A&A 580, A112 (2015)
DOI: 10.1051/0004-6361/201425370

Astronomy & Astrophysics

THOR: The H I, OH, Recombination line survey of the Milky Way

The pilot study: H I observations of the giant molecular cloud W43⋆

S. Bihr[1], H. Beuther[1], J. Ott[2], K. G. Johnston[3], A. Brunthaler[4], L. D. Anderson[5], F. Bigiel[6], P. Carlhoff[7], E. Churchwell[8], S. C. O. Glover[6], P. F. Goldsmith[9], F. Heitsch[10], T. Henning[1], M. H. Heyer[11], T. Hill[12], A. Hughes[1], R. S. Klessen[6,13,14], H. Linz[1], S. N. Longmore[15], N. M. McClure-Griffiths[16], K. M. Menten[4], F. Motte[17], Q. Nguyen-Lu'o'ng[18], R. Plume[19], S. E. Ragan[1], N. Roy[4,20], P. Schilke[7], N. Schneider[21], R. J. Smith[6], J. M. Stil[19], J. S. Urquhart[4], A. J. Walsh[22], and F. Walter[1]

(Affiliations can be found after the references)

Received 19 November 2014 / Accepted 11 May 2015

ABSTRACT

To study the atomic, molecular, and ionized emission of giant molecular clouds (GMCs) in the Milky Way, we initiated a large program with the *Karl G. Jansky* Very Large Array (VLA): "THOR: The H I, OH, Recombination line survey of the Milky Way". We map the 21 cm H I line, 4 OH lines, up to 19 Hα recombination lines and the continuum from 1 to 2 GHz of a significant fraction of the Milky Way (l = 15°−67°, $|b| \leq 1°$) at an angular resolution of ∼20″. Starting in 2012, as a pilot study we mapped 4 square degrees of the GMC associated with the W43 star formation complex. The rest of the THOR survey area was observed during 2013 and 2014. In this paper, we focus on the H I emission from the W43 GMC complex. Classically, the H I 21 cm line is treated as optically thin with properties such as the column density calculated under this assumption. This approach might yield reasonable results for regions of low-mass star formation, however, it is not sufficient to describe GMCs. We analyzed strong continuum sources to measure the optical depth along the line of sight, and thus correct the H I 21 cm emission for optical depth effects and weak diffuse continuum emission. Hence, we are able to measure the H I mass of this region more accurately and our analysis reveals a lower limit for the H I mass of $M = 6.6_{-1.8} \times 10^6\ M_\odot$ ($v_{\rm LSR}$ = 60−120 km s^{-1}), which is a factor of 2.4 larger than the mass estimated with the assumption of optically thin emission. The H I column densities are as high as $N_{\rm H\,I} \sim 150\ M_\odot\,{\rm pc}^{-2} \approx 1.9 \times 10^{22}\,{\rm cm}^{-2}$, which is an order of magnitude higher than for low-mass star formation regions. This result challenges theoretical models that predict a threshold for the H I column density of ∼10 $M_\odot\,{\rm pc}^{-2}$, at which the formation of molecular hydrogen should set in. By assuming an elliptical layered structure for W43, we estimate the particle density profile. For the atomic gas particle density, we find a linear decrease toward the center of W43 with values decreasing from $n_{\rm H\,I} = 20\ {\rm cm}^{-3}$ near the cloud edge to almost 0 cm^{-3} at its center. On the other hand, the molecular hydrogen, traced via dust observations with the *Herschel* Space Observatory, shows an exponential increase toward the center with densities increasing to $n_{\rm H_2} > 200\,{\rm cm}^{-3}$, averaged over a region of ∼10 pc. While atomic and molecular hydrogen are well mixed at the cloud edge, the center of the cloud is strongly dominated by H_2 emission. We do not identify a sharp transition between hydrogen in atomic and molecular form. Our results, which challenge current theoretical models, are an important characterization of the atomic to molecular hydrogen transition in an extreme environment.

Key words. ISM: clouds – ISM: structure – ISM: atoms – stars: formation – radio lines: ISM – surveys

1. Introduction

It is well known that stars form in giant molecular clouds (GMCs; Mac Low & Klessen 2004; McKee & Ostriker 2007; Dobbs et al. 2014; Offner et al. 2014), which primarily consist of molecular hydrogen. Yet it is still under debate whether molecular hydrogen is actually necessary for star formation or whether molecular hydrogen and stars only form under the same conditions side by side (Glover & Clark 2012). The density within these clouds is high enough (particle density: $n > 1{-}100$ cm^{-3}, column density: $N > 1{-}100\ M_\odot\,{\rm pc}^{-2} \sim 10^{20}{-}10^{22}\,{\rm cm}^{-2}$) for molecular hydrogen to become self-shielded from the interstellar radiation field, which would otherwise dissociate the H_2 molecules (e.g., Dobbs et al. 2014). Hence, molecular clouds form in the interior of large clouds of neutral hydrogen, which themselves are the environment of molecular clouds. Another open question is the fraction of neutral hydrogen within molecular clouds and how this affects physical conditions (Krčo & Goldsmith 2010; Goldsmith & Li 2005). Furthermore, the correlation between the transition between these two fundamental states of hydrogen and the corresponding physical conditions is also still unclear.

As cold molecular hydrogen is challenging to observe directly, it is difficult to study its distribution in detail. Classically, the low-J rotational transitions of CO are used as a tracer for H_2, however, recent simulations and observations show that a large amount (∼40%) of H_2 is CO-dark and therefore not well traced by CO (Pineda et al. 2013; Smith et al. 2014). Another approach to study molecular hydrogen is via observations of thermal dust emission or dust extinction (Lada et al. 2007; Molinari et al. 2010; Kainulainen & Tan 2013; Kainulainen et al. 2013). In contrast, the 21 cm spin-flip transition of hydrogen offers a well-known method to measure the atomic gas content. Even

⋆ The H I data cubes are only available at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/580/A112

though the 21 cm line is well studied (e.g., Radhakrishnan et al. 1972; Gibson et al. 2000, 2005a; Taylor et al. 2003; Heiles & Troland 2003a,b; Strasser & Taylor 2004; Goldsmith & Li 2005; Stil et al. 2006; Kalberla & Kerp 2009; McClure-Griffiths et al. 2012; Roy et al. 2013a; Fukui et al. 2015; Murray et al. 2014; Motte et al. 2014), it is difficult to disentangle the different contributions of the cold and warm H I in emission and absorption for different spin temperatures and optical depths. In addition, radio continuum emission at the frequency of the H I emission line can also suppress the intensity of the observed H I emission, an effect that is especially significant for the Galactic plane.

To address cloud formation, H I to H_2 formation as well as many other issues in ISM studies of the Milky Way, we initiated the THOR project: the H I, OH, Recombination line survey of the Milky Way (Beuther et al., in prep.). We are using the *Karl G. Jansky* Very Large Array (VLA) in a large program to observe the H I line at 21 cm, four OH lines at 1612, 1665, 1667, and 1720 MHz, 19 Hα radio recombination lines (RRL) and the continuum from 1–2 GHz, which is split in 8 sub-bands. These observations result in an angular resolution of 15–20″. We were awarded more than 200 h observing time to map a significant fraction of the Milky Way (Galactic longitude l = 15–67°, Galactic latitude $|b| \leq 1°$) for the semester 2013A and 2014B. To test our observing strategy and data reduction, we began with a pilot study of a 2° × 2° field around W43, which was completed in 10 h. In this paper, we describe the pilot study, which was observed in the semester 2012A. As the wealth of this data set is immense, we focus on the H I line here, but consecutive papers will focus on other aspects of the pilot study (e.g., Walsh et al. 2015; Johnston et al., in prep.). The full survey will be presented in Beuther et al. (in prep.).

The field chosen for the pilot study is around the massive star-forming complex W43 (l = 29.2–31.5°, $|b| \leq 1°$). This region is situated at the intersection of the Galactic bar and the first spiral arm (Scutum-Centaurus Galactic arm, Nguyen Luong et al. 2011; Carlhoff et al. 2013), leading to complex kinematic structures and possibly high star formation activity. The complex W43 is referred to as a Galactic mini-starburst region (Motte et al. 2003; Bally et al. 2010) and shows a star formation rate of ~0.1 $M_\odot\,\mathrm{yr}^{-1}$ (Nguyen Luong et al. 2011) or 5–10% of the star formation rate in the entire Milky Way. Motte et al. (2014) found velocity gradients in CO and H I 21 cm emission. These velocity gradients could indicate large scale velocity flows, which could cause vast star formation activity. Motte et al. (2014) also do not find a threshold for the H I column density, which is proposed by theoretical models (Krumholz et al. 2008, 2009). They argue that we see several transition layers of H I and H_2 along the line of sight and that the theoretical models are not suited to describe complicated molecular cloud complexes such as W43. The center of W43 harbors a large H II region, which is fueled by a Wolf-Rayet and OB star cluster (Liszt et al. 1993; Lester et al. 1985; Blum et al. 1999). Furthermore, W43 exhibits several high-mass starless molecular clumps, which are still in an early stage of star formation (Beuther et al. 2012; Louvet et al. 2014). Some massive dense clumps can potentially form young massive clusters, progenitors of globular clusters (Louvet et al. 2014; Nguyen-Lu'o'ng et al. 2013). The Bar and Spiral Structure Legacy survey (BeSSeL; Brunthaler et al. 2011; Reid et al. 2014) determined the distance to W43 to be 5.5 ± 0.4 kpc from parallax measurements of methanol and water masers (Zhang et al. 2014). This result has to be treated cautiously, as none of the four masers used for the parallax measurements are spatially directly associated with W43-Main.

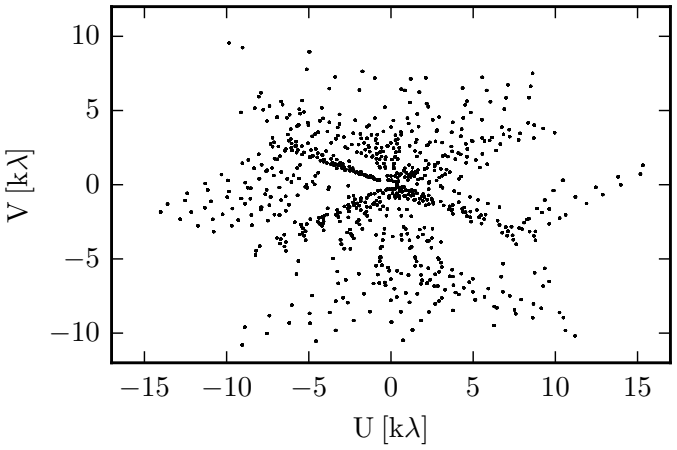

Fig. 1. UV coverage of one pointing (centered at RA 18:47:54.125746, Dec –03:28:42.90941, J2000) after 4 × 2 min observing time.

2. Observations and data reduction

2.1. VLA Observations

We mapped a 2°×2° field around W43 (l = 29.2–31.5°, $|b| \leq 1°$) during the 2012A semester with the *Karl G. Jansky* Very Large Array (VLA) in New Mexico in C-configuration (Project 12A-161). As we used Nyquist sampling at 1.42 GHz with a primary beam size of 32′, we needed 59 pointings to cover the 4 square degree mosaic. We chose a hexagonal geometry for the mosaic, similar to the VGPS survey (Stil et al. 2006). This results in a smooth areal sensitivity function with fluctuations of less than 1% in the inner region and a decreasing sensitivity toward the edges of our field. We observed each pointing for 4 × 2 min, which results in an overall observation time of ten hours, including ~2 h overhead for flux, bandpass, and complex gain calibration. The resulting uv-coverage for one pointing, after 4 × 2 min of observing time, is shown in Fig. 1. We performed the observations in two blocks each with 5 h observing time in April 2012. We chose the quasar 3C 286 as a flux and bandpass calibrator and the quasar J1822-0938 as a complex gain calibrator, which was observed every ~13 min. Observing at L-band and using the new WIDAR correlator, we were able to simultaneously observe the H I 21 cm line, 4 OH lines (1612, 1665, 1667 and 1720 MHz) and 12 Hα RRL. For the pilot field, we had the spectral capability to observe 12 Hα RRL, however, for the full THOR survey we were able to observe 19 Hα RRL. The continuum, consisting of eight spectral windows between 1 and 2 GHz, was observed in full polarization. For the H I 21 cm line, we used a bandwidth of 2 MHz with a channel width of 1.953 kHz. This results in a velocity range of ±200 km s^{-1} and a channel spacing of $\Delta v \approx 0.41$ km s^{-1}.

2.2. Calibration

To edit and calibrate the data, we use CASA (version 4.1.0) with a modified VLA pipeline[1] (version 1.2.0). The pipeline does automatic flagging for, e.g., zeros or shadowing of antennas. We manually performed additional flagging for radio frequency interference (RFI) and bad antennas. The pipeline also applies the bandpass, flux, and complex gain calibrator to the data. We do not use Hanning smoothing and do not recalculate the data

[1] https://science.nrao.edu/facilities/vla/data-processing/pipeline

weights ("statwt") at the end of the pipeline. The implemented modifications help to improve the quality check, and we do some flagging and editing by hand subsequently. A full description of our quality check method will be presented in our forthcoming overview paper (Beuther et al., in prep.).

2.3. Imaging and deconvolution

The H I emission and absorption covers a large range of angular scales, which is challenging for data reduction as the VLA C-Array filters out most of the large scale structure. Therefore, we combined our THOR data with the VGPS data (VLA D-Array and Green Bank Telescope data, Stil et al. 2006) to overcome missing flux problems and to reconstruct the large scale structure. We tested three different methods to combine the THOR and VGPS data. First, we tried to combine the visibility of the VGPS and THOR data and clean them together. Second, we tried the "feather" command in CASA and third, we used the images of the VGPS survey as a starting model (parameter "modelimage") for the clean process in CASA. After testing these three methods, we choose the third method, as this method provides the best results considering noise, side lobes, and recovery of large scale structure. We compared point source peak fluxes of the combined images to the VGPS data to check for consistency. The difference is at the 10% level. The clean process was stopped at the 5σ noise level.

We smoothed over three channels to reduce the noise, resulting in a velocity resolution of 1.24 km s^{-1}. This was the best compromise between computational time, noise, and a sufficient spectral resolution to distinguish absorption/emission features. The synthesized beam was set to 20″, which is slightly larger than the best resolution achievable (∼16″). The weighting parameter was set to robust = 0.5, which is a combination of natural and uniform weighting. These methods result in an rms of ∼14 K (9 mJy beam^{-1}) for emission free channels and up to a factor of 2 or 3 more in channels with strong emission due to systematic errors, such as side lobes. The dynamical range of our data set is ∼100−200, depending on the region. In the following, the H I absorption and small scale structure is based on the THOR data, whereas the large scale emission is based on the VGPS data.

The 21 cm continuum data are taken from the H I data cube for high and low velocities (−80 to −50 and 135 to 155 km s^{-1}), which are not affected by H I emission or absorption. Therefore the data reduction and imaging for the 21 cm continuum data is the same as for the H I data, and we can avoid systematic errors due to different data reduction methods.

2.4. H_2 column density

As we will compare the H I and H_2 column density in Sect. 5.4, we need an estimate of the H_2 content. We use dust observations from the *Herschel* Space Observatory to assess the H_2 column density. These observations are based on the HiGal survey (Molinari et al. 2010). The H_2 column density map is taken from Fig. 9 of Nguyen-Lu'o'ng et al. (2013), which was derived by SED fitting methods described by Hill et al. (2011). As the dust observations have no velocity resolution, we see all the dust and thus gas along the line of sight. Carlhoff et al. (2013) showed that the *Herschel* dust data are similar within the uncertainties to the CO data at the velocity range of W43 ($v_{\rm LSR}$ = 60−120 km s^{-1}). Hence, the *Herschel* dust observations are dominated by the emission from W43 and the contributions of other regions along the line of sight can be neglected. We refrain from using CO data to estimate the molecular content of W43, as CO does not trace all molecular hydrogen (e.g., Smith et al. 2014) and CO becomes optically thick for the dense interior of W43 (Carlhoff et al. 2013).

3. H I radiative transfer

In this section, we explain the methodology we used to determine the spin temperature, optical depth and column density of the neutral hydrogen. Even though the H I 21 cm line is a well-known probe of these quantities, the arrangement of different H I components with different temperatures along the line of sight can complicate these measurements. First we explain the classical methods to determine the column density via optically thin emission and H I continuum absorption (HICA). Subsequently, we outline the limitations of these methods and describe our approach to correct for optical depth and weak continuum emission. Although the basics of this approach are discussed in the literature (e.g., Wilson et al. 2010), we modified the classical approach to account for the continuum emission and optical depth. Therefore, we outline the calculations in more detail.

All following values are frequency dependent. To keep the equations and description simple, we drop the frequency dependencies in the equations. We emphasize, however, that the following emission/absorption mechanisms only works for identical frequencies and, therefore, identical velocities.

3.1. Column density

The classical approach to determine the H I column density is given by Wilson et al. (2010) as

$$\frac{N_{\rm H}}{\rm cm^{-2}} = 1.8224 \times 10^{18} \, \frac{T_{\rm S}}{\rm K} \int_{-\infty}^{\infty} \tau(v) \, d\left(\frac{v}{\rm km\,s^{-1}}\right), \tag{1}$$

where $N_{\rm H}$, $T_{\rm S}$ and τ are the column density, the spin temperature, and the velocity dependent opacity, respectively. The spin temperature $T_{\rm S}$ describes the relative population of the spin states of the hydrogen atom (Wilson et al. 2010). As $T_{\rm S}$ is the equivalent of the excitation temperature for molecules, it is only equal to the kinetic temperature in local thermodynamic equilibrium (i.e., when there are enough collisions to thermalize the gas). We assume in the following that the spin temperature $T_{\rm S}$ does not vary significantly within one velocity channel. In most cases, neither $T_{\rm S}$ nor the optical depth are known. The simplest assumption is that the H I emission is optically thin (see Sect. 3.2).

3.2. Optically thin H I emission

Under the optically thin assumption, without background continuum emission, the expression for the brightness temperature $T_{\rm B}$ simplifies to the following:

$$T_{\rm B} = T_{\rm S} \, (1 - \exp(-\tau)) \approx T_{\rm S} \, \tau. \tag{2}$$

This simplification provides a linear relation between the column density (Eq. (1)) and the brightness temperature. Hence, we can estimate the column density directly from the measured brightness temperature $T_{\rm B}$. This method is used in several studies and is well described in the literature (e.g., Lee et al. 2012; Motte et al. 2014; Wilson et al. 2010). Below, we discuss its limitations and describe a procedure to overcome them.

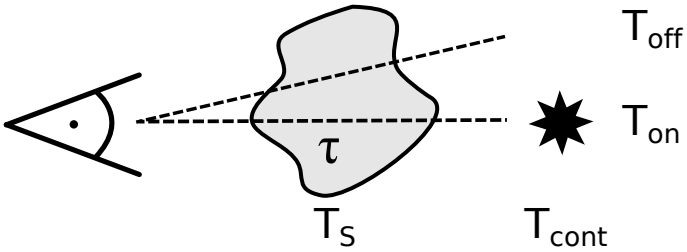

Fig. 2. Sketch of the arrangement of the H I cloud with spin temperature T_S and optical depth τ and the continuum source in the background. The on and off positions are marked.

3.3. H I optical depth

The H I continuum absorption method is the classical method to determine properties of the cold neutral medium (Strasser & Taylor 2004; Heiles & Troland 2003a,b; Strasser et al. 2007). This method is based on strong continuum sources (T_B > 300 K), such as Galactic H II regions, active galactic nuclei (AGNs), or extragalactic jets. As the brightness temperature of these continuum sources is larger than the typical spin temperature of the H I clouds ($T_S \sim 100$ K), we observe the H I cloud in absorption. In addition, the absorption spectrum is dominated by the cold neutral medium, as the H I absorption coefficient is proportional to T^{-1}.

The classical observing strategy is the "on-off" observation (see Fig. 2). The "on-source" points toward the continuum source and reveals the absorption spectrum (T_{on}), whereas the "off-source" points slightly offset from the continuum source and reveals the emission spectrum (T_{off}). During data reduction, the continuum is not subtracted from the H I spectrum. Hence, we can measure T_{cont} for channels that are not affected by the H I line. The general assumption is that the on-source and off-source spectra originate from the same cloud with the same properties. Therefore, it is important to have these two positions as close together as possible.

The measured on-source and off-source brightness temperatures are

$$\begin{aligned} T_{on} &= T_S\,(1 - \exp(-\tau)) + T_{cont}\,\exp(-\tau), \\ T_{off} &= T_S\,(1 - \exp(-\tau)). \end{aligned} \tag{3}$$

Hence, the optical depth is

$$\tau = -\ln\left(\frac{T_{on} - T_{off}}{T_{cont}}\right). \tag{4}$$

The spin temperature can then be determined from

$$T_S = \frac{T_{off}}{1 - e^{-\tau}}. \tag{5}$$

The advantage of the HICA method is the direct measurement of the optical depth and the spin temperature. The challenge of the HICA method is the need for strong continuum sources. Since most of the strong continuum sources are point sources, it is not possible to map the entire Milky Way, but instead the result is an incomplete grid of measurements. This way, it is not possible to map the intrinsic structure of individual clouds. Furthermore, the spin temperature does not describe the actual temperature of the cloud, but is a mean of the cold and warm component, weighted by their column densities (Strasser & Taylor 2004). Consequently, the derived spin temperature is an upper limit for the cold component.

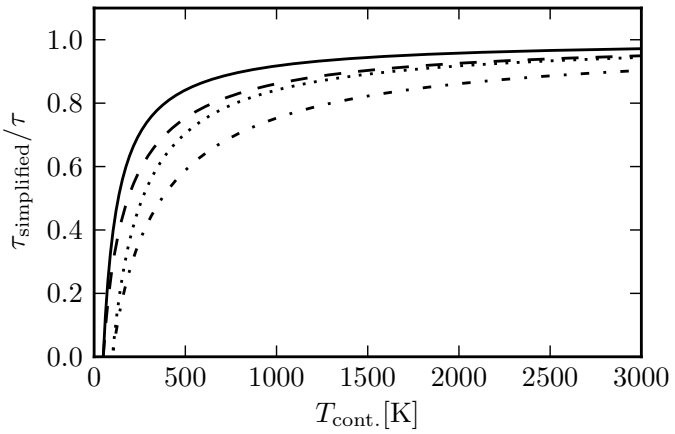

Fig. 3. Optical depth simplification. The simplified optical depth is calculated by neglecting the spin temperature of the cloud. The solid and dashed lines represent a spin temperature of 50 K with optical depths of 1 and 2, respectively. The dotted and dash-dotted lines represent a spin temperature of 100 K with optical depths of 1 and 2, respectively.

As some Galactic continuum sources, such as W43-Main, are extended, it is difficult to determine a proper off position. Nevertheless, it is still possible to determine the optical depth. As we use the VLA C+D array observations, including the continuum without the GBT data, we filter out most smooth large scale structure. The VLA C+D array data show H I emission of less than 30 K. Therefore, we can neglect the emission of the H I cloud in Eq. (3) and set $T_{off} = T_S\,(1 - \exp(-\tau)) = 0$ and calculate the optical depth without any off position as

$$\tau_{simplified} = -\ln\left(\frac{T_{on}}{T_{cont}}\right). \tag{6}$$

Even without the effect of filtering out the H I emission due to the interferometer, this simplification holds for strong continuum sources (T_{cont} > 500 K). Figure 3 shows the relation between the simplified and actual optical depth as a function of the continuum brightness temperature, not taking the filtering of the interferometer into account. It clearly shows that even for high spin temperatures ($T_S = 100$ K) and high optical depth ($\tau = 2$), we miss less than 10% of the optical depth for a strong continuum source ($T_{cont} \approx 2500$ K). Therefore, we are able to measure the optical depth for the H II region in the center of W43, even though we cannot determine a proper off position.

We also consider the effect of optically thick clouds. For these optically thick clouds, the absorption spectra approaches zero. Due to the rms of the spectra, it is also possible that the absorption spectra become negative, which is physically not meaningful. We set the optical depth to a lower limit for all absorption values that are close to zero. If the absorption spectrum T_{on} is smaller than 5 times the rms, we set the optical depth to

$$\tau_{lower-limit} = -\ln\left(\frac{5 \cdot \sigma(T_{on})}{T_{cont}}\right). \tag{7}$$

We find that the optical depth saturates for the main velocity range of W43-Main, which has consequences for our interpretation and conclusions.

The uncertainty of the optical depth (Eq. (4)) depends on T_{on}, T_{off} and T_{cont}. To estimate the uncertainty, we assume an uncertainty of 20 K for all three quantities. The uncertainty of the optical depth is shown in Fig. 4 as a function of the optical depth itself for three different continuum brightness values. It shows

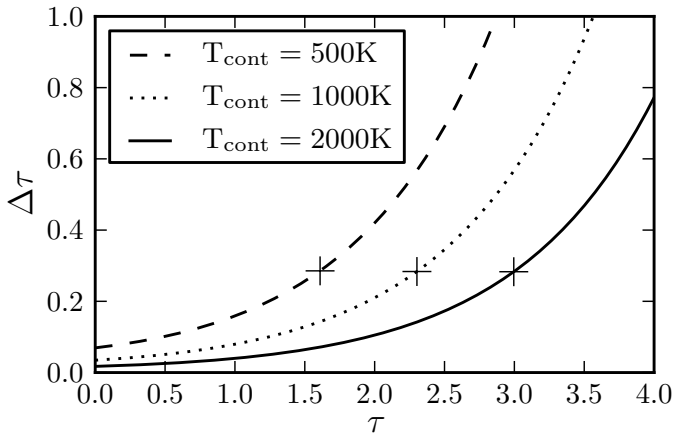

Fig. 4. Absolute uncertainty of the optical depth as a function of the optical depth itself for three different continuum background sources. The dashed, dotted, and solid line represent continuum sources with a brightness temperature of $T_{\rm cont}$ = 500 K, 1000 K and 2000 K, respectively. The crosses on each line show the lower limit of the optical depth that we can observe for the given continuum brightness temperature. For larger optical depths, the absorption spectra saturates (see Sect. 4.1 for more detail).

that the uncertainty increases significantly for increasing optical depths. However, we are not able to measure these high optical depths, as the absorption spectra saturates for optically thick clouds, and we report lower limits (see Eq. (7)). These lower limits are marked with crosses in Fig. 4. Up to the lower limit of the optical depth the corresponding uncertainty is ∼0.3 for the three different continuum brightness values. For strong continuum sources, such as W43-Main, the uncertainty of the optical depth is ∼10% up to the lower limit of ∼3.

3.4. Column density corrections

In addition to distinct, small continuum sources, such as H II regions or extragalactic jets, we also find weak diffuse continuum emission in the Galactic plane. This component has a strength between 10 and 50 K. Therefore, it is not strong enough to induce absorption features (HICA), but nevertheless it can influence the H I emission. To overcome this problem, the classical approach is to subtract this weak continuum emission during data reduction. Hence, we subtract $T_{\rm cont}$ in Eq. (3), and we find

$$\begin{aligned} T_{\rm B} &= T_{\rm S}(1-\exp(-\tau)) + T_{\rm cont}\ \exp(-\tau) - T_{\rm cont} \\ &= (T_{\rm S} - T_{\rm cont})\ (1-\exp(-\tau)). \end{aligned} \tag{8}$$

This shows that even if we subtract the continuum from our data, it can still influence the measured brightness temperature. If we neglect the weak continuum emission, our measured H I emission is suppressed and thus the calculated column density is underestimated. Therefore it is important to investigate the influence of the weak diffuse continuum emission.

In the following, we assess the assumption of optically thin emission and the influence of weak diffuse continuum emission on the determination of the measured column density for appropriate model clouds. We assume two model clouds, which have a spin temperature of 50 K and 100 K, respectively. We use Eq. (1) to calculate the expected column density $N_{\rm H\,I}$ for different optical depths in one velocity channel (dv = 1.24 km s^{-1}). Furthermore, we assume the cloud to be optically thin and use Eq. (2) to calculate the brightness temperature of the H I emission. Using this result, we can calculate the column density of the cloud, but this time with the simplification of optically thin emission. Hence, we call it the observed column density $N_{\rm H\,I}(observed)$. The solid lines in Fig. 5 show the ratio of the expected and observed column density for a range of optical depths. Obviously, for small optical depths ($\tau < 0.1$), our assumption of optically thin emission is sufficient and we observe more than 90% of the expected column density. For larger optical depths ($\tau > 0.1$), however, we miss a significant fraction of the column density (>40%).

In addition, we add weak continuum emission in the background, which changes the brightness temperature according to Eq. (8) and therefore suppresses the H I emission. Nevertheless, if we still assume optically thin emission and do not consider the weak continuum emission, we can calculate the observed column density. The dashed and dotted lines in Fig. 5 show this case for continuum emission of 10 K and 30 K, respectively. Even for small optical depths ($\tau < 0.1$) we miss a significant fraction of the column density, which depends on the ratio of the spin temperature and the continuum emission brightness temperature. In the worst case ($T_{\rm S}$ = 50 K and $T_{\rm cont}$ = 30 K), we observe only 40% of the expected column density, even for small optical depths.

To summarize, we measure the brightness temperature of the H I emission $T_{\rm B}$, as well as the brightness temperature for the continuum $T_{\rm cont}$ and combine this information with the optical depth τ, which we measure toward strong continuum sources. This allows us to calculate the corrected column density, by combining Eqs. (1) and (8), which yields

$$N_{\rm H} = 1.8224 \times 10^{18} \left(\frac{T_{\rm B}}{1-e^{-\tau}} + T_{\rm cont}\right) \int_{-\infty}^{\infty} \tau(v)\, {\rm d}v, \tag{9}$$

where the column density $N_{\rm H}$ is given in units cm^{-2}, the brightness temperature $T_{\rm B}$ and the continuum brightness temperature $T_{\rm cont}$ are measured in K, and the velocity dv is given in km s^{-1}.

3.5. Continuum correction for strong point sources

For the H I continuum absorption (HICA) toward strong continuum sources, we also have to consider weak diffuse continuum emission, which contributes to the on and off positions. Therefore we have to modify the equation for the on and off positions (Eq. (3)) by adding another term for the weak diffuse continuum emission, which we call $T_{\rm cont,dif}$. This does not change our result for the optical depth (Eq. (4)), as the optical depth depends solely on the difference between $T_{\rm on}$ and $T_{\rm off}$. For the calculated spin temperature, however, we have to modify Eq. (5) to the following:

$$T_{\rm S} = \frac{T_{\rm off} - T_{\rm cont,dif}\ {\rm e}^{-\tau}}{1-{\rm e}^{-\tau}}. \tag{10}$$

The effect of the diffuse continuum correction is discussed in Sect. 5.2.

4. Results

4.1. Optical depths determined using compact continuum sources

As described in Sect. 3.3, we can use strong continuum sources in the background as light-houses that shine through foreground H I clouds and create absorption spectra. Below we characterize the continuum sources.

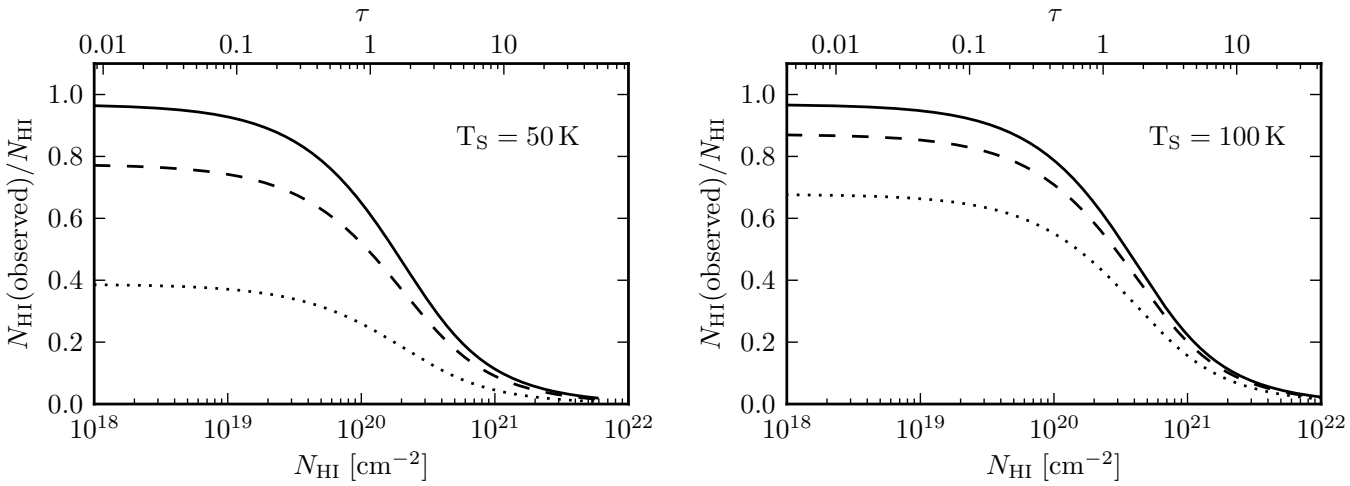

Fig. 5. Ratio of observed and expected column density for a model cloud with spin temperatures of $T_{\rm S}$ = 50 K and 100 K for the *left and right panel*, respectively. The solid lines show the ratio for the assumption of optically thin emission. The dashed and dotted lines show the ratio for the assumption of optically thin emission and a diffuse continuum source with brightness 10 K and 30 K, respectively.

Table 1. Continuum point sources that were used for H I absorption studies.

Name	RA(2000) (h:m:s)	Dec(J2000) (°: ′ : ″)	Maj × Min (″ × ″)	$T_{\rm cont}$ K	$T_{\rm cont,dif}$ K	Far Scutum arm	Spectral index	Location
G31.388−0.384	18:49:59.1	−1.32.55.9	21.3 × 20.3	2055 ± 24	20.2	yes	−0.82	ext. gal
G31.411+0.307	18:47:34.1	−1.12.45.5	24.0 × 23.0	914 ± 26	19.4	no	0.63	gal.
G30.234−0.138	18:47:00.4	−2.27.52.9	21.7 × 20.7	853 ± 23	42.1	no	0.00	gal.
G30.534+0.021	18:46:59.3	−2.07.27.9	28.0 × 21.5[a]	653 ± 21	35.9	no	0.21	gal.
G31.242−0.110	18:48:44.7	−1.33.13.7	25.1 × 21.2	573 ± 22	30.1	no	0.58	gal.
G29.090+0.512	18:42:35.8	−3.11.04.7	20.0 × 10.0	571 ± 44	20.5	yes	−1.03	ext. gal
G30.744+1.008	18:43:51.4	−1.29.14.8	23.1 × 20.1	472 ± 36	16.9	yes	−0.90	ext. gal
G30.699−0.630	18:49:36.3	−2.16.27.2	22.3 × 19.6	403 ± 18	23.4	yes	−1.11	ext. gal

Notes. Continuum point sources in the observed field with brightness temperatures $T_{\rm cont} > 400$ K. The names and spectral indexes are based on work presented in Johnston et al. (in prep.) and are preliminary results. Coordinates, major, and minor axes of the 2D fit, maximum brightness temperature $T_{\rm cont}$ of the continuum point source and the diffuse weak continuum temperature around the point source are given. The column "far Scutum arm" indicates whether absorption features at $v_{\rm LSR} \approx -40$ km s^{-1} are present in the H I spectra. The spectral index and the absorption of the far Scutum arm are used to determine the location of the point sources, which is given in the last column. [(a)] This source consists of two blended sources, which explains the large eccentricity.

To determine the optical depth of the H I accurately, we need continuum sources in the background which are brighter than the typical H I spin temperature (see Sect. 3.3). We extract all continuum sources in our field that have a brightness temperature $T_{\rm cont} > 400$ K and a point-like structure, which yields eight point sources. The analysis of extended sources follows in Sect. 4.4. We use a two-dimensional Gaussian to fit the position and size of the continuum source. To measure the continuum brightness temperature, we average over high- and low-velocity channels in our H I data cubes (−80 to −50 and 135 to 155 km s^{-1}) that are not affected by the H I line. The results are shown in Table 1.

To determine the off spectrum ($T_{\rm off}$), we average the emission spectrum around the continuum source within an annulus with inner and outer radius of 60″ and 120″, respectively. The upper two panels of Fig. 6 show a typical emission and absorption spectrum. We use Eq. (4) to calculate the optical depth for each channel. To avoid unrealistic optical depths, we only calculate the optical depth for those channels in which the emission/absorption is five times larger than the corresponding noise. The gray shaded areas in the upper three panels in Figs. 6 and 7 show the 5 sigma level. The optical depth is set to zero for channels, where the emission/absorption is below five times the corresponding noise.

As explained in Sect. 3.3, we have to consider the effect of optically thick clouds for which we can determine only lower limits for the optical depth. The gray shaded areas in the third panels of Figs. 6 and 7 show these lower limits of the optical depth.

4.2. H I spin temperature toward compact continuum sources

For each channel that allows for an optical depth measurement, we compute the spin temperature using Eq. (10). The spin temperature is shown in the fourth panel of Figs. 6 and 7. This method reveals absorption features in 655 channels for all eight continuum sources. The median of the spin temperature is 97.5 K and the distribution of all absorption features is shown in Fig. 8. For large optical depths, the spin temperature approaches the brightness temperature of the off position, however, for large optical depths ($\tau \gtrsim 2$) we can only report lower limits for τ. Hence, we overestimate the spin temperature in the optically thick regime. Nevertheless, this overestimation is small, as for optical depths of $\tau = 2$ the measured spin temperature is $T_{\rm spin} \approx 1.16 \cdot T_{\rm off}$ (see Eq. (5)). Therefore we overestimate $T_{\rm spin}$ by at most ∼ 15%. We discuss these results in detail in Sect. 5.2.

4.3. Location of continuum sources

To characterize and understand the H I absorption spectrum toward the continuum point sources, it is important to know the

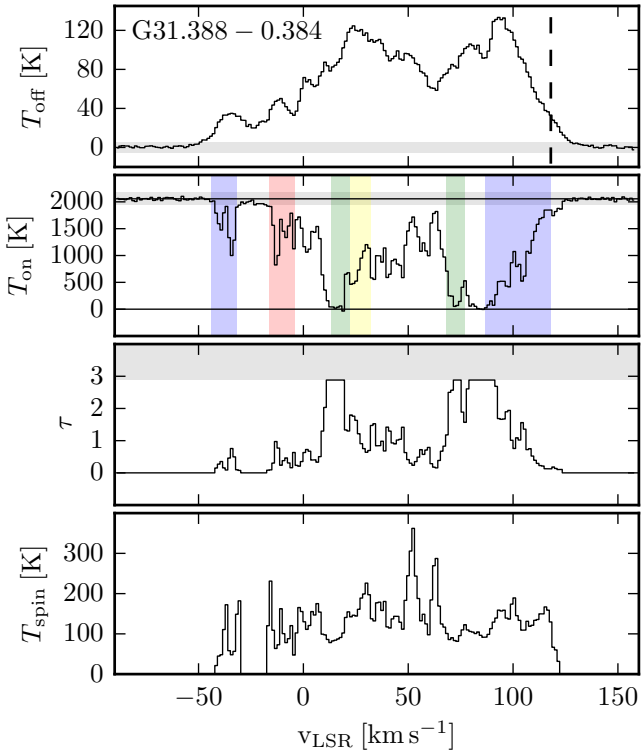

Fig. 6. H I emission and absorption spectrum of the extragalactic continuum point source G31.388−0.384. The emission spectrum is shown in the first panel and is measured in an annulus around the point source with inner and outer radii of 60″ and 120″, respectively (corresponds to 3 and 6 times the restoring beam). The second panel presents the absorption spectrum toward the point source. The color shaded areas represent the approximate velocities of the Milky Way spiral arms (Vallée 2008) in blue, red, green, and yellow, the Scutum-Centaurus, Cygnus, Sagittarius, and Perseus arm, respectively; the dashed black line in the first panel indicates the tangential velocity. In the *first two panels*, the gray shaded area indicates the 5σ noise level. The *third panel* shows the optical depth computed using Eq. (4) and the gray shaded area indicates the saturated optical depth limit, computed using Eq. (7). In *the fourth panel*, the spin temperature is presented, which is computed using Eq. (10).

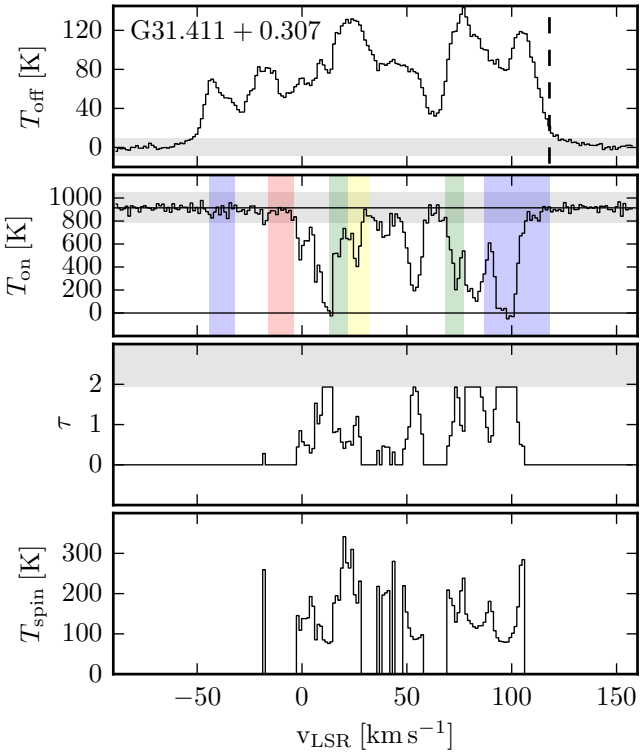

Fig. 7. Same layout as Fig. 6, except for the Galactic continuum source G31.411+0.307.

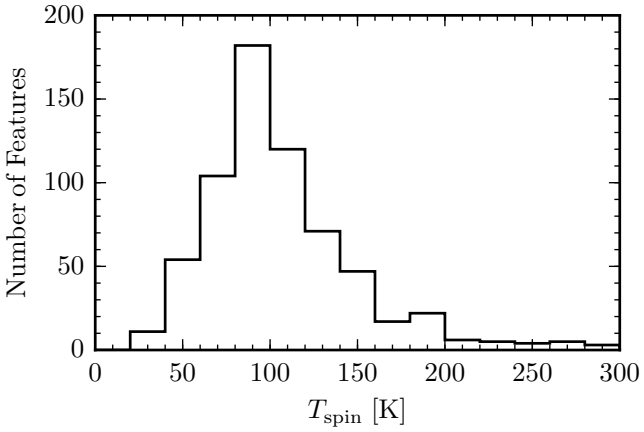

Fig. 8. Distribution of the spin temperature for all absorption features (655) toward point-like continuum sources. The median of the spin temperature is 97.5 K.

location of the continuum source. We can distinguish between extragalactic and Galactic point sources. The Galactic sources are mostly H II regions, whereas the extragalactic point sources are radio lobes from extragalactic jets or AGN. To distinguish between them, we have two criteria: the spectral index and H I absorption of the far Scutum-Centaurus arm. As we have to consider, e.g., primary beam effects or different spatial filtering of the interferometer for different frequencies, it is very difficult to determine accurate spectral indexes (e.g., Rau et al. 2014; Bhatnagar et al. 2011, 2013). For this analysis, we calculated the spectral indices, using flux measurements in the two most separated spectral windows at 1.05 and 1.95 GHz and are based on work presented in Johnston et al. (in prep.).

Four sources in our sample have negative spectral indices (Table 1), which is typical of synchrotron radiation from extragalactic jets or AGNs. The remaining four sources show flat or positive spectral indices, which indicates free-free emission from potentially optically thick H II regions. These four free-free emission sources also show no H I absorption feature for the far Scutum-Centaurus arm. Therefore, these sources reside in the Milky Way. Furthermore, we can study the H I absorption spectra and molecular emission spectra of the Galactic point sources to estimate their distance and overcome the near-far distance ambiguity. This is done by Anderson et al. (2014) and they find that, e.g., the source G31.411+0.307 has a near distance of 6.6 kpc. For this source, we also see a sharp cutoff in the absorption spectra at $v_{\rm lsr} \approx 100$ km s^{-1} (see Fig. 7), which corresponds to the molecular line velocity reported in Anderson et al. (2014).

Figure 6 shows an example of an extragalactic continuum source. The characteristic absorption of the far Scutum-Centaurus arm at $v_{\rm lsr} \approx -40$ km s^{-1} is clearly observable. Furthermore, the absorption and emission spectra show a similar cutoff for high velocities, when approaching the tangential point of the near Scutum-Centaurus arm. On the other hand, the absorption spectrum of Fig. 7 neither shows the absorption of the far Scutum-Centaurus arm, nor approaches zero at the velocity

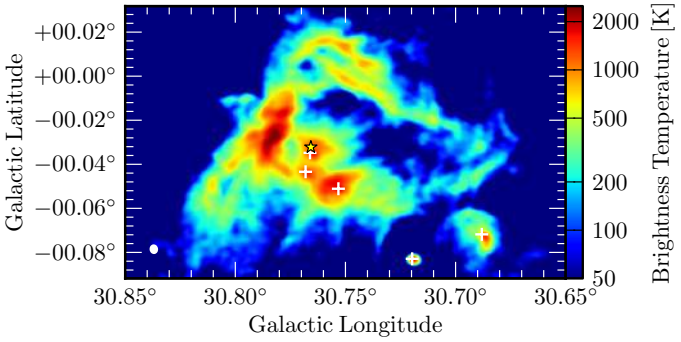

Fig. 9. Continuum emission at 1.4 GHz of W43-Main. The yellow star indicates the position of the OB cluster (Lester et al. 1985) and the white crosses marks the position of UCH II regions observed by the CORNISH survey (Hoare et al. 2012; Purcell et al. 2013).

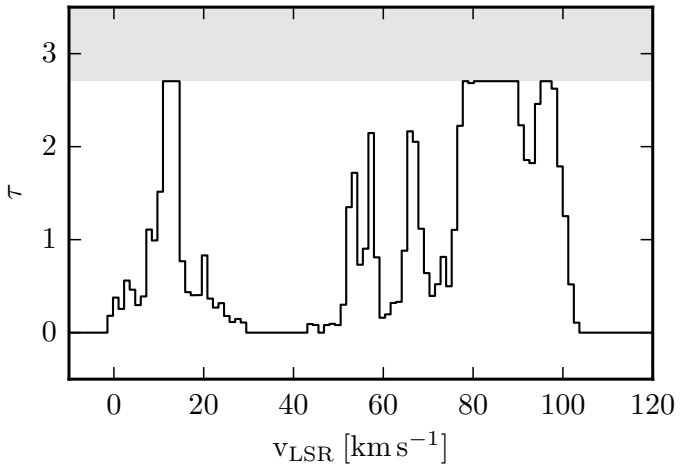

Fig. 10. Optical depth of W43-Main, calculated using Eq. (6). The gray shaded area indicates the maximum measurable optical depth of $\tau_{\mathrm{lower-limit}} = 2.7$ calculated using Eq. (7).

of the tangential point. This is typical for Galactic continuum sources.

4.4. Extended continuum sources – W43-Main

Besides point-like continuum sources, our observed field contains three strong ($T > 800$ K) extended continuum sources. The most prominent source is the well-known H II region around W43-Main (Lester et al. 1985), which is shown in Fig. 9. This region has an angular extent of ∼300″, which corresponds to ∼8 pc at a distance of 5.5 kpc. As our resolution is 20″, we are able to resolve the internal structure well. It is known that the nebula is ionized by an OB cluster and the observed continuum signal is the result of free-free emission. Figure 9 shows the continuum emission of W43-Main and the yellow star marks the position of the OB cluster (J2000, 18h47m36s, −1°56′33″, Lester et al. 1985). Furthermore, several UCH II regions can be found in the CORNISH survey, which are marked with white crosses (Hoare et al. 2012; Purcell et al. 2013). The maximum brightness temperature for this region is $T_{\mathrm{cont}} \sim 2200$ K and, therefore, we are able to calculate the optical depth, even though we cannot determine a proper off position (see Sect. 3.3 for details). The H I absorption spectra shows a cutoff at $v_{\mathrm{lsr}} \sim 100$ km s^{-1} (see Fig. 10), which marks the velocity of the continuum source. We measure radio recombination line emission at the same velocity (Johnston et al., in prep.; Anderson et al. 2011). Therefore, the continuum source W43-Main is situated at $v_{\mathrm{lsr}} \approx 100$ km s^{-1}, which is close to the tangent point velocity of the Scutum-Centaurus arm (Nguyen Luong et al. 2011).

Furthermore, the observed field contains two other extended continuum sources: W43-South and the supernova remnant SNR G029.7-00.2 with continuum brightness temperatures of at most $T_{\mathrm{cont}} \approx 1840$ K and $T_{\mathrm{cont}} \approx 850$ K, respectively.

4.5. Optical depth of W43-Main

In Sect. 3.3, we described that we can determine the optical depth for strong continuum sources using the absorption spectrum. As the brightness of the continuum source W43-Main is $T_{\mathrm{cont}} \sim 2200$ K, the uncertainty for the optical depth measurement is ∼10% (see Fig. 3). Figure 10 shows the optical depth of W43-Main and the gray shaded area indicates the lower limit of our measurement with $\tau_{\mathrm{lower-limit}} = 2.7$. The optical depth peak at $v_{\mathrm{lsr}} \approx 10$ km s^{-1} can be allocated to the near Sagittarius arm and therefore is not connected to the actual star-forming region W43. In contrast, it is difficult to allocate the distinct absorption features between 50 and 80 km s^{-1} and it is not clear whether they are spatially connected to W43. The prominent absorption feature of W43 is situated between 80 and 100 km s^{-1}. In this region, our measurement is saturated and, therefore, we can only report lower limits for the optical depth.

As we resolve the strong continuum source W43-Main spatially, we can determine the optical depth along different lines of sight and thus investigate the spatial distribution of absorption features as done in detail by Liszt et al. (1993). We refrain from this kind of study as we are mostly interested in the velocity range of W43 ($v_{\mathrm{lsr}} = 80$–110 km s^{-1}) at which the absorption features are mostly saturated preventing a detailed study of the spatial distribution. We instead measured the optical depth toward the strongest continuum peak to estimate the maximum optical depth possible (see Fig. 10), which is nevertheless a lower limit.

4.6. H I column density with optically thin assumption

In this section, we assess the H I column density and the H I mass of the GMC associated with W43. These two quantities depend strongly on the chosen velocity range. Nguyen Luong et al. (2011) defined for the "main" and "complete" velocity range of W43 values of v_{lsr}(main) = 80–110 km s^{-1} and v_{lsr}(complete) = 60–120 km s^{-1}. In the following, we use the "complete" velocity range, but we stress that a different velocity range can significantly change the column density and mass.

As described in Sect. 3.2, we can determine the column density by assuming optically thin H I emission. For this method, the column density is proportional to the observed H I brightness temperature. The column density map assuming optically thin emission is shown in the top right panel of Fig. 11. To estimate the mass, we integrate this column density over the main region of W43 ($l = 29.0$–$31.5°$, $|b| \leq 1°$). We mask all regions, where the emission spectrum has negative values, and exclude them. Given a distance of 5.5 kpc (Zhang et al. 2014), our analysis finds an H I mass of $M = 2.7 \times 10^6\, M_{\odot}$ ($l = 29.0$–$31.5°$, $|b| \leq 1°$, $v_{\mathrm{lsr}} = 60$–120 km s^{-1}).

Weak diffuse continuum emission can influence the H I emission spectrum and therefore the H I column density calculation needs to be modified according to Eq. (8) (Sect. 3.4). The top right panel of Fig. 11 shows this effect clearly. The color

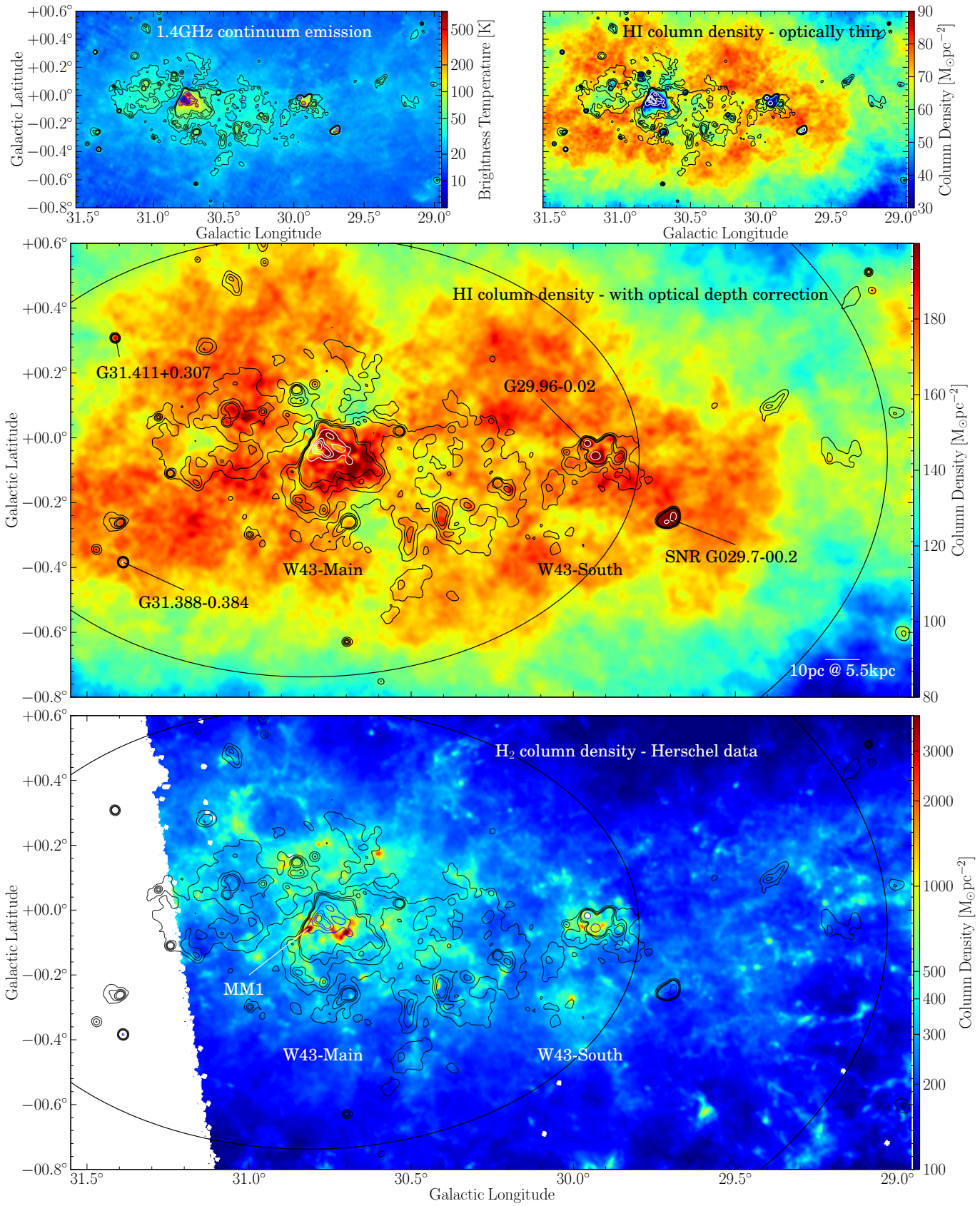

Fig. 11. *Top left panel*: continuum emission at 21 cm in Kelvin. *Top right and middle panels*: H I column density for optically thin assumption and optical depth corrections, respectively. The continuum and H I emission data are based on VGPS, whereas the optical depth correction used in the *middle panel* is based on the THOR data. The *bottom panel* shows the H_2 column density from the HiGAL data (Nguyen-Lu'o'ng et al. 2013). In all panels, the black and white/blue contours present the continuum emission at 21 cm (black contours show levels of 10, 30, and 70 K; white/blue contours show levels of 200, 400, 600, and 800 K). The black ellipses in the *middle and lower panels* have an equivalent radius of $r = 80$ and 140 pc. Several important objects are marked.

represents the column density determined by the optically thin assumption and the black contours indicate the weak diffuse continuum emission. These two components show a clear anti-correlation. However, this anti-correlation is the result of the expected observational effect in which H I emission is suppressed by weak continuum emission (see Sect. 3.4). To overcome this problem, we have to consider the optical depth, which we discuss in the next section.

4.7. H I column density with optical depth correction

With the measurement of the optical depth and the weak continuum emission, we can correct the H I emission as explained in Sect. 3.4. This allows for a more accurate determination of the column density, which is shown in the middle panel of Fig. 11. We chose the "complete" velocity range of $v_{\rm lsr} = 60{-}120\,{\rm km\,s^{-1}}$. As we correct for the optical depth, we are able to observe a larger H I column density. We also correct for the weak continuum emission around W43-Main. This correction removes the anti-correlation between the continuum emission and the H I column density. Hence, we determine the H I mass to be $M = 6.1 \times 10^6\,M_\odot$ ($l = 29.0{-}31.5°$, $|b| \leq 1°$, $v_{\rm lsr} = 60{-}104\,{\rm km\,s^{-1}}$).

The optical depth spectrum of W43-Main (Fig. 10) shows, that the absorption ends abruptly at 100 km s^{-1}, but the H I emission as well as CO emission (Nguyen Luong et al. 2011; Carlhoff et al. 2013) reveals features up to $v_{\rm lsr} = 110\,{\rm km\,s^{-1}}$. The reason for the abrupt drop in the absorption spectrum is not the absence of H I, but the location along the line of sight of the continuum source at this velocity (see Sect. 4.3 for further details). As we do not see H I in absorption for velocities larger than $100\,{\rm km\,s^{-1}}$, we are also not able to determine the corresponding optical depth. Therefore we cannot apply our corrections to the H I column density for velocities larger than $100\,{\rm km\,s^{-1}}$. Hence, the velocity range for the previously mentioned H I mass is only $v_{\rm lsr} = 60{-}104\,{\rm km\,s^{-1}}$. Nevertheless we can determine the H I mass for the velocity range from $v_{\rm lsr} = 104{-}120\,{\rm km\,s^{-1}}$ using the optically thin assumption. This reveals a H I mass of $M = 0.5 \times 10^6\,M_\odot$. Hence, the total H I mass for W43 in the velocity range $v_{\rm lsr} = 60{-}120\,{\rm km\,s^{-1}}$ is $M = 6.6 \times 10^6\,M_\odot$ with the optical depth correction for the velocity range $v_{\rm lsr} = 60{-}104\,{\rm km\,s^{-1}}$ and optically thin assumption for the velocity range $v_{\rm lsr} = 104{-}120\,{\rm km\,s^{-1}}$. This is 2.4 times larger than the H I mass determined with the optically thin assumption.

The limitations and uncertainties of our determined H I column density and H I mass with the optical depth corrections are discussed further in Sect. 5.3.

4.8. H I self absorption

The H I self absorption (HISA) method uses the diffuse broad H I emission background of the Milky Way as illumination for colder foreground clouds (e.g., Gibson et al. 2000, 2005a,b; Li & Goldsmith 2003; McClure-Griffiths et al. 2006). Dark regions on maps and narrow absorption features in spectra reveal these HISA features. The terminology "self absorption" can be misleading: the H I emission and absorption can occur in the same cloud or at the same position, but does not have to. The advantage of this method is that it is possible to map entire absorption clouds and study their intrinsic structure. In contrast, the disadvantage is that a sufficient background emission with the same velocity as the absorbing cloud is necessary to detect HISA features. Therefore this method misses a large portion of the cold H I clouds. The differentiation between actual HISA features and the lack of H I emission can also be challenging. Gibson et al. (2005b), however, present an efficient method to detect HISA features automatically. Another disadvantage is that the optical depth and spin temperature can only be measured together and further assumptions are needed to disentangle these two values.

As described in the previous section, we correct for the optical depth effects and weak continuum emission. This correction does not account for locally confined HISA features, as we assume a uniform optical depth for the entire W43 region. The HISA features could have a higher and spatially varying optical depth that we cannot measure. Furthermore, the weak and diffuse continuum emission around W43 makes the search for HISA features even more complicated. Hence, we refrain from searching and analyzing the possible HISA features around W43. Therefore, we are likely missing some cold neutral hydrogen in our analysis. We will conduct future HISA studies in other parts of the Milky Way THOR study with less diffuse continuum emission.

5. Discussion

5.1. Phases of the neutral atomic hydrogen

It is well known that the neutral atomic ISM has several phases that coexist side by side with very different properties (e.g., Clark 1965; Wolfire et al. 1995, 2003; Heiles & Troland 2003b). The main constituents are the cold neutral medium (CNM) and the warm neutral medium (WNM) with spin temperatures on the order of <100 K (Strasser et al. 2007) and ~10^4 K (Murray et al. 2014; Roy et al. 2013a), respectively. Furthermore, their density differs by two order of magnitude (CNM: $n_{\rm H} \sim 50$ cm^{-3}, WNM: $n_{\rm H} \sim 0.5$ cm^{-3}, e.g., Stahler et al. 2005). Because of the different spin temperature and density of the CNM and WNM, their corresponding optical depths are significantly different with typical values of $\tau_{\rm WNM} \sim 10^{-3}{-}10^{-4}$ (Murray et al. 2014) and $\tau_{\rm CNM} \gtrsim 0.1$ (Strasser & Taylor 2004). This is important for our interpretation.

Because of the low optical depth of the WNM, we see in absorption merely the CNM. Hence, the optical depth shown in Fig. 10 is the optical depth of the CNM. As the absorption spectra are strongly dominated by the CNM, we are not able to measure the optical depth of the WNM and we assume the WNM is optically thin. In contrast, we see a combination of the CNM and WNM for the H I emission. For the correction of the column density (see Sect. 4.7), we use the optical depth information from the absorption study to correct the H I emission data. As we do not distinguish between the two phases, we might combine two different quantities, namely the optical depth of the CNM with the emission of the CNM and the WNM. This could lead to an overestimation of the column density. In the following, we assess this effect.

If we assume a CNM cloud with varying $T_{\rm spin}$(CNM) ~ 20–80 K and varying optical depth surrounded by a WNM with $T_{\rm spin}$(WNM) ~ 7000 K (Murray et al. 2014) and optical depth τ(WNM) ~ 5×10^{-3}, we can calculate the column density of each component separately (CNM and WNM). Furthermore, we can calculate the brightness temperature that we would observe and apply our correction method described in Sect. 3.4. Finally, we can compare the actual column density of the CNM and WNM with the column density we would observe with our correction method and investigate the overestimation of the actual column density. The result is shown in Fig. 12, which shows the ratio of the observed and actual column density as a function of the optical depth of the CNM τ(CNM) for different

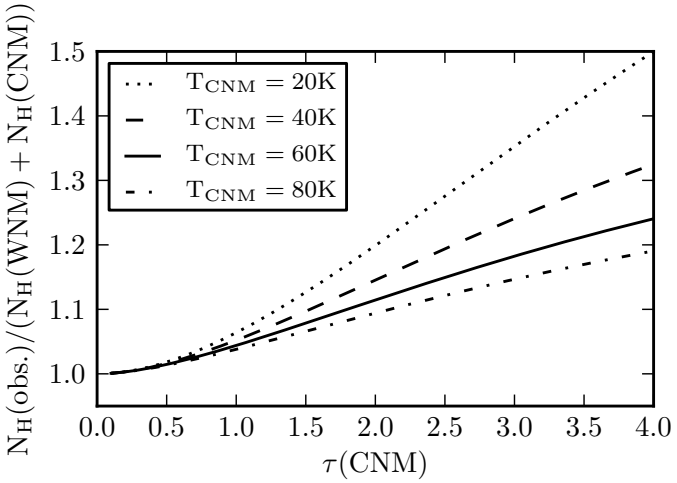

Fig. 12. Ratio of the observed and actual column density as a function of the CNM optical depth τ(CNM) for different CNM spin temperatures T_{spin}(CNM).

CNM spin temperatures T_{spin}(CNM). Figure 12 shows that we measure the column density accurately for the optically thin case and we overestimate the column density for larger optical depths. However, even for the extreme case (T_{spin}(CNM) ∼ 20 K, τ(CNM) ∼ 3), we overestimate the column density by at most 1.35. This effect is smaller than the underestimation of the column density due to saturated optical depths and therefore we only consider a single cold component in the following. Similar results in simple simulations for different combination of CNM and WNM fraction and a wide range of N_H and τ were found by Chengalur et al. (2013) and Roy et al. (2013a,b).

5.2. H I spin temperature measurements toward point sources

The spin temperature is the best measure of the kinetic temperature of the H I gas along the line of sight. Hence, it is a principal determinant of the physical processes that occur in the H I gas.

As described in Sect. 3.5, we corrected the H I spin temperature for diffuse weak continuum emission. If we neglect this correction we would systematically overestimate the H I spin temperature and our sample would have a median H I spin temperature of 110 K. That means the median value would be ∼12 K higher compared to our corrected value of 97.5 K (see Fig. 8).

Similar studies of the spin temperature with similar methods can be found in the literature. For example, Strasser & Taylor (2004) report a median spin temperature of 120 K. This is ∼20 K higher than our value of 97.5 K, but they do not consider the diffuse weak continuum emission and therefore probably overestimate the spin temperature.

Heiles & Troland (2003a,b) developed an extensive method to fit Gaussian components to the absorption and emission spectra. Spectral features shown in absorption and emission are assigned to the cold neutral medium (CNM), whereas emission-only features are assigned to the warm neutral medium (WNM). Using this method, they are able to distinguish these two phases and measure their properties, such as as the spin temperature, individually. However, they report that it is difficult to use their method for sources close to the Galactic plane ($|b| < 10$ deg), as multiple components can overlap and the Galactic rotation can broaden their profiles. Figures 6 and 7 illustrate this problem for our region and for these spectra it is impossible to find unique Gaussian components. As Heiles & Troland (2003a,b) are able to fit individual components, they find on average lower spin temperatures for the CNM in the range of ∼40−70 K. Given that we can not distinguish between the CNM and the WNM, the measured spin temperature is a column density weighted mean between these two components (Strasser & Taylor 2004). This could explain the increased spin temperature reported in Sect. 4.2. Another explanation for the increased spin temperature could be the strong radiation field of W43.

Further complications in our sample are the Galactic continuum sources. As these only trace the optical depth up to their location within the Milky Way, we miss the optical depth of the hydrogen behind the continuum source. On the other hand, the off-position measurement traces all the hydrogen along the line of sight. As the kinematic distance is uncertain due to near and far ambiguities, we cannot use the velocity to distinguish the distance of the emitting and absorbing hydrogen. Hence we underestimate the optical depth for the Galactic continuum sources and therefore overestimate the spin temperature. However, we do not see a significant difference in the mean spin temperature for Galactic and extragalactic continuum sources and we can neglect this effect.

5.3. Column density and mass estimate

In Sect. 4.7 we explained that we correct the H I column density for optical depth effects as well as for the diffuse continuum emission. This leads to a more accurate estimate of the column density and a more accurate mass estimate. Nevertheless these corrections have limitations which we discuss in the following section.

As we need a strong continuum source in the background to determine the optical depth, we can measure the optical depth only toward certain locations. For our H I column density correction, we used the strong continuum source W43-Main. Furthermore we assumed that the optical depth is the same for the entire cloud. This assumption might not hold, especially for the outer parts of the cloud. To investigate this effect, we compared the optical depth measurements for other sources at other positions. However, in the range of v_{lsr} = 80−110 km s^{-1}, we find that the optical depth is mostly saturated and determined by the corresponding lower limit. Two examples are given in the third panel of Figs. 6 and 7, which reveal a lower limit of $\tau_{lower-limit}$ = 2.9 and $\tau_{lower-limit}$ = 1.9, respectively. Other continuum sources that have a larger separation from the Galactic mid-plane show similar results. For example, the continuum source G30.699−0.630 has a Galactic latitude of $b \approx 0.6°$ and still shows a saturated optical depth for the velocity range of W43 with a lower limit of $\tau_{lower-limit}$ = 1.5. If we use this continuum source to correct for optical depth effects and the weak diffuse continuum emission, we determine a mass of $M = 4.8 \times 10^6\ M_\odot$ for the whole cloud (l = 29.0−31.5°, $|b| \leq 1°$, v_{lsr} = 60−120 km s^{-1}). Hence, by using W43-Main to correct the optical depth for the entire cloud, we do not overestimate the mass in the outer parts significantly, but at most by a factor of 1.4. We also use this mass estimate as a lower limit for the mass of W43.

The saturation of the optical depth is the second limitation we have to consider, especially for the inner part of the cloud. In this region, we underestimate the opacity and, therefore, the column density which would lead to a further underestimation of the mass. Estimating this effect is difficult and, therefore, we report only lower limits for the column density and the mass in the inner part of W43.

The third limitation is the distance of the continuum source that we use to determine the optical depth. As explained in Sect. 4.3, the continuum source W43-Main is located close to the tangential point of the Scutum-Centaurus arm at a distance of 5.5 kpc. Hence, we only see H I absorption as far as this distance, and miss all the H I that is located behind the continuum source but still within the cloud. If we assume that the continuum source is at the center of the cloud, we underestimate the column density by a factor of two. Another approach to overcome this limitation is to look at other more distant continuum sources. The continuum source G31.388−0.384 is extragalactic and has a comparable brightness temperature to W43-Main. Hence, it is an ideal candidate for this test. We used the optical depth shown in the third panel of Fig. 6 to correct the H I column density and measured the corrected mass for the same area and velocity range, as in Sect. 4.7. The absoprtion spectrum of W43-Main and G31.388−0.384 are similar, except for the velocity range $v_{\rm lsr} = 100-120\,{\rm km\,s^{-1}}$, which is missing in the W43-Main spectrum. We can also use the optical depth measurements of G31.388−0.384 to correct the H I emission. However, as these absorption spectra are similar, the corrected masses are the same within the uncertainties ($M = 6.9 \times 10^6\,M_\odot$ for the correction with the optical depth of G31.388−0.384).

Summing up, we might overestimate the mass in the outer part of W43, but underestimate the mass in the inner part of W43. Because of the position of the continuum source, we might underestimate the mass by a factor of two. As explained in Sect. 4.8, we also miss some H I due to self-absorption. Hence we report the H I mass of W43 to be a lower limit of $M = 6.6_{-1.8} \times 10^6\,M_\odot$.

Several H I mass estimates are given in the literature (Nguyen Luong et al. 2011; Motte et al. 2014). All of them are calculated with the assumption of optically thin emission.

Motte et al. (2014) measured the H I mass in the inner part of W43 ($l = 29.^\circ6 \sim 31.^\circ4$ and $b = -0.^\circ5 \sim 0.^\circ3$) and for the velocity range ($v_{\rm lsr} = 60-120\,{\rm km\,s^{-1}}$). The assumption of optically thin emission reveals an H I mass of $M = 0.9 \times 10^6\,M_\odot$. If we use our corrected H I column density to measure the mass in this area, we determine a H I mass of $M = 2.2 \times 10^6\,M_\odot$ (this value is smaller than the value given in Sect. 4.7 as we only consider the inner part of W43). As outlined above, we claim that this mass estimate is a lower limit and, therefore, the mass determined with the optically thin emission is at least a factor of 2.4 too small. This has implications for theoretical models, which we discuss in Sect. 5.6.

5.4. Spatial distribution of hydrogen

Beside the total H I mass, knowing the spatial distribution of the H I is crucial to understand the formation of W43. As we have corrected the column density map for optical depth effects and the weak diffuse continuum emission, we can use this corrected column density map (see Fig. 11) to investigate the H I spatial distribution in more detail, especially in the center.

Similar work was done by Motte et al. (2014). They measured the H I column density in rectangular annuli around the center of W43 ($l = 30.5^\circ$, $b = 0^\circ$) with an aspect ratio of 3:2 and find an increasing H I column density inward from $N_{\rm H\,I} \sim 40\,M_\odot\,{\rm pc^{-2}}$ at a distance of 170 pc to $N_{\rm H\,I} \sim 80\,M_\odot\,{\rm pc^{-2}}$ at a distance of 50 pc from the center (velocity range $v_{\rm lsr} = 60-120\,{\rm km\,s^{-1}}$). Since they assume optically thin H I emission, they underestimate the H I column density, especially in the central region (as discussed in Sect. 5.3).

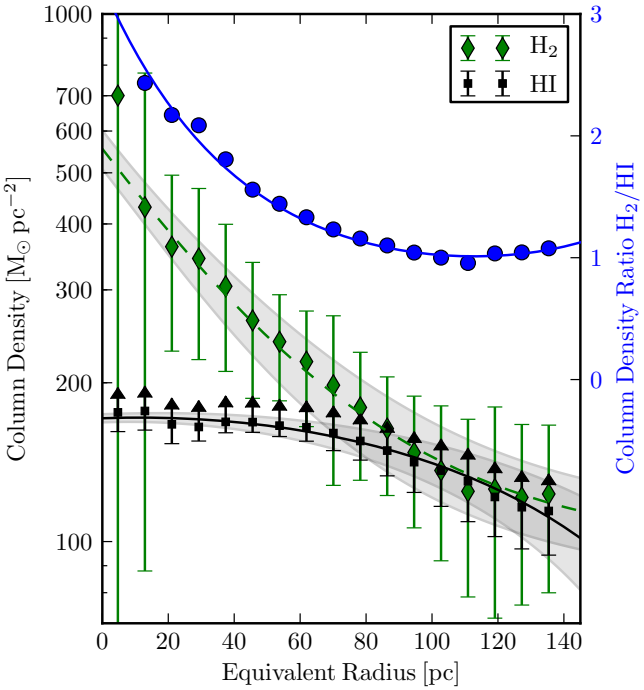

Fig. 13. Column density of H I and H_2 measured in elliptical annuli around W43 MM1. The x-axis presents the equivalent radius of these elliptical annuli. The black and green lines represent the fitted curves with the corresponding 1σ uncertainties shown as a gray shaded area. The larger symbols (diamonds for H_2, squares for H I) present the averaged value of the elliptical annuli and their corresponding variations within the annuli. As the optical depth spectra saturates, we can only determine lower limits for the H I column density. The blue dots and line show the H_2 and H I column density ratio of the data and the fitted curves, respectively.

To compensate for the approximate elliptical structure of W43, we choose elliptical annuli with an aspect ratio of 3:2 for major and minor axis, which fits the large scale structure of W43 well (see Fig. 11). As we focus on W43-Main, we choose the most massive submillimeter peak W43-MM1 ($l = 30.8175^\circ$ and $b = -0.0571^\circ$, Motte et al. 2003) as the center for the ellipses. The width of each annulus is 10 pc for the major axis and 6.6 pc for the minor axis. For each elliptical annulus, we calculated the equivalent radius $r = \sqrt{\rm major \cdot minor}$ and assigned these values as the distance to the center shown in Fig. 13.

The black squares in Fig. 13 represent the H I column density mean value and the corresponding standard deviation of each elliptical annulus. We confirm the result of Motte et al. (2014), which indicates that the H I column density rises inward. Our corrections allow us for the first time to study the central region of W43 ($r < 50$ pc) and, within the uncertainties, we report a flat column density distribution in this inner region. However, we mention that this flatness could also be due to the underestimation of the column density in the central part, as the optical depth saturates and therefore the column density is a lower limit.

Furthermore, the diamond symbols in Fig. 13 present the column density of the molecular hydrogen. The H_2 distribution is centrally concentrated and the column density rises steeply toward the center, which is different from the H I profile. The large uncertainties in the center are due to the clumpiness of the molecular hydrogen as the most prominent clumps, such as W43-MM1, are located in the first two elliptical annuli. Beside W43-Main, the second most prominent molecular clump

is W43-South (see Fig. 11). However, in this analysis we focus on W43-Main and, therefore, we choose W43-MM1 as the center for the ellipses. Furthermore, we masked W43-South for this analysis, as it would introduce a large uncertainty at an equivalent radius of $r \sim 50$ pc.

We use a nonlinear least square method ("curvefit" in the scipy package) to fit the mean values of the thin elliptical annuli with their corresponding uncertainties. For the H I column density we assume a quadratic radial distribution, i.e.,

$$N_{\mathrm{H\,I}}(r) = a \cdot r^2 + b \cdot r + c, \tag{11}$$

where $N_{\mathrm{H\,I}}(r)$ describes the H I column density and r describes the equivalent radius. The free parameters a, b and c have the fitted values of $a = -3.9 \pm 0.7 \times 10^{-3}\ M_\odot\,\mathrm{pc}^{-4}$, $b = 0.089 \pm 0.099\ M_\odot\,\mathrm{pc}^{-3}$ and $c = 171 \pm 3\ M_\odot\,\mathrm{pc}^{-2}$. For the H_2 distribution we assume an exponential function to fit the data as follows:

$$N_{\mathrm{H_2}}(r) = d \cdot \exp(-e \cdot r) + f, \tag{12}$$

where $N_{\mathrm{H_2}}(r)$ describes the H_2 column density and r describes the equivalent radius. The free parameters d, e, and f have the fitted values of $d = 458 \pm 44\ M_\odot\,\mathrm{pc}^{-2}$, $e = 0.022 \pm 0.004\,\mathrm{pc}^{-1}$ and $f = 97 \pm 15\ M_\odot\,\mathrm{pc}^{-2}$.

The black and green lines in Fig. 13 present the column density fits and the gray shaded areas represent their corresponding uncertainties. Both, the H I and H_2 distributions are well fit by the assumed functions. As the uncertainties for the innermost part of the H_2 distribution ($r < 20$ pc) are large, the fitted curve deviates from the data points and this area has to be treated cautiously. We also tested different functions to fit the data, such as functions with only two free parameters, or a Gaussian distribution for the H I distribution, but the results were similar within the uncertainties. Finally, we chose the aforementioned functions as they could reproduce well the ratio of the H_2 and H I column density. This ratio is shown in blue in Fig. 13. The circles represent the ratio of the data points, whereas the solid line represents the ratio of the fitted curves. The ratio stays fairly constant at $N_{\mathrm{H_2}}/N_{\mathrm{H\,I}} \approx 1$ for $140\,\mathrm{pc} > r > 60$ pc. For this region, the H I column density also rises to its maximum value of $N_{\mathrm{H\,I}} \approx 170\ M_\odot\,\mathrm{pc}^{-2}$. Further inward ($r < 60$ pc), the H I column density stays constant at this maximum value, whereas the H_2 column density rises sharply. Hence, the H_2/H I ratio also rises sharply to values above three. Summing up, the column density measurements imply that we have a mixture of H I and H_2 in the outskirt of the cloud and a molecular dominated region in the center. In the following, we investigate this structure for the particle density.

As mentioned before, we find a flat column density for the H I distribution toward the center of W43-Main, but what does this imply for the actual density in the center? If we assume that W43 has an elliptical shape, we can decompose the cloud into different layers, similar to an onion. Furthermore, we assign the appropriate column density to each layer with the information given in Fig. 13. As the column density is additive, the appropriate column density of each layer is the measured column density at the position of the considered layer minus the column density of all layers further outside. Hence, a flat H I column density distribution toward the center (see Fig. 13) means that the layers in the center have no or a very small column density and therefore also a very small particle density. In the following section, we use the concept of a elliptical layered structure to determine the actual particle density and show that indeed the H I particle density drops toward the center of W43 within this model.

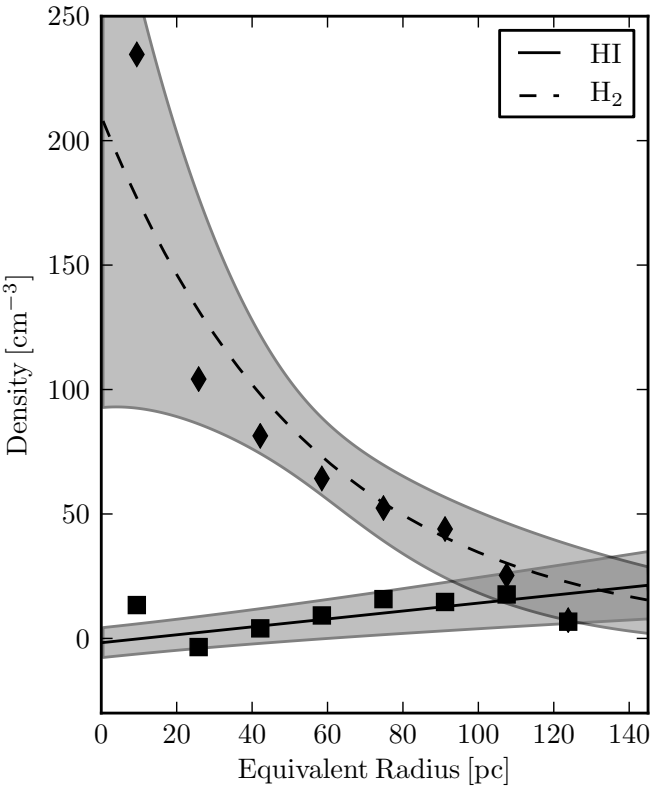

Fig. 14. Density of H I and H_2 as a function of the distance toward the center of W43 (equivalent radius). The diamonds and the dashed line represent the molecular hydrogen, whereas squares and the solid line represent the atomic hydrogen. The corresponding 3σ uncertainties from the fit introduced in Eq. (14) are shown as gray shaded areas.

5.5. Linking column density to particle density

While the column density neglects the third dimension, it does not necessarily reflect the actual particle density. Hence, we have to estimate the third dimension. Using the elliptical, layered structure for W43 previously described and the results presented in Fig. 13, we estimate the particle density in each layer, i.e.,

$$n_{\mathrm{i}} = \frac{(N_i - N_{i+1})}{2\ (r_{i+1} - r_i)}, \tag{13}$$

where n_i describes the particle density of layer i; N_i and N_{i+1} describes the column density of layer i and layer $i+1$, respectively; and r_i and r_{i+1} describes the equivalent radius of layer i and $i+1$. The factor of two accounts for the two layers of the elliptical annuli in the front and back of the cloud. The result for this calculation is shown in Fig. 14 as squares and diamond data points for the H I and H_2 density, respectively. To increase the signal to noise ratio, we used larger elliptical annuli with a major axis of 20 pc and a minor axis of 13.3 pc.

In addition, we can use the fitted curves of the column density to estimate the density as well, by inserting Eqs. (11) and (12) into Eq. (13). Hence, the radial distribution of the H I and H_2 densities are

$$\begin{aligned} n_{\mathrm{H\,I}} &= -0.5\,(a\ (r_i + r_{i+1}) - b), \\ n_{\mathrm{H_2}} &= \frac{d}{2\ (r_{i+1} - r_i)}\ (\exp(-e \cdot r_i) - \exp(-e \cdot r_{i+1}))\,, \end{aligned} \tag{14}$$

where $n_{\mathrm{H\,I}}$ and $n_{\mathrm{H_2}}$ describes the particle density of H I and H_2, respectively; r_i describes the equivalent radius of layer i; and a,

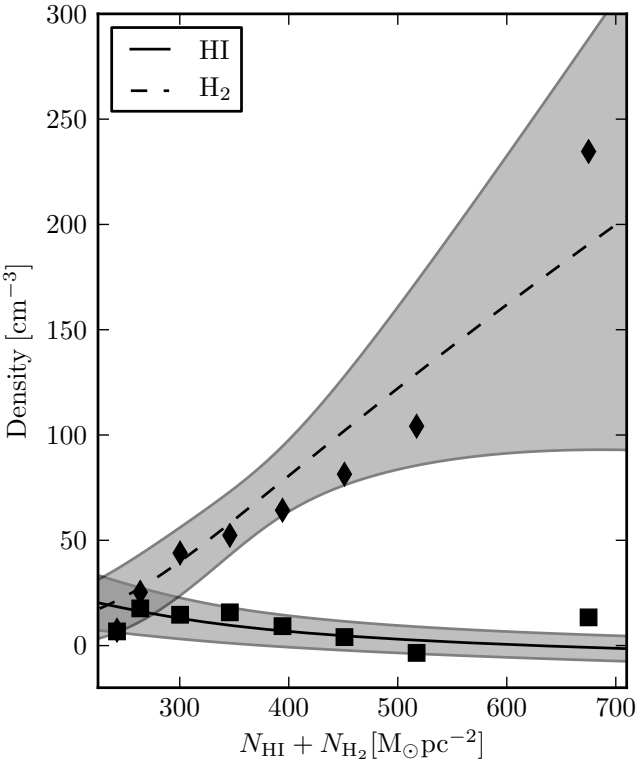

Fig. 15. Particle density of H I and H_2 as a function of the total column density ($N_{HI} + N_{H_2}$). The diamonds and the dashed line represents the molecular hydrogen, whereas squares and the solid line represents the atomic hydrogen. The corresponding 3σ uncertainties from the fit introduced in Eqs. (11) and (12) are shown as gray shaded areas.

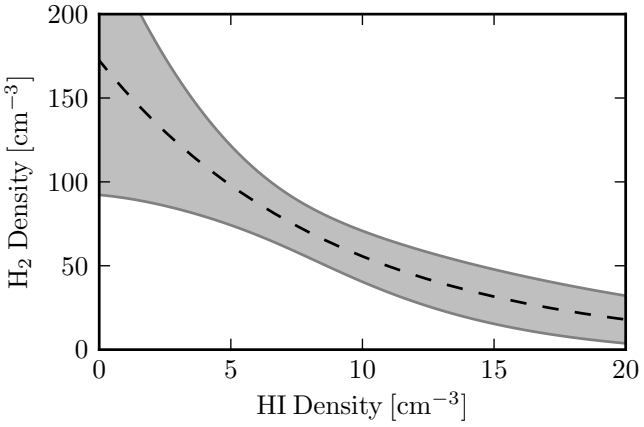

Fig. 16. Particle density of molecular hydrogen as a function of the atomic hydrogen. The black curve represents the best-fit curves from Fig. 14 and the gray shaded area shows the 3σ uncertainty for the H_2 fit.

b, c, and d are the free fit parameters introduced in Eqs. (11) and (12). In Fig. 14, the calculated density distributions are shown as black lines with the corresponding uncertainties as a gray shaded areas. We note that the density distributions of H I and H_2 are very distinct. While the H I distribution follows a simple linear relation, the H_2 distribution shows an exponential increase toward the center. The particle density shows a mixture of the atomic and molecular hydrogen in the outskirts of W43 ($r > 100$ pc) and a molecular dominated interior similar to the measurements of the column density. However, the particle density of the atomic hydrogen drops to almost zero toward the center, which results in the observed constant column density.

Another way to present our results is shown in Fig. 15. This figure shows the particle density of H I and H_2 introduced in Eq. (14) as a function of the total column density ($N_{HI} + N_{H_2}$). High column density regions, i.e., the central region of W43, are dominated by molecular hydrogen. On the other hand, we find an equivalent mixture of atomic and molecular hydrogen for low column density regions, which represent the envelope of W43.

In addition, we study the correlation between the H I and H_2 density. To do so, we replace the equivalent radius in Fig. 14 and plot the H_2 density as a function of the H I density. Figure 16 shows the corresponding plot with the uncertainties for the H_2 density. The statistical and systematical uncertainties for this plot are relatively large, but nevertheless we see an anti-correlation between the H I and H_2 density. This can be explained with a simple model that H_2 forms out of H I in the innermost part of W43. However, we do not detect a sharp transition between H_2 and H I predicted by Krumholz et al. (2008, 2009). In the following section, we discuss possible implications.

5.6. Column density threshold for atomic hydrogen

A fundamental question of molecular cloud formation is how does molecular hydrogen form out of neutral atomic hydrogen and what are the corresponding conditions (Dobbs et al. 2014). It is well known that the density must be high enough to shield the molecular hydrogen from the interstellar radiation field, to avoid dissociation back to its atomic form (Hollenbach & Tielens 1997; McKee & Ostriker 2007; Krumholz et al. 2008, 2009; Mac Low & Glover 2012).

To describe this transition, Krumholz et al. (2008, 2009) suggest an analytic model, which is based on the assumption of a spherical cloud that is embedded in an isotropic external radiation field. They approximate that the transition between H I in the envelope and H_2 in the center occurs in an infinitely thin shell. An important result is that an H I column density of $\sim 10\ M_\odot\ \mathrm{pc}^{-2}$ is necessary to shield the molecular hydrogen from the interstellar radiation field. If the cloud reaches this critical H I column density the formation of molecular hydrogen is efficient enough that most of the hydrogen goes into molecular form and the column density of H I remains relatively constant at this level. Therefore we should not expect to observe H I column densities larger than $\sim 10\ M_\odot\ \mathrm{pc}^{-2}$, contradicting our results presented in Fig. 13. Furthermore, they show that this H I column density threshold is independent of the external radiation field, but has a weak dependence on the metallicity of the gas.

This model has three free parameters. First, the ratio of the measured CNM density n_{CNM} to the minimal CNM density n_{min}, i.e.,

$$n_{CNM} = \phi_{CNM} \cdot n_{min}. \quad (15)$$

The minimal CNM density is determined by the pressure balance with the warm neutral medium (WNM) and has a typical value of $n_{min} \approx 7\text{–}8\ \mathrm{cm}^{-3}$. As the range of pressure balance between the CNM and the WNM is limited, the maximum CNM density can be at most $\sim 10 \cdot n_{min}$. Hence, ϕ_{CNM} can vary between 1 and 10. Krumholz et al. (2009) assume $\phi_{CNM} \approx 3$ for their fiducial value. The second free parameter is the ratio of the CNM density n_{CNM} to the molecular density n_{mol}, i.e.,

$$n_{mol} = \phi_{mol} \cdot n_{CNM}. \quad (16)$$

Krumholz et al. (2009) suggests that this ratio is $\phi_{mol} \approx 10$ and should not vary significantly. The last free parameter is the

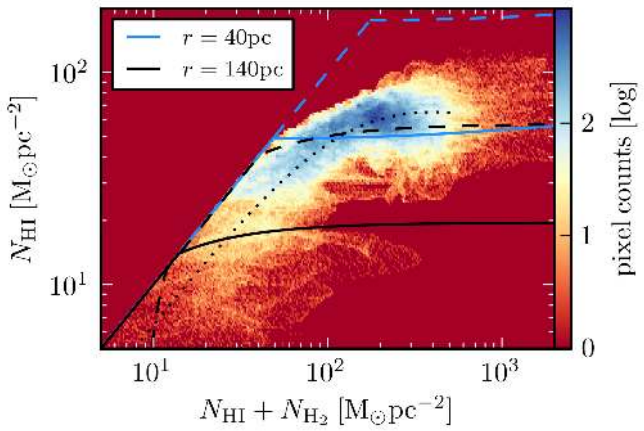

Fig. 17. Column density of the atomic hydrogen ($v_{\rm lsr}$ = 60−120 km s^{-1}) as a function of the total hydrogen column density ($N_{\rm HI} + N_{\rm H_2}$). The column density for the envelope ($r > 140$ pc, $N_{\rm HI} = 107 \pm 20\ M_\odot$ pc^{-2}, $N_{\rm H_2} = 116 \pm 18\ M_\odot$ pc^{-2}) is subtracted. The solid and dashed lines describe the theoretical model of Krumholz et al. (2008, 2009) for solar values ($G_0 = 1$, $Z = 1$) and for more realistic values of W43 ($G_0 = 5$, $Z = 1.4$), respectively. The blue and black colors represent the model parameters at different equivalent radii of $r = 40$ and 140 pc, respectively. The black dotted line shows the fitted curve of the elliptical annuli analysis introduced in Sect. 5.3.

metallicity Z. This model is supported by extragalactic observations (Krumholz et al. 2009) as well as observations in the Perseus molecular cloud by Lee et al. (2012). They find a uniform H I column density of $N_{\rm HI} \sim 6$−$8\ M_\odot$ pc^{-2} for H_2 column densities up to $N_{\rm H_2} \sim 80\ M_\odot$ pc^{-2}. To fit their data to the model of Krumholz et al. (2009), they fixed $\phi_{\rm mol} = 10$ and $Z = 1.0\,Z_\odot$ and fitted $\phi_{\rm CNM}$, which reveals values of $\phi_{\rm CNM} \approx 6$−10.

The complex W43 is a more extreme test case for this theory as we have a much larger mass reservoir as well as an OB cluster that acts as a strong radiation source in the center. For this analysis, we extracted the column density of H I and H_2 for each pixel. As a basis for the H I column density, we chose the optical depth corrected version with a velocity range of $v_{\rm lsr}$ = 60−120 km s^{-1}. For the H_2 column density, we again use the HiGal data (see Sect. 2.4). Even though the data base is the same as that used in Sect. 5.5, we stress that the method is different. Here we conduct a pixel by pixel comparison, whereas in Sect. 5.5 we averaged the column density over elliptical annuli. In addition, we focus merely on the inner part of W43 with $r < 140$ pc and, therefore, we have to subtract the column density of the envelope. We assess the column density at $r = 140$ pc as the envelope column density and subtract this value. The column density for the H I and H_2 envelopes are $N_{\rm HI} = 107 \pm 20\ M_\odot$ pc^{-2} = $1.3 \pm 0.3 \times 10^{22}$ cm^{-2} and $N_{\rm H_2} = 116 \pm 18\ M_\odot$ pc^{-2} = $7.3 \pm 1.1 \times 10^{21}$ cm^{-2}, respectively.

Figure 17 shows the result of the pixel by pixel comparison of the H I and H_2 column density. Since we subtracted the column density of the envelope for this analysis, the values are smaller than in Fig. 13. For better readability, we do not show each single pixel comparison, but we performed a pixel binning. We do not observe the predicted threshold for the H I column density of ∼10 $M_\odot$ pc^{-2}, instead our data show high H I column density values, peaking between $N_{\rm HI} = 50$−$80\ M_\odot$ pc^{-2}. Nevertheless we will try to fit the analytic model of Krumholz et al. (2009) to our data.

As described in Sect. 5.5, we use elliptical annuli to estimate the particle density. We can use this information (Fig. 14) to determine the model parameters $\phi_{\rm CNM}$ and $\phi_{\rm mol}$ using Eqs. (15) and (17), respectively. Given that the H I and H_2 densities vary as a function of equivalent radius, the model parameters $\phi_{\rm CNM}$ and $\phi_{\rm mol}$ also vary. We calculate the model parameters for three different distances r = 40, 80, and 140 pc, which characterize the molecular dominated interior, the transition region, and the well-mixed outer area. To calculate $\phi_{\rm CNM}$, we have to know $n_{\rm min}$, which is given by Krumholz et al. (2009) as

$$n_{\rm min} \approx 31 \frac{G_0}{1 + 3.1\,Z^{0.365}}\ {\rm cm}^{-3}, \tag{17}$$

where G_0 and Z are the far-ultraviolet radiation intensity (in units of the Habing field) and the total metallicity, both normalized to their values in the solar neighborhood. For W43, we have a slightly higher metallicity of $Z = 1.4\ Z_\odot$ (Motte et al. 2014) and a large radiation field, which could be up to $>100\,G_0$ (Beuther et al. 2014) in the central region. We assume a more moderate radiation field for the outer regions of W43 with values around 5−$10\,G_0$. In Table 2, we present the H I and H_2 density as well as the model parameters for different distances. The parameter $\phi_{\rm CNM}$ is calculated for solar values ($G_0 = 1$, $Z = 1$) and for more realistic values of W43 ($G_0 = 5$, $Z = 1.4$) that results in $n_{\rm min} = 7.5$ cm^{-3} and $n_{\rm min} = 34.4$ cm^{-3}, respectively. The solid blue and black lines in Fig. 17 show the theoretical column density for solar values and for the equivalent radius of $r = 40$ and 140 pc, respectively, whereas the dashed lines represent the more realistic values of W43 ($G_0 = 5$, $Z = 1.4$). The dashed black line, which represents the outer area of W43 ($r = 140$ pc) with a moderate radiation field ($G_0 = 5$, $Z = 1.4$) might fit the data. However, the model parameters might have unrealistically low values of $\phi_{\rm CNM} = 0.58$ and $\phi_{\rm mol} = 0.85$. Using a stronger radiation field increases $n_{\rm min}$ and therefore decreases $\phi_{\rm CNM}$ to even lower values. Hence, values with smaller distances and/or higher radiation fields predict column densities that are too high.

We do not have a conclusive answer as to why the model suggested by Krumholz et al. (2008, 2009) does not describe W43, but we suggest that the strong UV radiation produced by the central OB cluster (Blum et al. 1999) and several further clusters in the environment are responsible for the dissociation of the molecular hydrogen. Another explanation was given by Motte et al. (2014). They performed a similar analysis and their conclusion was that the analytical model by Krumholz et al. (2008, 2009) cannot describe a complicated molecular cloud complex, as we see several transition layers between H I and H_2 along a single line of sight and the assumption of a simple spherical cloud, without internal radiation sources breaks down.

Table 2. Densities and model parameter.

r [pc]	$n_{\rm HI}$ [cm^{-3}]	$n_{\rm H_2}$ [cm^{-3}]	$\phi_{\rm mol}$	$\phi_{\rm CNM}$ ($G_0 = 1$)	$\phi_{\rm CNM}$ ($G_0 = 5$)
140	20 ± 4	17 ± 5	0.85	2.7	0.58
80	11 ± 3	49 ± 5	4.4	1.5	0.32
40	5 ± 2	101 ± 8	20.2	0.7	0.14

Notes. H I and H_2 particle density for different equivalent radii with corresponding 1σ uncertainty, extracted using the information given in Fig. 14 and using Eq. (14). The given uncertainties are the statistical uncertainties, but do not take the systematical uncertainties due to the saturation of the optical depth spectra into account. The model parameters $\phi_{\rm mol}$ and $\phi_{\rm CNM}$ are calculated using Eqs. (17) and (15), respectively. The parameter $\phi_{\rm CNM}$ is calculated for solar values ($G_0 = 1$, $Z = 1$) and for more realistic values of W43 ($G_0 = 5$, $Z = 1.4$).

6. Conclusions

The H I, OH, Recombination line survey of the Milky Way (THOR) is a Galactic plane survey covering a large portion of the Galactic disk (l = 15–67°, $|b| \leq 1°$). We use the VLA to observe the 21 cm H I line, 4 OH lines, 19 Hα recombination lines and the continuum from 1–2 GHz. We present the H I data of the pilot field centered on the GMC associated with W43 (l = 29.2–31.5°, $|b| \leq 1°$). The main results can be summarized as:

1. We measured the average spin temperature of the neutral hydrogen gas along the line of sight toward eight strong continuum sources. Half of them are Galactic and half of them are extragalactic. We find a median value of T_S = 95.7 K, which is in agreement with other studies.
2. We can estimate the optical depth for the H I line toward strong continuum sources at various locations in W43. The measured optical depth saturates at the main velocity component of W43 (v_{lsr} = 80–100 km s^{-1}) with lower limits of $\tau \sim 2.7$ in the center. Hence, the derived H I masses based on optically thin emission strongly underestimates the hydrogen content. Employing further corrections for the weak and diffuse continuum emission, we obtain a lower limit for the H I mass of $M = 6.6_{-1.8} \times 10^6\ M_\odot$ for a velocity range of v_{lsr} = 60–120 km s^{-1} and an area of l = 29.0–31.5°and $|b| \leq 1°$. This is a factor of ~2.4 larger than the H I mass estimates with the assumption of optically thin emission.
3. The measured H I column density exceeds values of $N_{HI} \sim 150\ M_\odot\,pc^{-2}$ over much of the inner region with $r < 80$ pc. This is an order of magnitude larger than for low-mass star-forming regions such as Perseus.
4. As we corrected the H I column density for optical depth effects and the weak continuum emission, we are able to study the H I distribution spatially even in the innermost part of W43. We assumed an elliptical layered structure for the GMC associated with W43. This allows us to reconstruct the particle density of H I and we find a linearly decreasing H I density toward the center of W43 with values from n_{HI} = 20 to almost 0 cm^{-3}. Furthermore, we compared our results to the density of molecular hydrogen based on *Herschel* continuum data. The density of molecular hydrogen shows an exponential increase toward the center of W43 with values rising from n_{H_2} = 15 to 200 cm^{-3}. For smaller clumps, the density of H_2 can rise to even higher values.
5. We compared our measurements to the analytic model suggested by Krumholz et al. (2008, 2009). Our data does not show a sharp transition between H I and H_2, and nor do we find the predicted threshold for the H I column density of ~10 $M_\odot$ pc^{-2}. Based on these models, larger H I column densities should not exist, as molecular hydrogen should form for such high H I column densities. To fit the model, we have to assume low model parameters, which may indicate that the model is not applicable in a region with such a high radiation field. We suggest that the addition of an internal radiation field from a central cluster may be required to describe the observations. Thus, this work has shown that the H I content of W43 and its relation to H_2 challenges current models of H_2 formation.

Acknowledgements. We like to thank the anonymous referee for an excellent report with a lot of helpful comments that improved the paper. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. R.S.K., S.C.O.G., R.J.S., and S.E.R. acknowledge financial support by the Deutsche Forschungsgemeinschaft (DFG) via the Heidelberg Sonderforschungsbereich SFB 881 The Milky Way System (subprojects B1, B2, and B8) as well as the Schwerpunktprogramm SPP 1573 "Physics of the Interstellar Medium". R.S.K. further acknowledges support from the European Research Council under the European Communitys Seventh Framework Programme (FP7/2007–2013) via the ERC Advanced Grant STARLIGHT (project number 339177). We gratefully acknowledge help from the team of the Pete V. Domenici Science Operations Center (SOC) in Socorro for their help on data reduction during an extended visit in 2013. S.B. is a fellow of the International Max Planck Research School for Astronomy and Cosmic Physics (IMPRS) at the University of Heidelberg and acknowledges support.

References

Anderson, L. D., Bania, T. M., Balser, D. S., & Rood, R. T. 2011, ApJS, 194, 32
Anderson, L. D., Bania, T. M., Balser, D. S., et al. 2014, ApJS, 212, 1
Bally, J., Anderson, L. D., Battersby, C., et al. 2010, A&A, 518, L90
Beuther, H., Tackenberg, J., Linz, H., et al. 2012, A&A, 538, A11
Beuther, H., Ragan, S. E., Ossenkopf, V., et al. 2014, A&A, 571, A53
Bhatnagar, S., Rau, U., Green, D. A., & Rupen, M. P. 2011, ApJ, 739, L20
Bhatnagar, S., Rau, U., & Golap, K. 2013, ApJ, 770, 91
Blum, R. D., Damineli, A., & Conti, P. S. 1999, AJ, 117, 1392
Brunthaler, A., Reid, M. J., Menten, K. M., et al. 2011, Astron. Nachr., 332, 461
Carlhoff, P., Nguyen Luong, Q., Schilke, P., et al. 2013, A&A, 560, A24
Chengalur, J. N., Kanekar, N., & Roy, N. 2013, MNRAS, 432, 3074
Clark, B. G. 1965, ApJ, 142, 1398
Dobbs, C. L., Krumholz, M. R., Ballesteros-Paredes, J., et al. 2014, Protostars and Planets VI, 3
Fukui, Y., Torii, K., Onishi, T., et al. 2015, ApJ, 798, 6
Gibson, S. J., Taylor, A. R., Higgs, L. A., & Dewdney, P. E. 2000, ApJ, 540, 851
Gibson, S. J., Taylor, A. R., Higgs, L. A., Brunt, C. M., & Dewdney, P. E. 2005a, ApJ, 626, 195
Gibson, S. J., Taylor, A. R., Higgs, L. A., Brunt, C. M., & Dewdney, P. E. 2005b, ApJ, 626, 214
Glover, S. C. O., & Clark, P. C. 2012, MNRAS, 421, 9
Goldsmith, P. F., & Li, D. 2005, ApJ, 622, 938
Heiles, C., & Troland, T. H. 2003a, ApJS, 145, 329
Heiles, C., & Troland, T. H. 2003b, ApJ, 586, 1067
Hill, T., Motte, F., Didelon, P., et al. 2011, A&A, 533, A94
Hoare, M. G., Purcell, C. R., Churchwell, E. B., et al. 2012, PASP, 124, 939
Hollenbach, D. J., & Tielens, A. G. G. M. 1997, ARA&A, 35, 179
Kainulainen, J., & Tan, J. C. 2013, A&A, 549, A53
Kainulainen, J., Ragan, S. E., Henning, T., & Stutz, A. 2013, A&A, 557, A120
Kalberla, P. M. W., & Kerp, J. 2009, ARA&A, 47, 27
Krumholz, M. R., McKee, C. F., & Tumlinson, J. 2008, ApJ, 689, 865
Krumholz, M. R., McKee, C. F., & Tumlinson, J. 2009, ApJ, 693, 216
Krčo, M., & Goldsmith, P. F. 2010, ApJ, 724, 1402
Lada, C. J., Alves, J. F., & Lombardi, M. 2007, Protostars and Planets V (Tucson: University of Arizona Press), 3
Lee, M.-Y., Stanimirović, S., Douglas, K. A., et al. 2012, ApJ, 748, 75
Lester, D. F., Dinerstein, H. L., Werner, M. W., et al. 1985, ApJ, 296, 565
Li, D., & Goldsmith, P. F. 2003, ApJ, 585, 823
Liszt, H. S., Braun, R., & Greisen, E. W. 1993, AJ, 106, 2349
Louvet, F., Motte, F., Hennebelle, P., et al. 2014, A&A, 570, A15
Mac Low, M.-M., & Glover, S. C. O. 2012, ApJ, 746, 135
Mac Low, M.-M., & Klessen, R. S. 2004, Rev. Mod. Phys., 76, 125
McClure-Griffiths, N. M., Dickey, J. M., Gaensler, B. M., Green, A. J., & Haverkorn, M. 2006, ApJ, 652, 1339
McClure-Griffiths, N. M., Dickey, J. M., Gaensler, B. M., et al. 2012, ApJS, 199, 12
McKee, C. F., & Ostriker, E. C. 2007, ARA&A, 45, 565
Molinari, S., Swinyard, B., Bally, J., et al. 2010, A&A, 518, L100
Motte, F., Schilke, P., & Lis, D. C. 2003, ApJ, 582, 277
Motte, F., Nguyên Luong, Q., Schneider, N., et al. 2014, A&A, 571, A32
Murray, C. E., Lindner, R. R., Stanimirović, S., et al. 2014, ApJ, 781, L41
Nguyen Luong, Q., Motte, F., Schuller, F., et al. 2011, A&A, 529, A41
Nguyen-Lu'o'ng, Q., Motte, F., Carlhoff, P., et al. 2013, ApJ, 775, 88
Offner, S. S. R., Clark, P. C., Hennebelle, P., et al. 2014, Protostars and Planets VI, 53
Pineda, J. L., Langer, W. D., Velusamy, T., & Goldsmith, P. F. 2013, A&A, 554, A103
Purcell, C. R., Hoare, M. G., Cotton, W. D., et al. 2013, ApJS, 205, 1
Radhakrishnan, V., Murray, J. D., Lockhart, P., & Whittle, R. P. J. 1972, ApJS, 24, 15
Rau, U., Bhatnagar, S., & Owen, F. N. 2014, ArXiv e-prints [arXiv:1403.5242]
Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2014, ApJ, 783, 130

Roy, N., Kanekar, N., Braun, R., & Chengalur, J. N. 2013a, MNRAS, 436, 2352
Roy, N., Kanekar, N., & Chengalur, J. N. 2013b, MNRAS, 436, 2366
Smith, R. J., Glover, S. C. O., Clark, P. C., Klessen, R. S., & Springel, V. 2014, MNRAS, 441, 1628
Stahler, S. W., Palla, F., & Palla, F. 2005, The Formation of Stars (Physics Textbook), 1st edn. (Wiley-VCH)
Stil, J. M., Taylor, A. R., Dickey, J. M., et al. 2006, AJ, 132, 1158
Strasser, S., & Taylor, A. R. 2004, ApJ, 603, 560
Strasser, S. T., Dickey, J. M., Taylor, A. R., et al. 2007, AJ, 134, 2252
Taylor, A. R., Gibson, S. J., Peracaula, M., et al. 2003, AJ, 125, 3145
Vallée, J. P. 2008, AJ, 135, 1301
Walsh, A. J., Beuther, H., Bihr, S., et al. 2015, MNRAS, submitted
Wilson, T. L., Rohlfs, K., & Hüttemeister, S. 2010, Tools of Radio Astronomy (Astronomy and Astrophysics Library), softcover reprint of hardcover 5th ed. 2009 edn. (Springer)
Wolfire, M. G., Hollenbach, D., McKee, C. F., Tielens, A. G. G. M., & Bakes, E. L. O. 1995, ApJ, 443, 152
Wolfire, M. G., McKee, C. F., Hollenbach, D., & Tielens, A. G. G. M. 2003, ApJ, 587, 278
Zhang, B., Moscadelli, L., Sato, M., et al. 2014, ApJ, 781, 89

[1] Max Planck Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany
e-mail: name@mpia.de
[2] National Radio Astronomy Observatory, PO Box O, 1003 Lopezville Road, Socorro, NM 87801, USA
[3] School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, UK
[4] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
[5] Department of Physics and Astronomy, West Virginia University, Morgantown, WV 26506, USA
[6] Universität Heidelberg, Zentrum für Astronomie, Institut für Theoretische Astrophysik, Albert-Ueberle-Str. 2, 69120 Heidelberg, Germany
[7] 1. Physikalisches Institut, Universität zu Köln, Zülpicher Str. 77, 50937 Köln, Germany
[8] Department of Astronomy, University of Wisconsin, Madison, WI 53706, USA
[9] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
[10] Department of Physics and Astronomy, University of North Carolina-Chapel Hill, Chapel Hill, NC 27599-3255, USA
[11] Department of Astronomy, University of Massachusetts, Amherst, MA 01003-9305, USA
[12] Joint ALMA Observatory, Alonso de Cordova 3107, 763-0355 Vitacura, Santiago, Chile
[13] Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
[14] Department of Astronomy and Astrophysics, University of California, 1156 High Street, Santa Cruz, CA 95064, USA
[15] Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK
[16] Australia Telescope National Facility, CSIRO Astronomy and Space Science, Marsfield, NSW 2122, Australia
[17] Laboratoire AIM, CEA/DSM-CNRS-Université Paris Diderot, IRFU/Service d'Astrophysique, Saclay, 91191 Gif-sur-Yvette, France
[18] Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada
[19] Department of Physics and Astronomy, University of Calgary, 2500 University Drive NW, Calgary, AB T2N 1N4, Canada
[20] Department of Physics, Indian Institute of Technology Kharagpur, 721302 Kharagpur, India
[21] Université de Bordeaux, Laboratoire d'Astrophysique de Bordeaux, CNRS/INSU, 33270 Floirac, France
[22] International Centre for Radio Astronomy Research, Curtin University, GPO Box U1987, Perth WA 6845, Australia