 4m (74 MHz) nonthermal emission. The G49.1-0.4, G49.0-0.3, and G48.9-0.3 regions, however, show 4m absorption signatures and may be in the foreground. There are clumps aligned along the 68 km $s^{-1}$ filamentary cloud with very high CO and H I velocities (Koo & Moon 1997b,a; Brogan et al. 2013), indicating that the SNR is interacting with the molecular gas.

The clumps at G49.1-0.3, $\sim 68$ km $s^{-1}$ are either lower density ($n < 1.5 \times 10^4$ cm$^{-3}$) and in the background of the H II region or high density ($n > 1.5 \times 10^5$ cm$^{-3}$), low-column density and in the foreground. The 62 km $s^{-1}$ clumps have densities a few times higher, $n \sim 4 \times 10^4$ cm$^{-3}$, and are clearly in the foreground of the continuum emission because their absorption depths are $\sim 2.5$ K, which cannot occur for absorption against the CMB. Figure 19 shows a model spectrum fitted assuming the continuum lies between the two molecular velocity components. The relative strength of the $^{13}$CO and the $H_2CO$ also suggests that the 68 km $s^{-1}$ component is behind the continuum.

We are seeing molecular gas both in front of and behind the supernova. This geometry can be readily confirmed by looking for molecular absorption at much lower frequencies where the SN synchrotron emission dominates over the H II region free-free emission, i.e. the 335 and 71 MHz

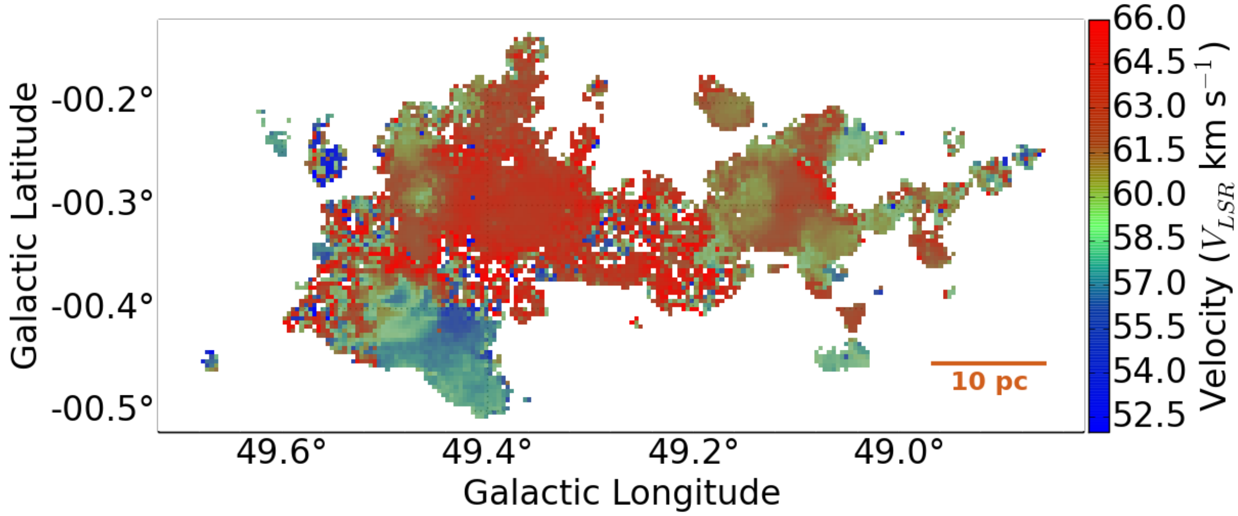



(a)

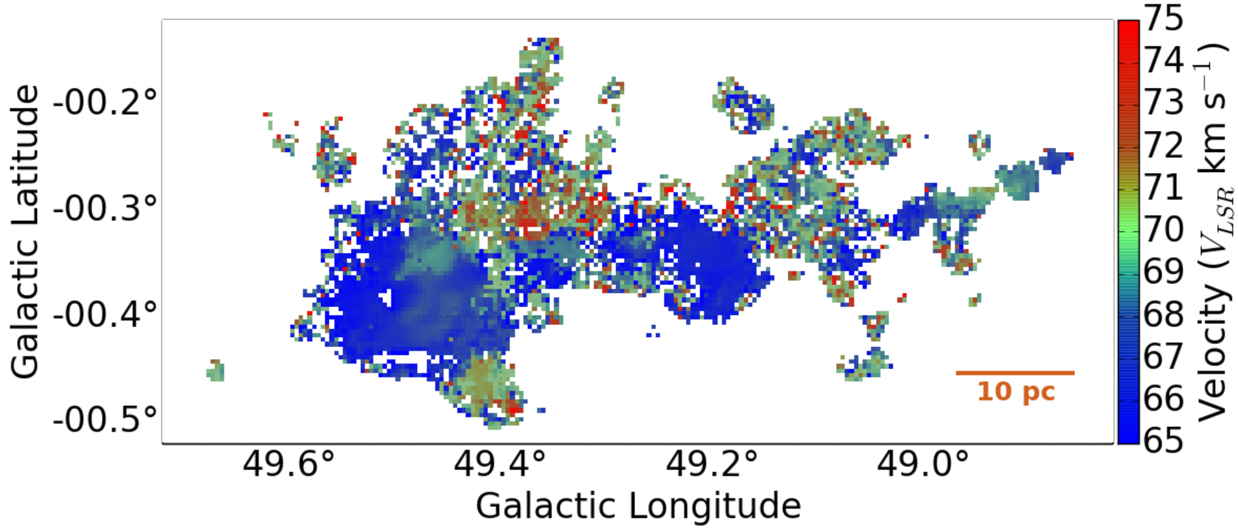



(b)

Fig. 16.— (left) Velocity of the peak $H_2CO$ $1_{10} - 1_{11}$ signal (deepest absorption) at 1 km s$^{-1}$ resolution (right) Velocity of the peak H110$\alpha$ emission as derived from Gaussian fits to each spectrum.




(a)




(b)

Fig. 17.— Maps of the fitted $H_2CO$ velocity components over the range $40 < v_{LSR} < 66$ km s$^{-1}$ (left) and $66 < v_{LSR} < 75$ km s$^{-1}$ (right). The regions that appear noisy have ambiguous multi-component decompositions. This is being refit

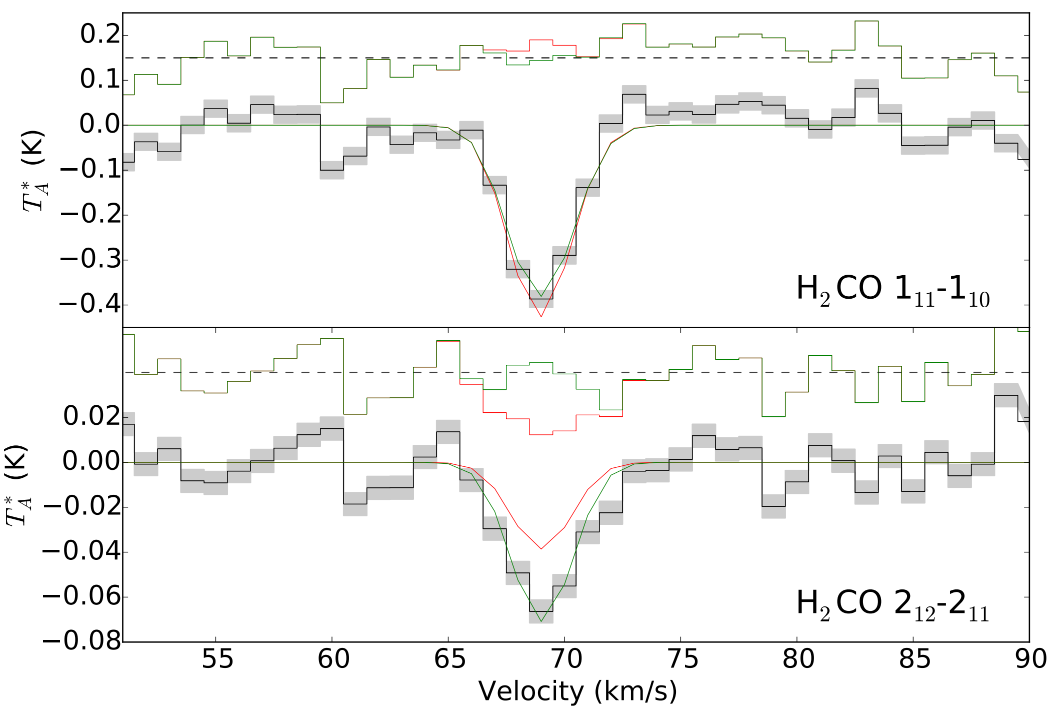



Fig. 18.— An example of the difference in models between a continuum source (red) and the CMB (green) as the background. The top plot shows the $1_{10} - 1_{11}$ line and the bottom shows the $2_{11} - 2_{12}$ line both with the continuum level set to zero in the plot but treated as a frozen parameter in the fit. The residuals are shown offset above the spectra, with the dashed line indicating the zero-residual level. The grey shaded regions show the 1-$\sigma$ error bars on each pixel. The model with the CMB as the only background is able to reproduce the absorption line, while the model with the H II region in the background cannot account for the depth of the $2_{11} - 2_{12}$ line. The reduced $\chi^2/n$ for the models are 14.1 (red) and 2.8 (green), evaluated only over the pixels where the model is greater than the local RMS.

p-$H_2CO$ lines.

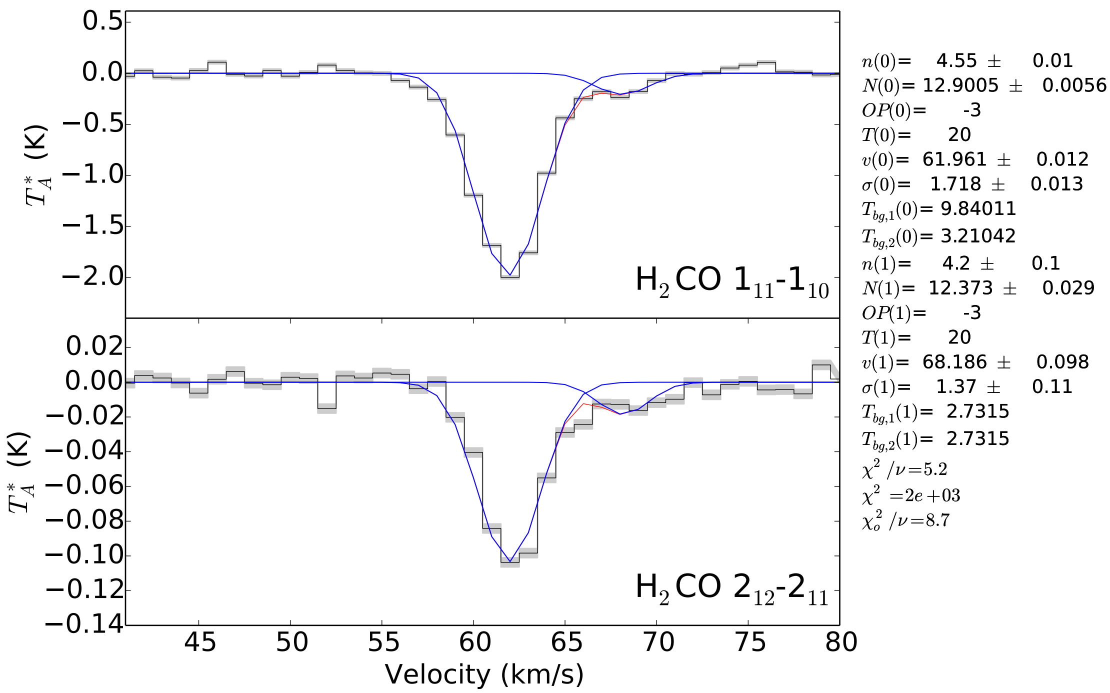



Fig. 19.— The spectrum extracted from G49.119-0.277 in a 55″ radius aperture, showing a model in which the continuum is *behind* the 63 km s$^{-1}$ component but in front of the 68 km s$^{-1}$ component. The legend gives the fit parameters along with $1-\sigma$ error bars. The parameters with no errors indicated (OPR, $T$, $T_{BG}$) are assumed or independently measured values.

### *4.5.3. G49.20-0.35 and G49.1-0.4*

Tian & Leahy (2013) focus on the H II regions G49.20-0.35 and G49.10-0.40 (called G49.10-0.38 in their work) to determine the relative geometry of the W51 C SNR and the W51 B H II/star-forming region. They observe that the high-velocity H I is not detected toward either of these sources, indicating that the H II regions must be behind the high-velocity H I features.

We detect $H_2CO$ $1_{10}-1_{11}$ at $\sim 58$ and $\sim 63$ km s$^{-1}$ toward G49.10-0.40, with line ratios that are consistent with the H II region being behind the molecular cloud complex. It also has an extreme RRL velocity, $v_{110\alpha} \approx 72$km s$^{-1}$, the most redshifted seen in the entire W51 region (see Figures 16 and 20).

G49.20-0.35 is also clearly behind the molecular cloud, as evidenced both by $H_2CO$ absorption depth and the IRDC absorption in the foreground. It has an RRL velocity $v_{110\alpha} \approx 70$ km s$^{-1}$.

Because both H II regions are extremely redshifted, they are most likely associated with the W51 B cloud complex, contrary to the interpretation by Tian & Leahy (2013) in which they are unrelated background clouds. The Galactic rotation curve doesn't allow for velocities red of $\sim 60$ km s$^{-1}$, and almost none of the molecular gas exceeds $\sim 70$ km s$^{-1}$ even on the wings. The H II regions are therefore probably shooting out the back side of the molecular cloud, perhaps accelerating ionized gas from the $\sim 66$ km s$^{-1}$ component further to the red.

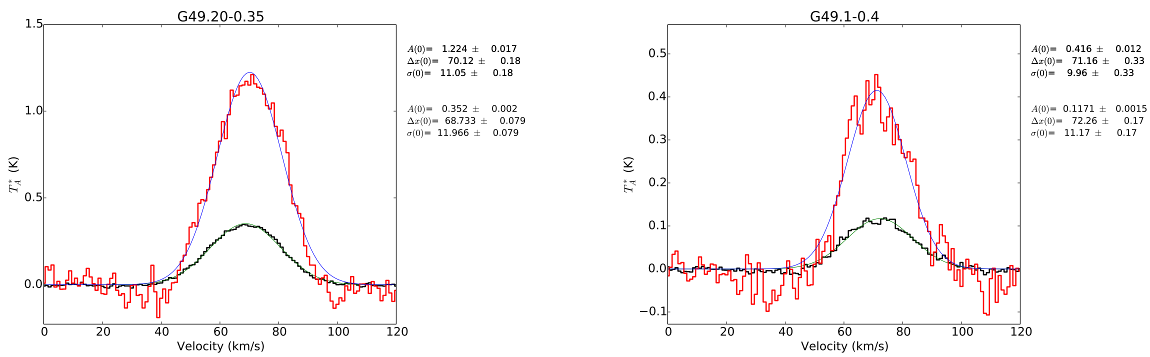



Fig. 20.— Fitted H110$\alpha$ (red) and H77$\alpha$ (black) spectra extracted from $55''$ apertures centered on G49.20-0.35 (left) and G49.1-0.4 (right). The best-fit Gaussian parameters are shown in the legends, with the lower legend corresponding to H77$\alpha$.

### *4.5.4. The 66 km s$^{-1}$ IRDC*

Between W51 A and W51 B, there is a component of the 68 km s$^{-1}$ cloud that is filamentary and in the foreground of all of the free-free emission. This cloud component is evident as an IRDC in the Spitzer GLIMPSE images from $\ell = 49.393$, $b = -0.357$ to $\ell = 49.207$, $b = -0.338$.

The H II region G49.20-0.35 is clearly behind the IRDC, though there are strong morphological hints that it is interacting with and truncated by the cloud.

### *4.5.5. G49.27-0.34*

The UCH II region G49.27-0.34, which was considered a candidate extended green object (EGO) and subsequently rejected for lack of $H_2$ emission (De Buizer & Vacca 2010; Lee et al. 2013), exhibits a second velocity component at $\sim 68$km s$^{-1}$, slightly but clearly redshifted of the rest of the IRDC. It contains a gas mass $\sim 2 \times 10^3 M_\odot$ based on the BGPS flux and using the assumptions outlined in Aguirre et al. (2011), suggesting that the high velocity could be due to infall or virialized gas within a deep potential. The virialized velocity width, given the radius and mass from the BGPS data, is $\sigma_{vir} = 8.8$km s$^{-1}$, while the measured $H_2CO$ linewidth is $FWHM(H_2CO) = 7.2$km s$^{-1}$,

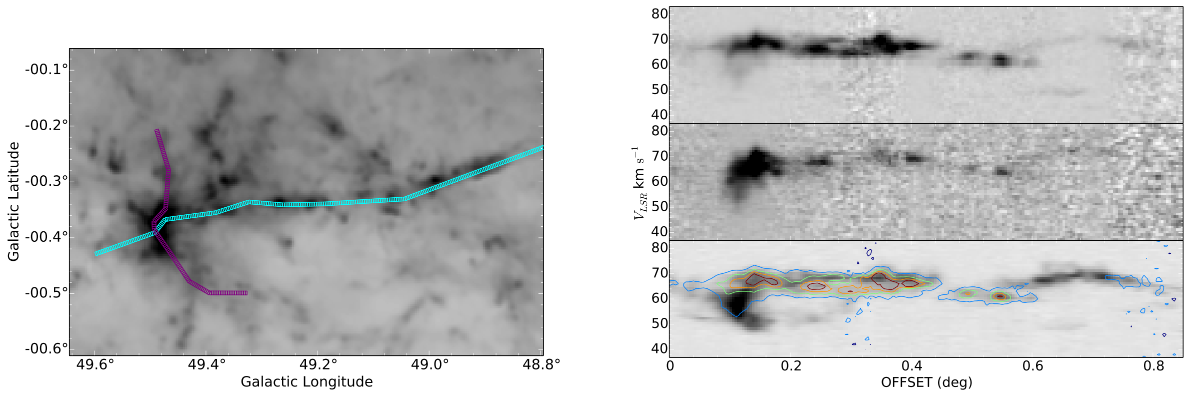



Fig. 21.— (left) A column density map fitted from the Herschel Hi-Gal data with two filament extraction regions superposed in cyan and red-blue (right) A position-velocity slice of the 68 km s$^{-1}$ cloud, shown in cyan in the left figure, which includes an infrared dark cloud and the interaction region with the W51C supernova remnant. (top) $H_2CO$ $1_{10} - 1_{11}$ observed optical depth (middle) $H_2CO$ $2_{11} - 2_{12}$ observed optical depth (bottom) $^{13}CO$ 1-0 emission from the GRS with $H_2CO$ $1_{10} - 1_{11}$ contours superposed. The weakness of the $H_2CO$ absorption on the right half of the cloud corroborates the geometry inferred from comparison of the $1_{10} - 1_{11}$ and $2_{11} - 2_{12}$ lines in Figure 18.

wider than in any other part of the cloud except W51 Main.

Both radio continuum and RRLs are detected toward this source. The H77$\alpha$ RRL velocity is $\sim 58$ km s$^{-1}$, significantly blueshifted from the molecular gas. The $H_2CO$ lines do not independently distinguish between the continuum source being in the front or back of the cloud, but the mean density from the BGPS mass and radius $n \sim 2.5 \times 10^4$ cm$^{-3}$ is within a factor of 2 of the $H_2CO$-derived density, $n \sim 1.4 \times 10^4$ cm$^{-3}$, if the continuum source is behind the gas, while the $H_2CO$-derived density is too low, $n \sim 2 \times 10^3$ cm$^{-3}$ if the continuum source is in front.

The implied geometry therefore has the H II region behind the molecular gas, plowing toward it at a velocity difference $\Delta v \sim 10$ km s$^{-1}$. Such a high velocity difference may indicate that the H II region is confined by the molecular gas and on a plunging orbit into the cloud.

#### *4.5.6. G49.4-0.3f, aka G49.34-0.34, aka IRAS 19209+1418*

The H II region centered at 49.34-0.34 was identified by Mehringer (1994) as part of the G49.4-0.3 complex. There are 3 distinct $H_2CO$ line components at 51, 63.70, and 68.47 km s$^{-1}$. The 51 km s$^{-1}$ component is behind the H II region; the $^{13}CO$ line is detected at comparable brightness at 51 km s$^{-1}$ and 63 km s$^{-1}$, while the $H_2CO$ $1_{10} - 1_{11}$ line is $\sim 10\times$ deeper at 63 km s$^{-1}$. The RRLs associated with this source are at $v_{LSR} = 58 \pm 1$ km s$^{-1}$.

The $H_2CO$ lines are moderately well-fit by the two-velocity-component model, but there is a relative excess of $2_{11} - 2_{12}$ absorption at 66 km s$^{-1}$. The extra absorption may indicate that there is a high-density, low-column component at this velocity.

The 8 $\mu$m GLIMPSE image shows that the 68 km s$^{-1}$ IRDC crosses in front of this source. Herschel Hi-Gal 70$\mu$m images reveal a ring structure that is hinted at in the 8 $\mu$m image. There is no evidence for interaction between the ring feature and the IRDC. This intriguing feature will likely be difficult to study in detail because the dusty, molecular gas feature lies in front of it.

### 4.6. G49.4-0.3

The collection of H II regions around G49.4-0.3 vaguely resembles a cartoon mouse. As noted in Carpenter & Sanders (1998), the molecular gas in this region is separated into two distinct components, one at 51 km s$^{-1}$ and the other at 64 km s$^{-1}$. The 64 km s$^{-1}$ component is in the foreground, while the 51 km s$^{-1}$ is in the background of most of the H II regions.

Both cloud components are in the foreground of the central H II regions at G49.36-0.31, the 'eyes' of the mouse. The density of the 51 km s$^{-1}$ component is an order of magnitude higher than that in the 64 km s$^{-1}$ component in this region, suggesting that the gas is being compressed by the H II region. The clean separation between the 64 and 51 km s$^{-1}$ cloud components suggests that they are not interacting at this location.

Based on the absorption line depths, the G49.38-0.30, IRAS 19207+1422, and G49.37-0.30 H II regions are behind the 51 km s$^{-1}$ cloud. The 8 $\mu$m absorption features are associated with the 64 km s$^{-1}$ cloud and are in front of all of the H II regions.

The 8 $\mu$m morphology of G49.42-0.31 is bubble-like, so it is plausible that the H II region is neither in front nor behind the 51 km s$^{-1}$ cloud but embedded within it, blowing a hole in the cloud.

### 4.7. Infrared Dark Cloud G49.47-0.27

The cloud to the north of W51 Main/IRS2 appears as a dark feature in Spitzer GLIMPSE 8 $\mu$m maps. It is detected in $H_2CO$ from 54 to 64 km s$^{-1}$. Throughout, it has a high $1_{10} - 1_{11}/2_{11} - 2_{12}$ ratio, $\gtrsim 7$ in most voxels, indicating a low density $n \lesssim 10^3$ cm$^{-3}$.

Centered at 60.6 km s$^{-1}$, the region has a line FWHM 5 - 7 km s$^{-1}$, indicating that it is quite turbulent, with 3D Mach number in the range $10 < \mathcal{M} < 20$ for an assumed $10 < T < 20$ K. At its centroid velocity, it is connected to the W51 Main cloud.

There is a previously unreported bubble HII region in the north part of this cloud, which we designate G49.47-0.26, with radius $\sim 70''$ (1.7 pc). The H II has RRL velocities $v_{lsr} \approx 50$km s$^{-1}$.

Because it is not detected in Brackett $\gamma$ emission (from the UWISH2 survey: Froebrich et al. 2011), it is most likely behind the cloud.

The Kang et al. (2009) Spitzer survey of YSOs in the region indicates that there are no YSOs within the boundaries of this cloud; it is very likely non-star-forming at present.

Because the cloud is continuous with the W51 Main region in velocity and is infrared-dark, it is most likely at the same distance as W51 Main and associated indirectly with the massive cluster forming region.

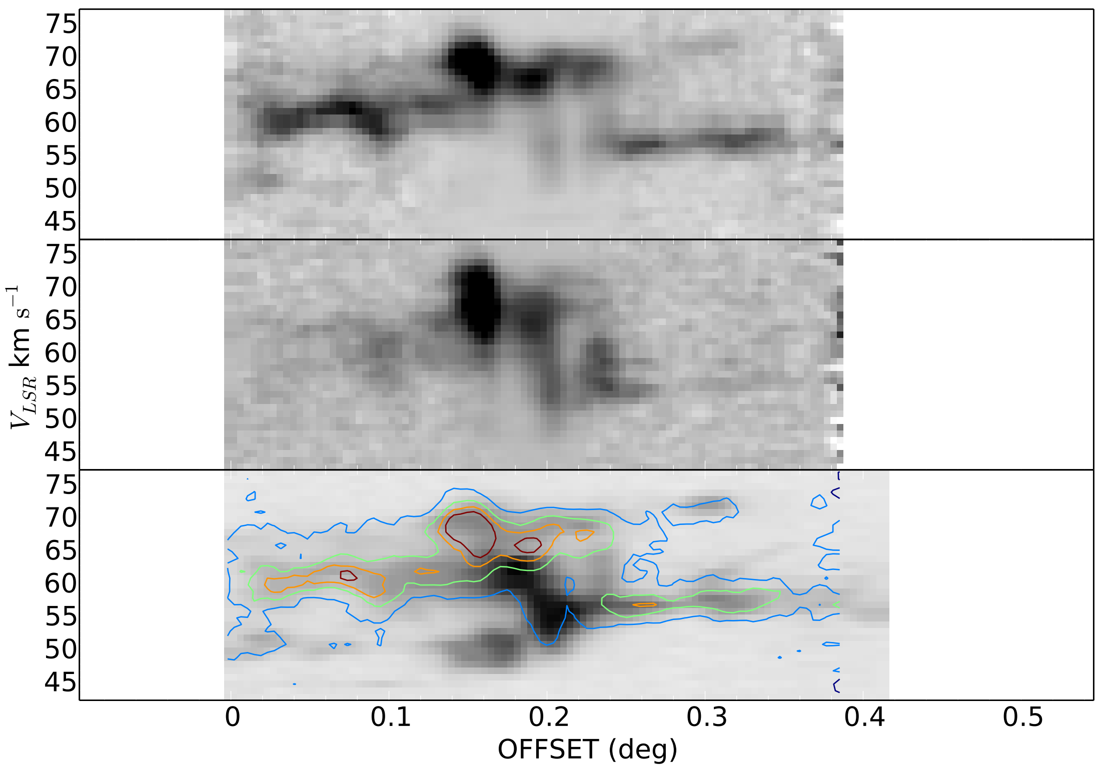



Fig. 22.— Position-velocity diagrams of filamentary structures to the north and south of W51 main. See Figure 21 for the extracted region.

### 4.8. W51 Main & W51 IRS 2

The W51 Main and IRS 2 spectra show that both have ionized gas components at $v_{LSR} \sim 55$ km s$^{-1}$. This velocity approximately coincides with the peak of the $^{13}$CO emission.

The $H_2CO$ $1_{10} - 1_{11}$ spectra are deepest at $\sim 68$ km s$^{-1}$, while the $2_{11} - 2_{12}$ have depths approximately equal between the $\sim 58$ km s$^{-1}$ and $\sim 68$km s$^{-1}$ components. The 55-60 km s$^{-1}$ components are too deep to be entirely behind the H II regions. This indicates that the 55 km s$^{-1}$

ionized gas must be embedded within the molecular cloud, with molecular gas on *both* sides of the ionized gas along the line of sight.

Because these are well-studied regions, the low spatial resolution $H_2CO$ spectra we present here add little new information about the gas kinematics. However, all of the velocity components observed in the W51 region are apparently kinematically connected to the W51 clusters.

### 4.9. The 40 km s$^{-1}$ clouds

There are clouds observed at 40 km s$^{-1}$ that show only weak $H_2CO$ absorption spread across nearly the entire region. These molecular clouds are behind nearly all of the H II regions in the W51 complex. There are additional 40 km s$^{-1}$ clouds clearly seen in H I absorption (Stil et al. 2006) that are not associated with these molecular clouds, but instead represent a foreground population of neutral atomic medium clouds.

## 5. A formation scenario for the W51 region

The W51 cloud complex has been discussed as both a collection of unrelated clouds and a tight complex of interacting clouds (Carpenter & Sanders 1998; Kang et al. 2010). The $H_2CO$ data presented here support the idea that the 68 km s$^{-1}$ cloud is a coherent entity and that it is interacting with other clouds associated with W51 A. The presence of W51 Main at the interaction point between multiple clouds hints that its great mass and star-forming potential was triggered by this cloud-cloud collision. It is likely that both observed clouds were pre-existing features in a larger, possibly atomic, cloud that underwent stretching and squeezing upon interaction with a spiral arm.

The line-of-sight length of the W51 complex is still uncertain, despite our constraints on the relative geometry of different regions. The best prospect for resolving the line-of-sight structure of the region is via precise constraints on distances to the individual regions. Spectrophotometric surveys of the individual stellar sub-clusters may be able to provide this and should be undertaken. Maser parallax observations of different zones may also provide differential distance estimates.

## 6. Discussion

The W51 cloud complex includes a full range of star forming conditions. In the west, W51 B, there is an older generation of stars including at least one supernova remnant. In the east, there is a pair of forming, still-embedded massive clusters. We have described the geometry of these regions and features of the cloud structures, now we will speculate on the broader implications of these observations.

The gas in the W51 B region, while clearly affected by the expanding W51 C supernova, is less dense than most of the gas in the W51 A region. The supernova feedback is, if anything, destructive; a 'collect-and-collapse' scenario does not fit the observed gas structure. Stellar feedback in the W51 region is therefore destructive, though so far not severely so.

The proximity of the 68 km s$^{-1}$ filamentary 'high velocity stream' and the W51 Main protocluster and their relative line-of-sight positions have been presented as evidence for a cloud-cloud collision (Kang et al. 2010). Examination of the $H_2CO$ line ratios has shown that the protocluster is embedded in the $\sim 55$ km s$^{-1}$ molecular cloud.

### 6.1. Gas density and its relation to star formation

The lowest density readily observed in our survey of W51 is $n \sim 10^4$ cm$^{-3}$. Lower densities can be detected with $H_2CO$ observations but require greater depth, especially in the $2_{11} - 2_{12}$ line.

In Ginsburg et al. (2013b), we measured the $H_2CO$ line ratio to high precision in a low-density, turbulent, non-star-forming GMC. While the mean density of this cloud was of order $n \sim 10^2$ cm$^{-3}$, the mass-weighted density as probed by $H_2CO$ was $n \approx 5 \times 10^3$ cm$^{-3}$. The density measurements in the W51 cloud exceed this value for the majority of the detections significant in both $1_{10} - 1_{11}$ and $2_{11} - 2_{12}$, indicating that there is a much higher fraction of high-density gas in this star-forming cloud than in the quiescent GMC in front of W49.

#### *6.1.1. Dense Gas Mass Fractions*

Because the $H_2CO$ densitometer yields a mass-weighted measurement of the gas volume density, it is difficult to connect directly to the total gas mass, which is the quantity of interest when determining bulk properties like star forming efficiency. However, because the $H_2CO$ and CO trace the same gas, the $H_2CO$-derived density can be applied to the total mass measured by CO. We assume that each $^{13}$CO PPV 'voxel' has a mass proportional to its integrated intensity and a density given by the $n(H_2)$ delivered from the $H_2CO$ densitometer.

The 'dense gas mass fraction' (DGMF) is an oft-quoted measurement used to argue about the speed of the star formation process, the existence of density threshold, and turbulent properties of the ISM (e.g. Fig. 5 of Krumholz et al. 2007; Battisti & Heyer 2014; Kainulainen et al. 2013; Juneau et al. 2009; Muraoka et al. 2009; Hopkins et al. 2013). However, these fractions are most often quoted as mass of gas at a *single* density divided by the total mass. We present an improvement on these measurements, showing the continuous distribution of the dense gas mass fraction.

Figure 23 shows the result of using our $H_2CO$ PPV density cubes to measure the DGMF from $^{13}$CO. We use a range of density thresholds from $\sim 10^3$ to $\sim 10^4$ cm$^{-3}$. At each density, we identify all pixels in the $^{13}$CO PPV cube above this threshold and sum those. We then divide by the total integrated $^{13}$CO brightness to get the mass fraction.

Figure 24 shows the same results, but for two individual regions: the W51 Main proto-cluster and the W51 B region. Within about 10 pc of W51 Main, around half of the mass is at density $n > 10^4$ cm$^{-3}$. By contrast, the rest of the molecular cloud shows a consistent fraction $f \sim 10\%$ with $n > 10^4$ cm$^{-3}$.

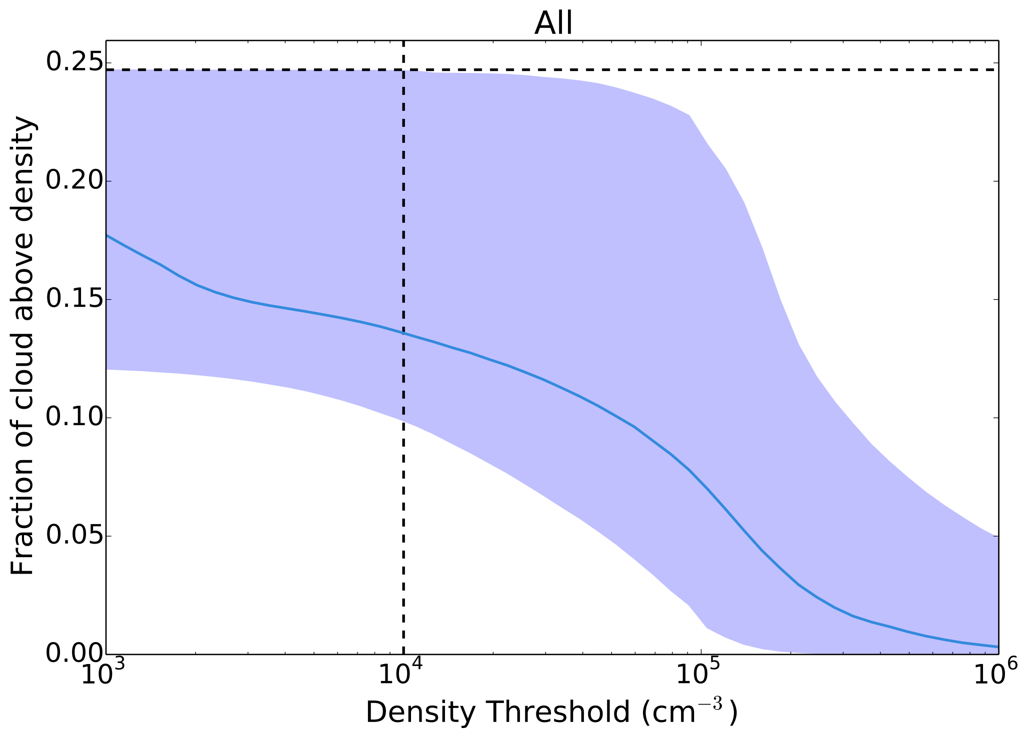



Fig. 23.— The 'dense gas mass fraction' as a function of volume density $n(\mathrm{H_2})$ cm$^{-3}$. The $y$-axis shows the sum of $^{13}$CO pixels from the GRS cube with $\mathrm{H_2CO}$-derived density above the value shown on the $x$-axis divided by the total. Both values are computed over the velocity range 40km s$^{-1}$ $< v_{LSR} <$ 75km s$^{-1}$. The blue shaded region shows the extent of plausible model fits at each density: effectively, this is the $\sim 1-\sigma$ error region. The vertical line at $n = 10^4$ cm$^{-3}$ indicates the approximate completeness limit. The horizontal line shows the fraction of $^{13}$CO flux detected at $> 2\sigma$ in both the $1_{10} - 1_{11}$ and $2_{11} - 2_{12}$ lines: it represents the upper limit of what could have been detected if, e.g., all $\mathrm{H_2CO}$ detections were toward regions with $n > 10^4$ cm$^{-3}$. The failure to converge to a fraction $f \rightarrow f_{max}$ indicates that there are some real detections of low-density gas.

The multi-density DGMF presented here can be compared to models of 'global' collapse. They are effectively a gas density cumulative distribution function. However, to understand the systematic effects of line-of-sight stacking of different velocity components (and corresponding radiative transfer issues), similar analysis should be performed on hydrodynamic simulations.

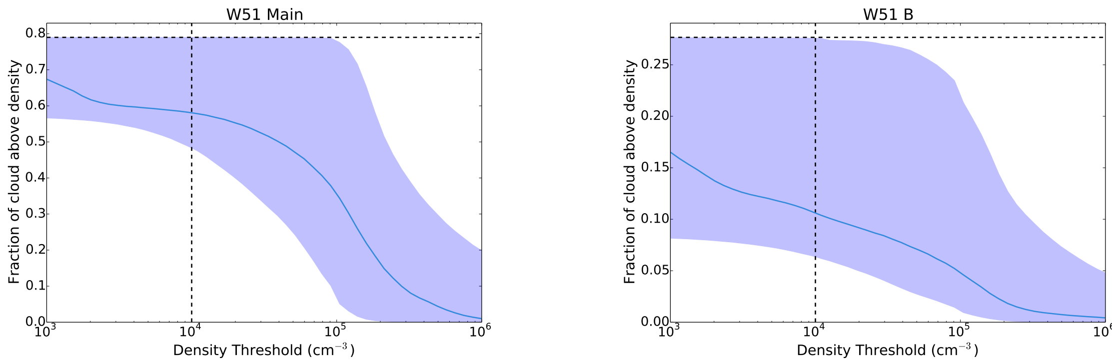



Fig. 24.— Same as Figure 23, but for two individual regions: W51 Main (left), the region $49.4° < \ell < 49.6°$, $-0.5° < b < -0.3°$, and W51 B (right) with $\ell < 49.4°$. Note that the $y$ axes have different ranges.

#### *6.1.2. Dense Gas Fraction assumptions & caveats*

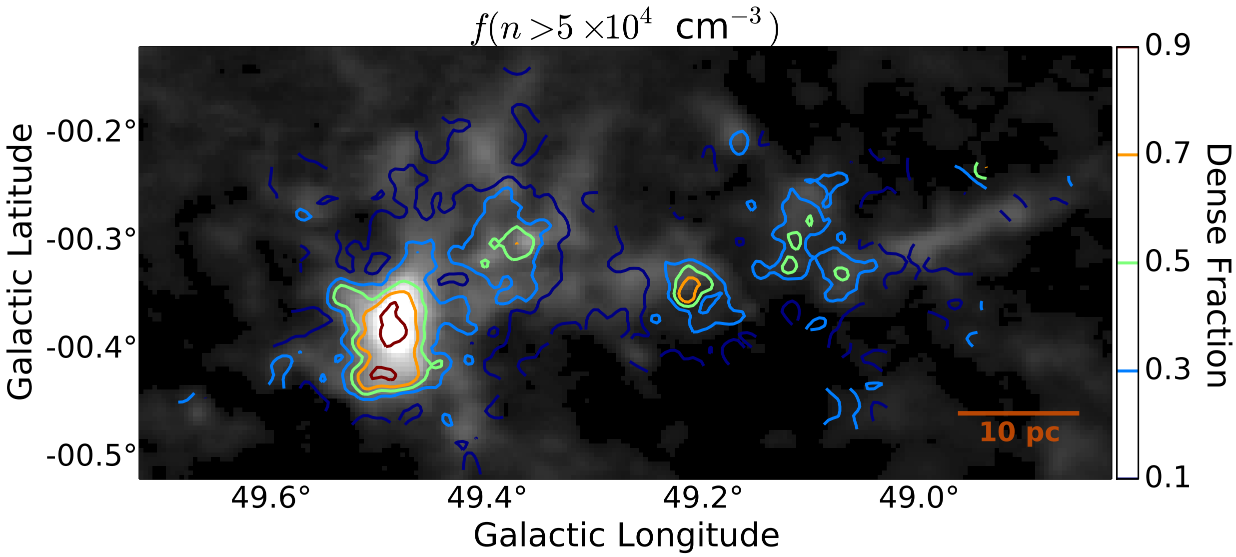
This analysis relies on the $^{13}$CO being optically thing and thermally excited, both of which are generally good assumptions for the majority of the mass. The molecular cloud probably includes no more than $\sim 20\%$ of its mass in CO-dark gas (Pineda & Teixeira 2013; Langer et al. 2013; Smith et al. 2014), which adds little to the overall uncertainty.

In Section 4.2, we discussed the various caveats and issues related to $H_2CO$ density fitting. To account for the full range of errors in that analysis, we have plotted the DGMF calculated using the minimum and maximum values of the $H_2CO$-derived density consistent with the data at the $1 - \sigma$ level in Figure 23 and 24.

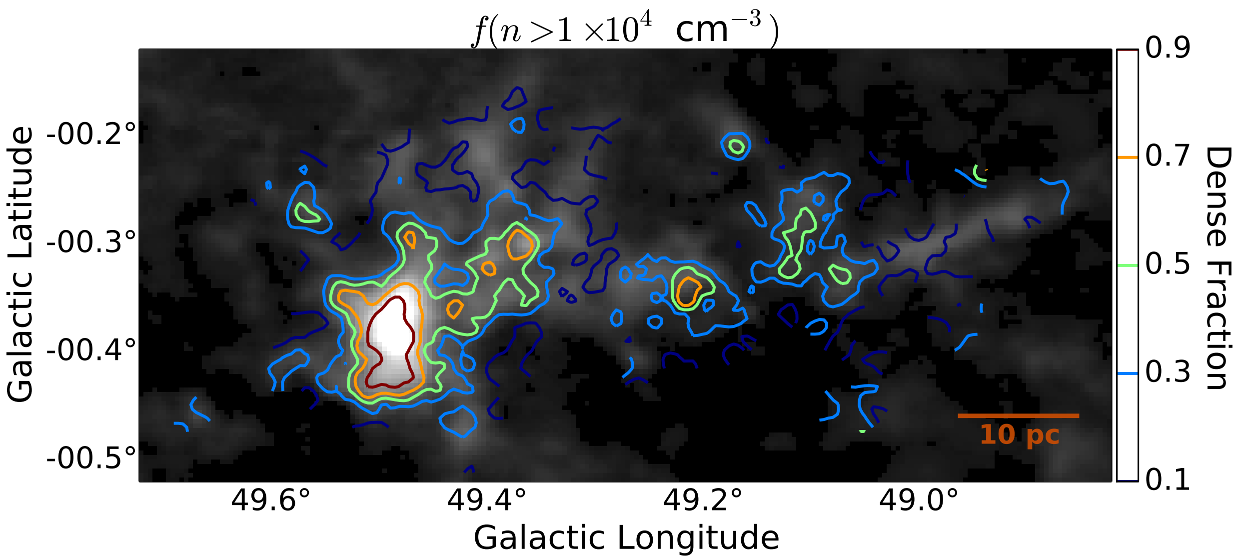
We have assigned *all* of the mass associated with a given PPV voxel with a single, fixed density in this analysis. There is certainly some mass at a lower density in each PPV pixel. This additional mass biases the measured DGMF higher than it should be, but probably only by a small amount at each threshold. This systematic bias could be characterized from simulations projected into PPV space.

### 6.2. Implications for the future evolution of W51

The low DGMF associated with the W51 B / 66 km s$^{-1}$ cloud indicates that it has a low star formation potential despite containing significant mass ($M \gtrsim 1 \times 10^5 \; M_\odot$). The presence of a supernova remnant and old, diffuse HII regions indicates that the cloud did previously (and recently) form stars, but is now being destroyed. The fact that this cloud contains $10 - 30\%$ of the total mass of W51, but has only a tiny fraction of its total mass above the purported star forming thresholds supports this story.




(a)




(b)

Fig. 25.— Contours of the dense gas mass fraction using two different thresholds overlaid on the integrated $^{13}$CO map. The regions with fraction $f > 0.5$ should be rapidly forming stars. The background image in both frames is the GRS $^{13}$CO image integrated over the range 40km s$^{-1}$ < $v_{lsr}$ < 75km s$^{-1}$, masked to include only pixels with $S > 0.5$ K. Unclosed contours should probably be ignored or removed.

This cloud is therefore a good region to examine the effects of different feedback mechanisms in parallel. It may also be a good location to examine how star formation comes to an end in the presence of massive star feedback.

By contrast, the W51 Main region has a dense gas fraction $\sim 1$, or at least $\sim 50\%$ out to nearly 5 pc. It has presumably only formed a small fraction of its total potential. It exceeds all of the various star formation and massive star formation thresholds (e.g. Lada et al. 2010; Krumholz & McKee 2008; Kauffmann & Pillai 2010), and therefore is expect to form $\sim 10^4 - 10^6$ $M_\odot$ of additional stars.

### 6.3. Implications for extragalactic observations of $H_2CO$

Although W51 is one of the most massive and active GMCs in the galaxy, containing 7% of the present-day massive star formation galaxy-wide (Urquhart et al. 2014), its star-forming gas mass is predominantly at a moderate density, $n \sim 5 \times 10^4$ cm$^{-3}$; there is very little gas above $10^6$ cm$^{-3}$ even in W51 Main. There were *no detections* of $H_2CO$ emission on the $\sim 1.25$ pc ($50''$) scales observed.

By contrast, in extragalactic observations of starburst galaxies, there have been detections of emission on $\sim 100$ pc scales. Mangum et al. (2013) report detections of $H_2CO$ $1_{10} - 1_{11}$ emission in NGC 3079, IC 860, IR 15107+0724, and Arp 220 on $\sim 10$ kpc scales. The implied local column densities from their analysis are modest, but the densities are extreme: their observations imply that the local-scale *chemical* conditions are comparable to W51, but the densities are different.

The only location in which densities $n \gtrsim 5 \times 10^5$ cm$^{-3}$ (comparable to $n_{\mathrm{H}_2}$(Arp 220), etc.) are observed in W51 are in the central W51 Main region. We don't observe emission because of the extremely bright continuum background source. It is therefore not possible to explain a $H_2CO$-emission galaxy by constructing it from collections of UCHII regions; such a galaxy would be continuum-bright and show only $H_2CO$ absorption. Instead, they must be assembled from huge quantities of high-density, non-star-forming gas. This result is in contradiction to the idea that Giant H II Regions are the 'building blocks' of starburst galaxies (e.g. Miura et al. 2014).

## 7. Conclusion

We have presented maps of the $H_2CO$ $1_{10}-1_{11}$ and $2_{11}-2_{12}$ and H77$\alpha$ and H110$\alpha$ lines covering the W51 star forming complex. The continuum data were compared with previous lower-resolution observations of the region, resulting yielding a consistent calibration.

The recombination lines have been used in conjunction with the continuum to estimate electron temperatures, with a consistent mean $T_e^* \sim 7500$ K.

The $H_2CO$ $1_{10} - 1_{11}/2_{11} - 2_{12}$ line ratio was used to measure gas volume densities. While systematic uncertainties in the modeling remain, the $H_2CO$ nonetheless yields a consistent picture of the W51 star forming region in which the central W51 Main proto-cluster is filled with gas of a mean density at least an order of magnitude greater than the rest of the cloud.

The $H_2CO$ lines and their ratios have also been used to constrain the geometry of the W51 GMC and the associated H II regions. The Galactic face-on view of W51 is presented in more detail than has previously been possible.

The data are made available in FITS cubes and images hosted at the CfA dataverse `doi: 10.7910/DVN/26818`, `http://thedata.harvard.edu/dvn/dv/W51_H2CO`. The entire reduction and analysis process and all associated code and scripts are made available via a git repository hosted on github: `https://github.com/keflavich/w51_singledish_h2co_maps`, with a snapshot of the publication version available as a tarball from figshare `TBD`.

**Acknowledgements**: We thank Xiaohui Sun for providing the Urumqi 6 cm Stokes I image prior to its availability on the survey website.

**Code Packages Used**:

- The GBT KFPA Pipeline `https://safe.nrao.edu/wiki/bin/view/Kbandfpa/ObserverGuide`
- aoIDL `http://www.naic.edu/~phil/download/aoIdl.tar.gz`
- astropy `www.astropy.org`
- astroquery `astroquery.readthedocs.org`
- sdpy `https://github.com/keflavich/sdpy`
- FITS_tools `https://github.com/keflavich/FITS_tools`
- aplpy `http://aplpy.github.io`
- image-registration `http://image-registration.rtfd.org`

## REFERENCES

Aguirre, J. E. et al. 2011, ApJS, 192, 4

Ao, Y. et al. 2013, A&A, 550, A135

Arnal, E. M. & Goss, W. M. 1985, A&A, 145, 369

Bastian, N. 2008, MNRAS, 390, 759

Battisti, A. J. & Heyer, M. H. 2014, ApJ, 780, 173

Beccari, G. et al. 2010, ApJ, 720, 1108

Bieging, J. H., Peters, W. L., & Kang, M. 2010, ApJS, 191, 232

Bressert, E., Ginsburg, A., Bally, J., Battersby, C., Longmore, S., & Testi, L. 2012, ApJ, 758, L28

Brogan, C. L., Frail, D. A., Goss, W. M., & Troland, T. H. 2000, ApJ, 537, 875

Brogan, C. L. et al. 2013, ApJ, 771, 91

Burkert, A. & Hartmann, L. 2012, ArXiv e-prints

Carpenter, J. M. & Sanders, D. B. 1998, AJ, 116, 1856

Clark, P. C. & Glover, S. C. O. 2013

Dale, J. E. & Bonnell, I. A. 2008, MNRAS, 391, 2

Dale, J. E., Bonnell, I. A., Clarke, C. J., & Bate, M. R. 2005, MNRAS, 358, 291

Dale, J. E., Ercolano, B., & Bonnell, I. A. 2012, MNRAS, 427, 2852

Dale, J. E., Ngoumou, J., Ercolano, B., & Bonnell, I. A. 2013

Davies, B. 2012, in Astronomical Society of the Pacific Conference Series, Vol. 465, Proceedings of a Scientific Meeting in Honor of Anthony F. J. Moffat, ed. L. Drissen, C. Rubert, N. St-Louis, & A. F. J. Moffat, 383

De Buizer, J. M. & Vacca, W. D. 2010, AJ, 140, 196

Froebrich, D. et al. 2011, MNRAS, 413, 480

Ginsburg, A., Bressert, E., Bally, J., & Battersby, C. 2012, ApJ, 758, L29

Ginsburg, A., Darling, J., Battersby, C., Zeiger, B., & Bally, J. 2011, ApJ, 736, 149

Ginsburg, A. et al. 2013a, ArXiv e-prints

—. 2013b, ApJS, 208, 14

Henkel, C., Walmsley, C. M., & Wilson, T. L. 1980, A&A, 82, 41

Hopkins, P. F., Narayanan, D., Murray, N., & Quataert, E. 2013, MNRAS

Joshi, Y. C. 2007, MNRAS, 378, 768

Juneau, S., Narayanan, D. T., Moustakas, J., Shirley, Y. L., Bussmann, R. S., Kennicutt, Jr., R. C., & Vanden Bout, P. A. 2009, ApJ, 707, 1217

Kainulainen, J., Federrath, C., & Henning, T. 2013, A&A, 553, L8

Kang, M., Bieging, J. H., Kulesa, C. A., Lee, Y., Choi, M., & Peters, W. L. 2010, ApJS, 190, 58

Kang, M., Bieging, J. H., Povich, M. S., & Lee, Y. 2009, ApJ, 706, 83

Kauffmann, J. & Pillai, T. 2010, ApJ, 723, L7

Koo, B.-C., Kim, K.-T., & Seward, F. D. 1995, ApJ, 447, 211

Koo, B.-C. & Moon, D.-S. 1997a, ApJ, 475, 194

—. 1997b, ApJ, 485, 263

Kruijssen, J. M. D. 2012, MNRAS, 426, 3008

Krumholz, M. R. et al. 2014

Krumholz, M. R., Klein, R. I., & McKee, C. F. 2007, ApJ, 665, 478

Krumholz, M. R. & McKee, C. F. 2008, Nature, 451, 1082

Kudryavtseva, N. et al. 2012, ApJ, 750, L44

Lada, C. J., Lombardi, M., & Alves, J. F. 2010, ApJ, 724, 687

Langer, W. D., Velusamy, T., Pineda, J. L., Willacy, K., & Goldsmith, P. F. 2013, ArXiv e-prints

Langston, G., Minter, A., D'Addario, L., Eberhardt, K., Koski, K., & Zuber, J. 2000, AJ, 119, 2801

Lee, K. I., Looney, L. W., Schnee, S., & Li, Z.-Y. 2013

Longmore, S. N. et al. 2012, ApJ, 746, 117

Mangum, J. G., Darling, J., Henkel, C., & Menten, K. M. 2013

Mangum, J. G., Emerson, D. T., & Greisen, E. W. 2007, A&A, 474, 679

Mangum, J. G. & Wootten, A. 1993, ApJS, 89, 123

Martin-Pintado, J., Wilson, T. L., Henkel, C., & Gardner, F. F. 1985a, A&A, 142, 131

Martin-Pintado, J., Wilson, T. L., Johnston, K. J., & Henkel, C. 1985b, ApJ, 299, 386

Mehringer, D. M. 1994, ApJS, 91, 713

Miura, R. E. et al. 2014, ApJ, 788, 167

Muraoka, K. et al. 2009, PASJ, 61, 163

Myers, A., Klein, R., Krumholz, M., & McKee, C. 2014, ArXiv e-prints

Parker, R. J. & Dale, J. E. 2013

Parsons, H., Thompson, M. A., Clark, J. S., & Chrysostomou, A. 2012, MNRAS, 424, 1658

Pineda, J. E. & Teixeira, P. S. 2013, ArXiv e-prints

Reed, B. C. 2006, JRASC, 100, 146

Rogers, H. & Pittard, J. M. 2013, MNRAS, 431, 1337

Sato, M., Reid, M. J., Brunthaler, A., & Menten, K. M. 2010, ApJ, 720, 1055

Smith, R. J., Glover, S. C. O., Clark, P. C., Klessen, R. S., & Springel, V. 2014, ArXiv e-prints

Stil, J. M. et al. 2006, AJ, 132, 1158

Sun, X. H., Han, J. L., Reich, W., Reich, P., Shi, W. B., Wielebinski, R., & Fürst, E. 2007, A&A, 463, 993

Sun, X. H., Reich, P., Reich, W., Xiao, L., Gao, X. Y., & Han, J. L. 2011a, A&A, 536, A83

Sun, X. H., Reich, W., Han, J. L., Reich, P., Wielebinski, R., Wang, C., & Müller, P. 2011b, A&A, 527, A74

Tian, W. W. & Leahy, D. A. 2013, ApJ, 769, L17

Urquhart, J. S., Figura, C. C., Moore, T. J. T., Hoare, M. G., Lumsden, S. L., Mottram, J. C., Thompson, M. A., & Oudmaijer, R. D. 2014, MNRAS, 437, 1791

van der Tak, F. F. S., Black, J. H., Schöier, F. L., Jansen, D. J., & van Dishoeck, E. F. 2007, A&A, 468, 627

Walmsley, C. M. 1990, A&AS, 82, 201

Wilson, T. L., Rohlfs, K., & Hüttemeister, S. 2009, Tools of Radio Astronomy (Springer-Verlag)

Young, K. E., Lee, J.-E., Evans, II, N. J., Goldsmith, P. F., & Doty, S. D. 2004, ApJ, 614, 252