A&A 657, A1 (2022)
https://doi.org/10.1051/0004-6361/202141265
© J. Syed et al. 2021

Astronomy & Astrophysics

The "Maggie" filament: Physical properties of a giant atomic cloud⋆

J. Syed[1], J. D. Soler[1], H. Beuther[1], Y. Wang[1], S. Suri[2,1], J. D. Henshaw[1], M. Riener[1], S. Bialy[3], S. Rezaei Kh.[1,4], J. M. Stil[5], P. F. Goldsmith[6], M. R. Rugel[7], S. C. O. Glover[8], R. S. Klessen[8,9], J. Kerp[10], J. S. Urquhart[11], J. Ott[12], N. Roy[13], N. Schneider[14], R. J. Smith[15], S. N. Longmore[16], and H. Linz[1]

[1] Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
e-mail: syed@mpia.de
[2] University of Vienna, Department of Astrophysics, Türkenschanzstrasse 17, 1180 Vienna, Austria
[3] Harvard Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA, 02138, USA
[4] Chalmers University of Technology, Department of Space, Earth and Environment, 412 93 Gothenburg, Sweden
[5] Department of Physics and Astronomy, The University of Calgary, 2500 University Drive NW, Calgary AB T2N 1N4, Canada
[6] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
[7] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
[8] Universität Heidelberg, Zentrum für Astronomie, Institut für Theoretische Astrophysik, Albert-Ueberle-Str. 2, 69120 Heidelberg, Germany
[9] Universität Heidelberg, Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, INF 205, 69120 Heidelberg, Germany
[10] Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany
[11] Centre for Astrophysics and Planetary Science, University of Kent, Canterbury CT2 7NH, UK
[12] National Radio Astronomy Observatory, PO Box O, 1003 Lopezville Road, Socorro, NM 87801, USA
[13] Department of Physics, Indian Institute of Science, Bengaluru 560012, India
[14] I. Physik. Institut, University of Cologne, Zülpicher Str. 77, 50937 Cologne, Germany
[15] Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK
[16] Astrophysics Research Institute, Liverpool John Moores University, IC2, Liverpool Science Park, 146 Brownlow Hill, Liverpool L3 5RF, UK

Received 7 May 2021 / Accepted 1 November 2021

ABSTRACT

Context. The atomic phase of the interstellar medium plays a key role in the formation process of molecular clouds. Due to the line-of-sight confusion in the Galactic plane that is associated with its ubiquity, atomic hydrogen emission has been challenging to study.
Aims. We investigate the physical properties of the "Maggie" filament, a large-scale filament identified in H I emission at line-of-sight velocities, $v_{\rm LSR} \sim -54\,{\rm km\,s^{-1}}$.
Methods. Employing the high-angular resolution data from The H I/OH Recombination line survey of the inner Milky Way (THOR), we have been able to study H I emission features at negative $v_{\rm LSR}$ velocities without any line-of-sight confusion due to the kinematic distance ambiguity in the first Galactic quadrant. In order to investigate the kinematic structure, we decomposed the emission spectra using the automated Gaussian fitting algorithm GAUSSPY+.
Results. We identify one of the largest, coherent, mostly atomic H I filaments in the Milky Way. The giant atomic filament Maggie, with a total length of 1.2 ± 0.1 kpc, is not detected in most other tracers, and it does not show signs of active star formation. At a kinematic distance of 17 kpc, Maggie is situated below (by ≈500 pc), but parallel to, the Galactic H I disk and is trailing the predicted location of the Outer Arm by $5-10\,{\rm km\,s^{-1}}$ in longitude-velocity space. The centroid velocity exhibits a smooth gradient of less than $\pm 3\,{\rm km\,s^{-1}}\,(10\,{\rm pc})^{-1}$ and a coherent structure to within $\pm 6\,{\rm km\,s^{-1}}$. The line widths of $\sim 10\,{\rm km\,s^{-1}}$ along the spine of the filament are dominated by nonthermal effects. After correcting for optical depth effects, the mass of Maggie's dense spine is estimated to be $7.2^{+2.5}_{-1.9} \times 10^5\,M_\odot$. The mean number density of the filament is $\sim 4\,{\rm cm^{-3}}$, which is best explained by the filament being a mix of cold and warm neutral gas. In contrast to molecular filaments, the turbulent Mach number and velocity structure function suggest that Maggie is driven by transonic to moderately supersonic velocities that are likely associated with the Galactic potential rather than being subject to the effects of self-gravity or stellar feedback. The probability density function of the column density displays a log-normal shape around a mean of $\langle N_{\rm H\,I}\rangle = 4.8 \times 10^{20}\,{\rm cm^{-2}}$, thus reflecting the absence of dominating effects of gravitational contraction.
Conclusions. While Maggie's origin remains unclear, we hypothesize that Maggie could be the first in a class of atomic clouds that are the precursors of giant molecular filaments.

Key words. ISM: clouds – ISM: atoms – ISM: kinematics and dynamics – ISM: structure – radio lines: ISM

⋆ Spectral decomposition datacubes 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/cat/J/A+A/657/A1

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Open Access funding provided by Max Planck Society.

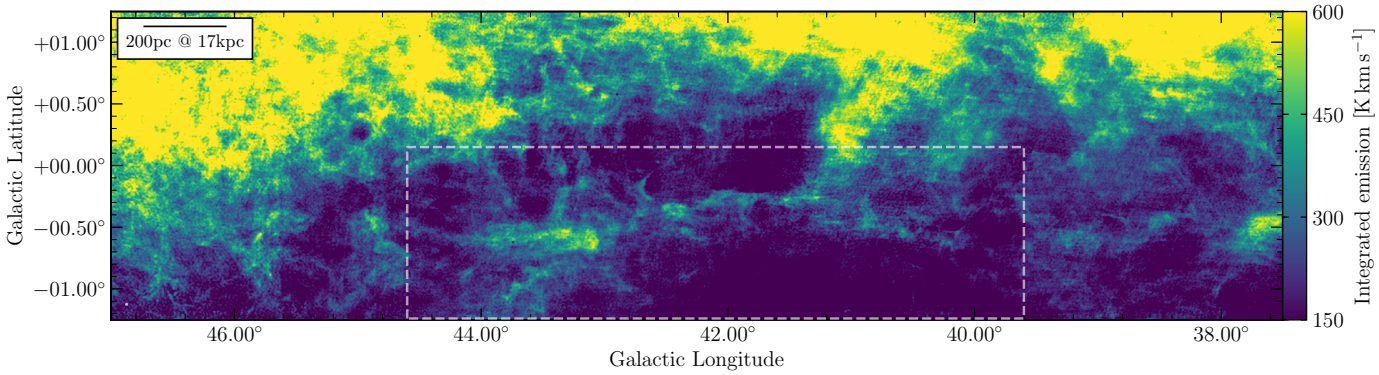

Fig. 1. Overview of Maggie. The map shows the integrated H I emission in the velocity interval between −57.5 and −48.5 km s^{-1}. Maggie is located below the Galactic midplane. The white dashed box containing Maggie marks the region that is considered in the subsequent analysis.

1. Introduction

Stars form in the cold, dense interiors of molecular clouds. The physical properties of these clouds therefore set the initial conditions under which star formation takes place. A key question in understanding the star formation process as part of the global interstellar matter cycle addresses the formation of large-scale molecular clouds out of the diffuse atomic phase of the interstellar medium (ISM; for a review see Ferrière 2001; Draine 2011; Klessen & Glover 2016). However, the atomic ISM and its dynamical properties are still observationally poorly constrained.

The ISM has a hierarchical structure and facilitates the formation of filaments that are governed by the Galactic potential on a large scale. Soler et al. (2020) present a network of H I filaments already evident in the diffuse atomic phase of the ISM that is structured mostly parallel to the Galactic plane. Only in a few cases is the orientation of the filaments locally no longer dictated by the overall drag of the Galactic disk, but rather it is determined by the effects of stellar feedback and strong magnetic fields. Soler et al. (2020) identified a unique H I filament that they have named "Maggie", after the largest river in Colombia, the Río Magdalena. Maggie is shown to be a highly elongated filamentary cloud (see Fig. 1) extending over ∼4° on the sky in Galactic longitude. Given its central velocity of $v_{\rm LSR} \approx -54\,{\rm km\,s^{-1}}$ and assuming circular motion, Maggie is located approximately 17 kpc away from us and has a length of more than 1 kpc.

In this paper, we present a detailed study of the Maggie filament and aim to understand its physical nature. Moreover, we investigate the physical and kinematic properties in comparison with large-scale molecular filaments. In this context, it is an open question if the properties of filamentary molecular clouds are generally inherited from an atomic counterpart in the diffuse ISM. Giant molecular filaments are to date the largest coherent entities identified in the Milky Way and have been subject of many recent studies investigating the evolution of large filaments with respect to the global dynamics of the Milky Way on the one hand and local stellar feedback on the other hand (Jackson et al. 2010; Ragan et al. 2014; Goodman et al. 2014; Wang et al. 2015, 2016, 2020b; Abreu-Vicente et al. 2016; Zucker et al. 2015, 2018).

The highly filamentary infrared dark cloud (IRDC) "Nessie", first identified in Jackson et al. (2010), is argued to be one of the first in a class of filaments whose morphology is likely governed by the structural dynamics of the Galaxy. Goodman et al. (2014) dubbed this type of filaments "bones" of the Milky Way – highly filamentary molecular clouds whose formation, evolution, and shape could be closely linked to the global spiral structure of the Galaxy.

Zucker et al. (2018) utilized a standardized approach to reanalyze a set of Galactic filaments identified in the literature. In doing that, it is possible to draw meaningful conclusions and reliably compare physical properties between filaments rather than being subject to the systematics of selection criteria and different methodology. They found in their sample of filaments that the giant molecular filaments (GMFs), first identified in Ragan et al. (2014) and Abreu-Vicente et al. (2016), have the highest masses (∼$10^5\,M_\odot$) while exhibiting the lowest column densities and star-forming activity, making them good candidates to be the immediate descendants of atomic ISM structures.

The GMFs are first identified as near and mid-infrared extinction features with a spatial extent of ≈100 pc. Ragan et al. (2014) and Abreu-Vicente et al. (2016) then confirm velocity contiguity of the filaments via ^{13}CO emission. Giant molecular filaments are not only associated with spiral arms but are also located in inter-arm regions.

The recently discovered Radcliffe wave (Alves et al. 2020) is a coherent 2.7 kpc long association of local molecular cloud complexes. It appears to be undulating above and below the Galactic midplane, and its three-dimensional shape (in position-position-position space) is well described by a damped sinusoidal wave. The Radcliffe wave provides a framework for future studies of molecular cloud formation and evolution with respect to the Galactic dynamics of the Milky Way.

Simulations find that giant molecular clouds often form as large filaments, and their formation can be related to the dynamics of the galaxy and position with respect to the spiral arm potentials. Smith et al. (2014a) use simulations of a four-armed spiral galaxy to investigate the formation of molecular gas structures. They find that high-density filaments tend to form in spiral arms while lower-density gas resides in long inter-arm spurs that are stretched by galactic shear. While the simulations account for the chemical evolution of the gas, they do not include the effects of self-gravity of the gas or stellar feedback.

Including both stellar feedback and self-gravity, smoothed particle hydrodynamics simulations of a high-resolution section of a spiral galaxy indicate that galactic shear between spiral arms plays a critical role in the formation of highly filamentary structures (Dobbs 2015; Duarte-Cabral & Dobbs 2016). Duarte-Cabral & Dobbs (2017) expand this analysis to the time evolution of these filaments, finding that giant molecular clouds tend to sustain their large filamentary shape before entering the spiral

arms, where they are prone to being disrupted by local events of star formation and stellar feedback.

The CloudFactory simulations presented in Smith et al. (2020) strongly suggest that spiral arms and differential rotation tend to arrange molecular clouds in long filaments that are likely to fragment while regions of clustered feedback randomizes the orientation of them. Molecular filaments that are dominated by galactic-scale forces also exhibit low internal velocity gradients and are tightly confined to the galactic plane. Filaments formed in regions of higher turbulence due to supernova (SN) feedback show a range of orientations with respect to the galactic plane and are more widely distributed.

The systematic search for atomic counterparts to molecular clouds poses a challenging task. By means of the 21cm-line of H I emission it is generally possible to probe large atomic clouds in the ISM but it has proven difficult to identify them as clearly defined objects in the inner Galactic plane. Traditionally, H I clouds are then either observed at high Galactic latitudes (e.g., Kalberla et al. 2016) or they must have velocities significantly different than those imposed by the Galactic rotation (e.g., Wakker & van Woerden 1997; Westmeier 2018).

As the analysis of H I emission in the Galactic plane within the solar circle suffers from the kinematic distance ambiguity, any H I cloud with $v_{\rm LSR} \gtrsim 0\,{\rm km\,s^{-1}}$ (in the first Galactic quadrant) might be the product of blending foreground and background components that correspond to the same line-of-sight velocity. However, at negative $v_{\rm LSR}$ velocities any H I structure can be identified with less line-of-sight confusion[1] (see e.g., Brand & Blitz 1993).

Additionally, the physical properties of atomic hydrogen are not straightforward to derive from emission studies alone. In thermal pressure equilibrium, theoretical considerations based on ISM heating and cooling processes predict two stable phases of atomic hydrogen at the observed pressures in the ISM, namely the cold neutral medium (CNM) and warm neutral medium (WNM; Field et al. 1969; McKee & Ostriker 1977; Wolfire et al. 2003; Bialy & Sternberg 2019). Observations of H I emission are thus generally attributed to both CNM and WNM, which have significantly different physical properties. In an attempt to observationally isolate the CNM from the bistable emission, H I self-absorption (HISA; Gibson et al. 2000; Li & Goldsmith 2003; Wang et al. 2020c; Syed et al. 2020) is a viable method to study H I clouds in the inner Milky Way but it heavily depends on the presence of sufficient background emission.

The CNM is observed to have temperatures $\lesssim$300 K and number densities of $\gtrsim n_{\rm min,CNM} = 10\,{\rm cm^{-3}}$ while the thermally stable WNM exceeds temperatures of ~5000 K with number densities $\lesssim n_{\rm max,WNM} = 0.1\,{\rm cm^{-3}}$ (Heiles & Troland 2003; Kalberla & Kerp 2009). For densities between $n_{\rm min,CNM}$ and $n_{\rm max,WNM}$, the gas is thermally unstable (denoted by UNM: unstable neutral medium) and it will move toward a stable CNM or WNM branch under isobaric density perturbations (Field 1965).

Throughout the scope of this paper, we use the analytic model presented in Wolfire et al. (2003) as a standard reference. According to this, the two stable atomic hydrogen phases can coexist in thermal equilibrium over a narrow range of pressures, $P_{\rm min} \leq P \leq P_{\rm max}$, that is governed by the total heating and cooling rates of the interstellar gas (see e.g., the phase diagrams shown in Fig. 7 in Wolfire et al. 2003). The dominant heating process under typical ISM conditions is photoelectric (PE) heating from dust grains and polycyclic aromatic hydrocarbons (PAHs). The WNM is mainly cooled by Lyα emission while the most efficient cooling mechanism of the CNM is metal-line fine-structure emission primarily excited by collisions with electrons and neutral hydrogen atoms. Assuming a constant ratio of the cosmic-ray (or X-ray) ionization rate, ζ, to the intensity of the UV interstellar radiation field, $I_{\rm UV}$, PE heating is proportional to $I_{\rm UV}$. If $I_{\rm UV}$ is then increased (decreased), the density at which metal-line cooling balances heating must increase (decrease) in the same manner. As a result, the pressure range at which both WNM and CNM occur shifts to higher or lower values in proportion to $I_{\rm UV}$.

Furthermore, the metallicity Z' and dust abundance are crucial parameters for the thermal balance. If the metallicity is moderately lowered, the equilibrium pressure range dictating the WNM and CNM properties decreases since the dust-to-gas ratio and associated PE heating rate decreases more rapidly with decreasing Z' than does the metal-line cooling[2] (Bialy & Sternberg 2019, see their Fig. 4).

The Maggie filament has been discovered in Soler et al. (2020) via H I emission in The H I/OH Recombination line survey of the inner Milky Way (THOR; Beuther et al. 2016; Wang et al. 2020a). It is identified as a large coherent filament using a Hessian matrix approach. This method allows to systematically identify filamentary structures by their spatial curvature that emerge as second derivative signatures in the H I emission maps of the THOR survey.

We follow up on this discovery and derive the physical properties of the Maggie filament in the subsequent analysis. This paper is organized as follows. We introduce the observations and data in Sect. 2 and outline the method we use for the spectral decomposition of the H I emission. We present the fundamental properties of Maggie and show the results of the kinematics and column density in Sect. 3. In Sect. 4 we examine Maggie for molecular counterparts and explore the possible formation process of Maggie. We also discuss the implications of the velocity and column density structure. Finally, we review the relationship between Maggie and giant molecular filaments and draw our conclusions in Sect. 5.

2. Observation and methods

2.1. H I 21 cm line and continuum

In the following analysis we employ the H I and 1.4 GHz continuum data from the THOR survey. A complete overview of the THOR survey and the data products is given in Beuther et al. (2016) and Wang et al. (2020a). We investigate the kinematics based on the H I emission data, where the 1.4 GHz continuum emission has been subtracted from the THOR H I observation during data reduction. The final H I emission data (THOR-H I) have been obtained from observations with the Very Large Array (VLA) in C- and D-configuration, as well as single dish observations from the Greenbank Telescope. The data have an angular resolution of 40″ and the noise in emission-free channels is ~4 K at the spectral resolution of $1.5\,{\rm km\,s^{-1}}$. The covered velocities range from -113 to $+163\,{\rm km\,s^{-1}}$.

In Sect. 3.3 we correct for the optical depth that we measured against discrete continuum sources. We therefore select the THOR H I+continuum data that consist of VLA C-array configuration data only (THOR-only). THOR-only data have

[1] Vice versa in the fourth Galactic quadrant.

[2] Here we do not consider the effect of cosmic ray heating that becomes the dominant heating mechanism at metallicities $Z' \lesssim 0.1$ (see Fig. 2 in Bialy & Sternberg 2019).

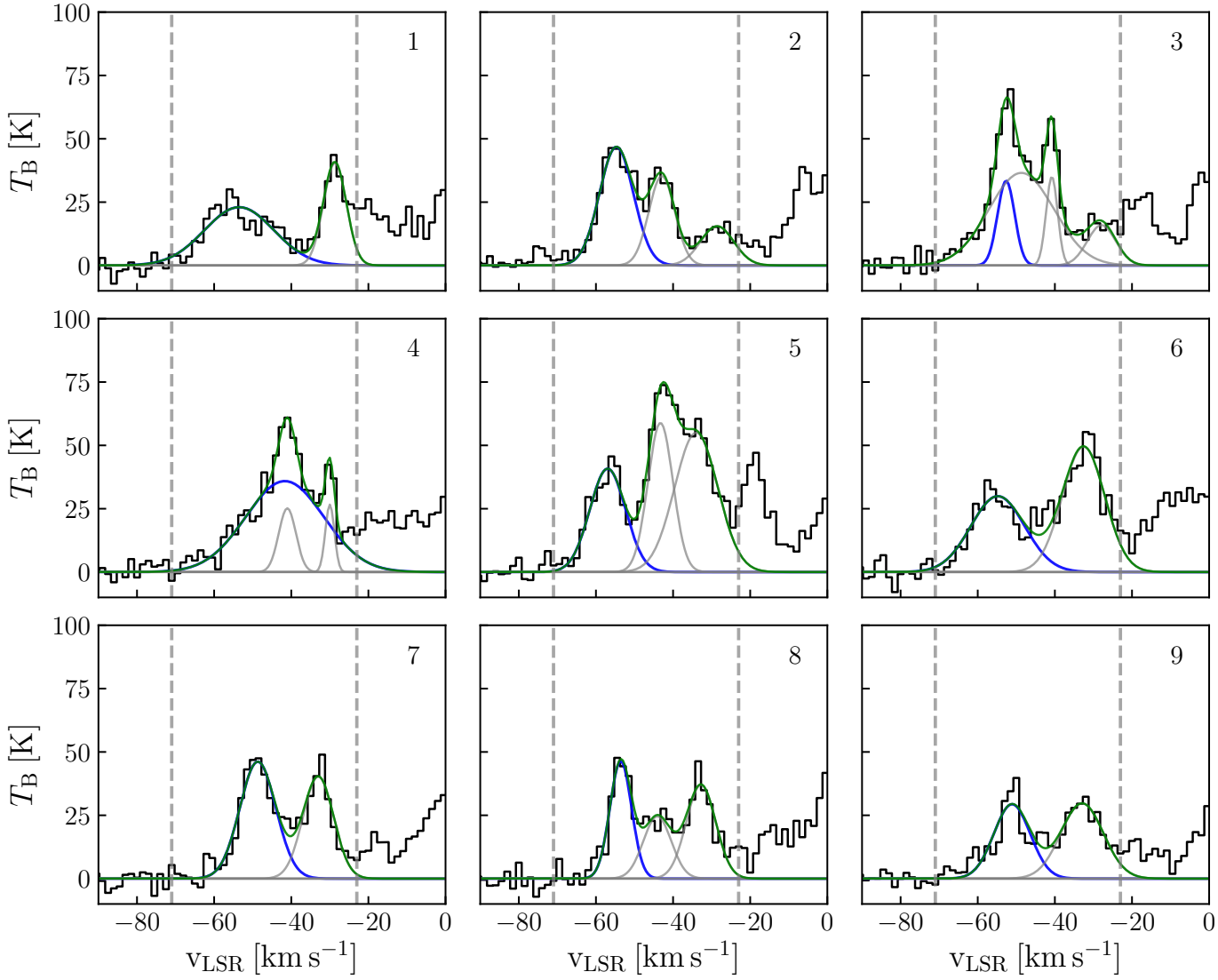

Fig. 2. H I spectra and GAUSSPY+ decomposition results along the filament. The black curves show H I spectra corresponding to the positions marked in the middle panel of Fig. 3. The vertical dashed lines at −71 and −23 km s^{-1} mark the velocity range taken into account for the GAUSSPY+ decomposition. The blue curve in each panel shows the first (i.e., lowest-centroid-velocity) component, which we attribute to Maggie. The gray curves show all additional components that were fit. The green curve represents the total spectrum fit in the considered velocity range.

a higher angular resolution of ∼14″, making them suitable to study absorption against discrete continuum sources. Since this data set consists of interferometric observations only, large-scale H I emission is effectively filtered out. This large-scale filtering due to the interferometer can be exploited when only discrete continuum sources are of interest.

2.2. Gaussian decomposition

The THOR H I emission spectra show Gaussian-like structures at negative velocities, with usually three or fewer components superposed. We therefore use the fully automated Gaussian decomposition algorithm GAUSSPY+[3] (Riener et al. 2019) to study the kinematics of the filament.

GAUSSPY+ is a multicomponent Gaussian decomposition tool based on the earlier GAUSSPY algorithm (Lindner et al. 2015) and provides an improved fitting routine and a fully automated means to decompose emission spectra using machine-learning algorithms. GAUSSPY+ automatically determines initial guesses for Gaussian fit components using derivative spectroscopy. To decompose the spectra, the spectra require smoothing to remove noise peaks while retaining real signal. The optimal smoothing parameters are found by employing a machine-learning algorithm that is trained on a subsection of the data set.

As Maggie is found at negative velocities and the H I emission toward the inner Galactic plane with $v_{LSR} \gtrsim 0$ km s^{-1} is ubiquitous in the spectra, it would not be sensible to fit the whole spectra with GAUSSPY+. Instead, to save computational resources and achieve a better decomposition performance, we masked all spectral channels at velocities ≤ -71 km s^{-1} and ≥ -23 km s^{-1} (see Fig. 2). Maggie has velocities around −54 km s^{-1} (see Fig. 3 and Sect. 3.2) and we aim to disentangle components that might blend in with Maggie.

However, it is essential to reliably estimate the noise in the spectra to obtain good fit results. GAUSSPY+ comes with an automated noise estimation routine as a preparatory step for the decomposition. For this step, we supplied the full spectra to

[3] https://github.com/mriener/gausspyplus

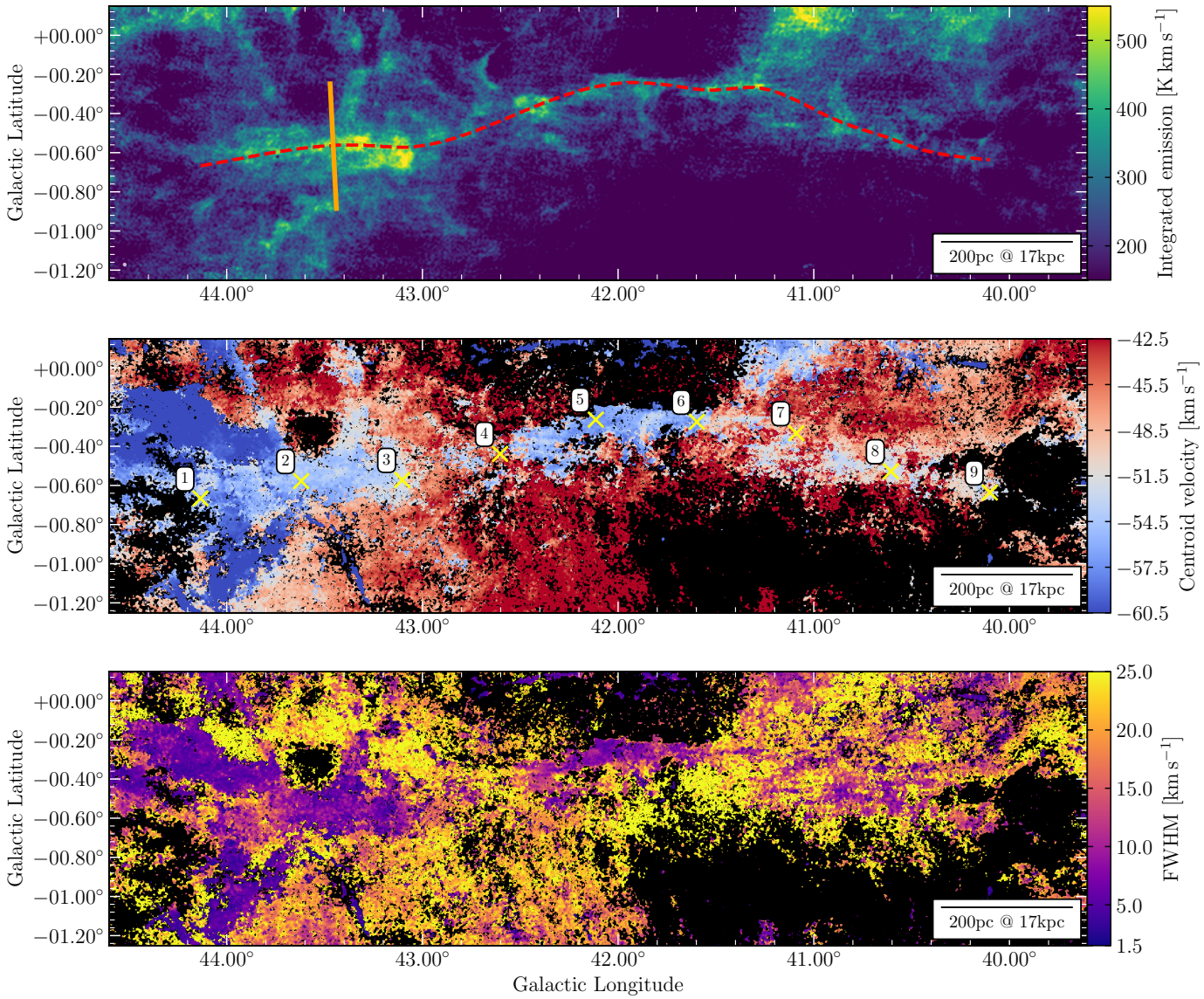

Fig. 3. Kinematic overview of the Maggie filament. *Top panel*: integrated H I emission in the velocity interval between −57.5 and −48.5 km s^{-1}. The red dashed curve marks the spine of the filament. The orange line perpendicular to the spine marks the cut of the average emission profile shown in black in the upper panel of Fig. 4. *Middle panel*: centroid velocities of the first (or single) component derived from the GaussPy+ decomposition of the emission spectra. The fits are restricted to the velocity range $-71 \leq v_{\rm LSR} \leq -23$ km s^{-1}. The yellow crosses mark the positions of the spectra shown in Fig. 2. *Bottom panel*: corresponding line widths in terms of the full width at half maximum (FWHM). Both maps are clipped at 5σ (~20 K).

GaussPy+ as the masked spectra might not contain enough noise channels. The noise map derived in that way was then used for the decomposition of Maggie.

We ran the GaussPy+ training step with 500 randomly selected spectra from the H I data to find the optimal smoothing parameters for the fitting, as recommended in Riener et al. (2019). For H I observations, we would expect both narrow and broad line widths owing to the multiphase (i.e., WNM-CNM-UNM) nature of H I emission. To account for that, we chose not to refit broad or blended components, parameters that can easily be adjusted in the GaussPy+ routine.

After the initial fitting, the algorithm applies a two-phase spatial coherence check that can optimize the fit by refitting the components based on the fit results of neighboring pixels. Figure 2 shows example spectra along the Maggie filament marked in Fig. 3, and the final fit results from the GaussPy+ decomposition. Between two and three components were typically fit by GaussPy+ in the given velocity range. In every pixel spectrum, we selected the component with the lowest centroid velocity as the "Maggie" component (see Sect. 3.2) since no significant portion of H I emission is found at $v_{\rm LSR} \lesssim -60$ km s^{-1}. Due to blended components and strong emission at higher velocities, a selection based on amplitude of the components fails to recover the structure identified in Soler et al. (2020) as the Maggie filament.

3. Results

3.1. Location, distance, and morphology

Figure 1 shows the integrated H I emission covering the Galactic plane at longitudes and latitudes $37.5° < \ell < 47.0°$ and $|b| < 1.25°$, respectively. The filament Maggie subtends an area from $(\ell, b) = (40.4, -0.3)°$ to $(44.2, -0.9)°$ in Galactic longitude

and latitude, and has a velocity around $-54\,\mathrm{km\,s^{-1}}$ with respect to the local standard of rest (LSR). In this region in position-position-velocity (p-p-v) space, the Galactic H I disk shows a warp toward higher latitudes (see e.g., Burton 1988; Dickey & Lockman 1990; Sparke 1993; Dickey et al. 2009), which places Maggie at a location significantly displaced from the Galactic midplane. The H I midplane lies even beyond the coverage of the THOR survey at $b \geq +1.25°$. Therefore, the midplane and Maggie are separated by at least $\Delta b \geq 1.6°$ on the plane of the sky.

In the forthcoming analysis we assumed the Galactic rotation model by Reid et al. (2019) and used the Bayesian distance calculator of the BeSSeL survey (Reid et al. 2016) to translate the location in p-p-v space into a physical distance. Maggie's estimated distance away from us is $d = 17 \pm 1$ kpc and its distance to the Galactic center is $R_{\mathrm{GC}} = 12 \pm 1$ kpc. This physical scale locates the filament ≥470 pc below the Galactic midplane, which is greater than the average H I scale height (~200 pc) at this Galactocentric radius (Kalberla & Kerp 2009).

The Maggie filament discloses a hub-like feature in the east, on which smaller-scale filaments appear to converge, and a tail that thins out toward the west. The northwestern part shows a connection to the midplane, potentially feeding off the H I material located at higher latitudes. Most of the filament, however, appears to be disconnected from the Galactic midplane material.

We define a backbone that runs through Maggie by selecting a spline based on visual inspection of the integrated emission (see top panel in Fig. 3). The length and width of Maggie are estimated using the Filament Characterization Package[4] (FilChap; Suri et al. 2019). These properties are not heavily affected by a potential misplacement of Maggie's spine.

The length of the filament as marked by the red spine in the top panel of Fig. 3 is $\ell_{\mathrm{tot}} = 1.2 \pm 0.1$ kpc assuming a distance of 17 kpc. The uncertainty in length is dominated by the uncertainty in distance. The width of the filament was estimated using the integrated emission map shown in Fig. 3. We took perpendicular slices of 200 pc width with a step size of 10 pc (~3 × beam) along the filament. We then averaged the radial emission profiles over three neighboring slices and fit Gaussian and Plummer-like functions to the averaged emission profile. Plummer-like functions (as functions of cylindrical radius r) have the form

$$I(r) \propto \frac{\rho_{\mathrm{c}} R_{\mathrm{flat}}}{\left[1 + (r/R_{\mathrm{flat}})^2\right]^{\frac{p-1}{2}}}, \tag{1}$$

where p is the power-law index, ρ_{c} is the density at the center of the filament, and R_{flat} is the characteristic radius of the flat inner portion of the density profile. More details about the fitting can be found in Suri et al. (2019).

Plummer-like functions are used to describe the radial column density (i.e., integrated emission in the optically thin limit) profile of a cylindrically shaped filament that has a flat density distribution up to R_{flat} and a power-law fall-off beyond (Arzoumanian et al. 2011). We used constant power-law indices $p = 2$ and 4 for the Plummer-like functions due to the degeneracy between ρ_{c} and R_{flat} (and therefore p) in finding the best fit solution (see Smith et al. 2014b). The power law index $p = 4$ describes the Ostriker (1964) model of an isothermal filament in hydrostatic equilibrium. The power-law index $p = 2$ is derived by Arzoumanian et al. (2011) as a free fit parameter from small-scale filamentary structures seen in *Herschel* data and might be attributed to nonisothermal or magnetized filaments (see Arzoumanian et al. 2011, and references therein).

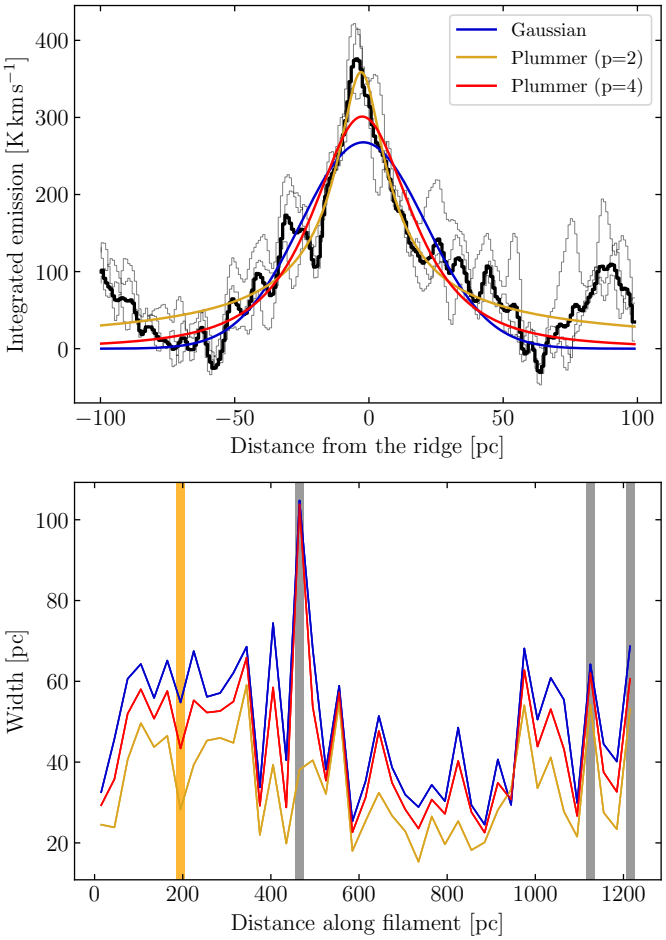

Fig. 4. Width of the Maggie filament. *Top panel*: example emission profile perpendicular to the filament. At each step along the filament, three neighboring emission profiles (gray) are taken to estimate an average profile (black). A fit baseline has been subtracted from each profile. The cut along which the average emission profile is measured is marked in the upper panel of Fig. 3. To derive the widths, we fit the average emission profiles with Gaussian (blue), and Plummer-like functions with index $p = 2$ (yellow) and $p = 4$ (red), respectively. *Bottom panel*: fit widths of Maggie are plotted against the distance along the filament. The colors correspond to the fits shown in the upper panel. The gray bars show the widths that are discarded as these profiles could not be fit properly. The orange bar marks the position of the example profile shown in the upper panel.

The fit widths along the filament are shown in Fig. 4, with zero distance at the outermost point in the east. The widths derived from Gaussian fits are systematically larger than the widths estimated by Plummer-like functions. While the Plummer-like function with an index of $p = 2$ is able to reliably reproduce individual peaks (upper panel of Fig. 4), the fits on average underestimate the width of the emission profiles. Gaussian fits do not fully recover the amplitude of the profiles but give a more accurate representation of the wings. As the noise and blending of different emission features make it difficult to determine which function represents the emission profiles best, we consider the fits equally good. The reduced chi-squared distributions of the fits are similar and do not allow any preference of a fitting function. Three positions along the filament (vertical gray bars in Fig. 4) are discarded as the fits gave no

[4] https://github.com/astrosuri/filchap

reliable results, regardless of the fitting function. For simplicity, we report the average width taken over the full length of the filament. Taking into account all fitting methods, the average width is $w = 40 \pm 6$ pc, which is well resolved by the beam of the THOR observations corresponding to a spatial scale of ≈3.3 pc. The uncertainty is dominated by the variance of the measured widths along the filament. The filament width is only weakly affected by the velocity integration range. If we integrate the emission over the velocity range –57.5 to –42.5 km s^{-1}, the average filament width becomes 44 pc, which is well within the uncertainty of our derived width. However, due to blending emission components toward higher velocities, the derivation of the filament width becomes increasingly difficult with increasing velocity integration range. We therefore constrain the integration range to roughly the lower and upper quartile of the velocity distribution (see Sect. 3.2).

One of the characteristics that is usually used to define filaments is the projected aspect ratio (see e.g., Zucker et al. 2015, 2018). Taking the average width, Maggie has an aspect ratio of (30:1). Although Maggie is identified through atomic line emission, as opposed to the selection of molecular filaments investigated by Zucker et al. (2018), Maggie roughly agrees with the overall aspect ratios of other large-scale filaments (Wang et al. 2015) and the identified Milky Way bones (Goodman et al. 2014; Zucker et al. 2015).

3.2. Kinematics

In the following section, we focus in detail on the kinematic properties of Maggie. We show in the middle and bottom panel of Fig. 3 the maps of the fit centroid velocities and line widths of the first fit component (corresponding to the lowest-velocity component of the GAUSSPY+ decomposition), to which we refer as the "Maggie component" in the following. Accordingly, the second fit component refers to the next component at higher centroid velocity. Overall, the peak velocities along the filament vary between approximately −60 and −45 km s^{-1} and exhibit an undulating pattern. For the most part, the velocity map clearly reflects the spine of the filament, as components off the spine have significantly larger velocities but are picked up as the first component in the absence of Maggie. Given the large spatial scales, the velocities along the spine show a smooth gradient of less than ±3 km s^{-1} (10 pc)$^{-1}$, a kinematic property that satisfies the bone criteria in Zucker et al. (2015).

The line widths as well as centroid velocities shown in Fig. 3 indicate a stark contrast between the filament and the background. Furthermore, the kinematics in terms of the centroid velocity and line width along the filament spine are moderately correlated, where narrow line width corresponds to lower centroid velocity. Figure 5 presents the relationship between the centroid velocity and line width along the 40 pc-wide spine of the filament. The Pearson correlation coefficient between the velocity and line width distribution is 0.76, implying that under the assumption of a linear dependence ∼60% of the variance in the line width distribution can be explained by the scatter present in the centroid velocity distribution. We note that this does not necessarily imply a causal relation between the line width and velocity. Rather, this correlation is likely to be due to contamination from background gas at positions where the filament emission is less well defined and difficult to disentangle from blending components.

Figure 6 shows the histogram of centroid velocities of Maggie, where all individual pixel positions along the 40 pc-wide spine of the filament are being sampled. The centroid velocities

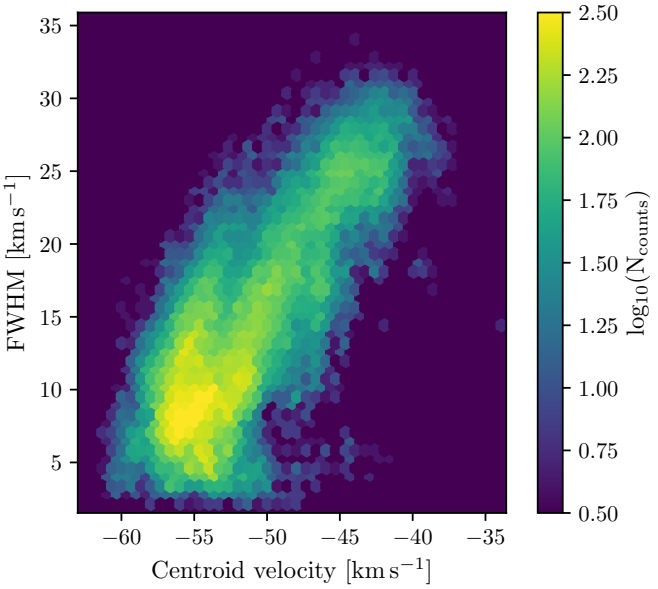

Fig. 5. Line width–centroid velocity correlation. The scatter plot shows the line width (FWHM) as a function of the centroid velocity at all pixel positions along the 40 pc-wide spine of Maggie.

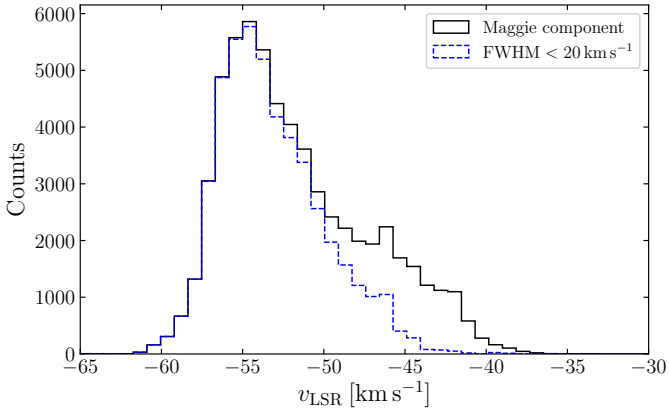

Fig. 6. Histograms of fit peak velocities of Maggie. The black histogram shows the distribution of the centroid velocities at all pixel positions inferred from the GAUSSPY+ fits along the 40 pc-wide spine of the filament. The blue histogram shows the peak velocity distribution corresponding to components with line widths $\Delta v < 20$ km s^{-1}.

along the filament exhibit a strong peak at −54 km s^{-1}, with some skewness toward −45 km s^{-1} that might be attributed to velocity blending of broad WNM components at higher velocities.

Figure 7 offers a position-velocity view of Maggie and its location with respect to the average H I emission. We also overplot the predicted locations of the spiral arms as given by Reid et al. (2019). The bulk of H I material and the second fit component of our spectral decomposition trace out the spine of the Outer arm remarkably well while Maggie exhibits an offset of 5–15 km s^{-1}. Given the distance from the midplane, neither spatially nor kinematically can Maggie be associated with any spiral arm structure. We note, however, that different spiral arm models can vary by 10 km s^{-1} or more in position-velocity space, a concern raised by Zucker et al. (2015) when classifying bones of the Milky Way.

The histogram of line widths in Fig. 8 shows a broad distribution, with line widths between 4 and 30 km s^{-1}. The histogram has a peak around 10 km s^{-1} and a shoulder that extends to

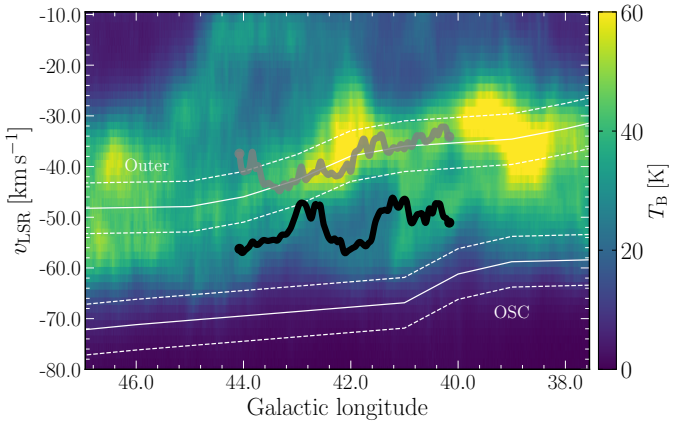

Fig. 7. Position-velocity diagram of H I. The colored background shows the H I emission averaged over $-1.25° < b < +1.25°$. The black and gray curve indicate the intensity-weighted average velocity of Maggie and the second fit component at each longitude as given by the GAUSSPY+ decomposition, respectively. The white line segments give the predicted spiral arm centroids of the Outer arm and Outer-Scutum-Centaurus (OSC) arm along with their uncertainties (dashed lines) in velocity (Reid et al. 2019).

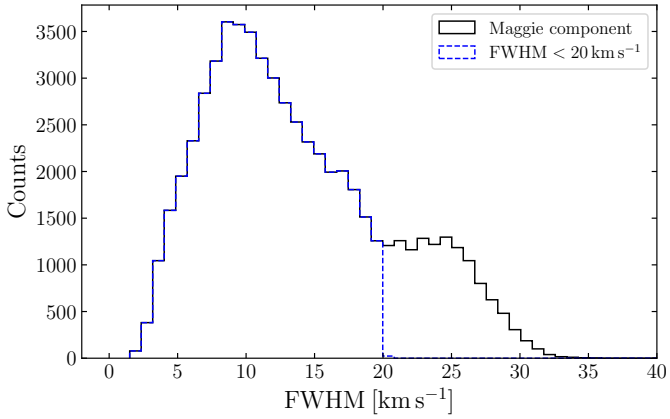

Fig. 8. Histograms of fit line widths of Maggie. The black histogram shows the distribution of the line widths at all pixel positions in terms of the FWHM inferred from the GAUSSPY+ fits along the 40 pc-wide spine of the filament. The blue histogram shows the line width distribution corresponding to components with line widths $\Delta v < 20\,{\rm km\,s^{-1}}$.

25 km s^{-1}. We additionally show in Figs. 6 and 8 the histograms corresponding to velocity components with a line width $\Delta v < 20\,{\rm km\,s^{-1}}$, discarding the shoulder in the line width distribution. The median values of the total centroid velocity distribution is $-52.7\,{\rm km\,s^{-1}}$, and only taking into account line widths $\Delta v < 20\,{\rm km\,s^{-1}}$, the median shifts to $-53.8\,{\rm km\,s^{-1}}$.

To study the kinematic structure with respect to the position along the filament we show in Fig. 9 the average velocity of each 40 pc-wide slice weighted by the amplitude at each pixel position. The uncertainties are estimated by the standard deviation of each slice. The wave-like velocity structure of Maggie is now more clearly visible. The Maggie velocities beyond $>-48.5\,{\rm km\,s^{-1}}$ around 400 pc distance (corresponding to the region around spectrum 4 in Fig. 2) are difficult to clearly separate from the second fit component. We note that the mean velocity and particularly line width have to be treated with some caution as the velocity blending makes it difficult to distinctively identify Maggie in this part of the region. We revisit the velocity

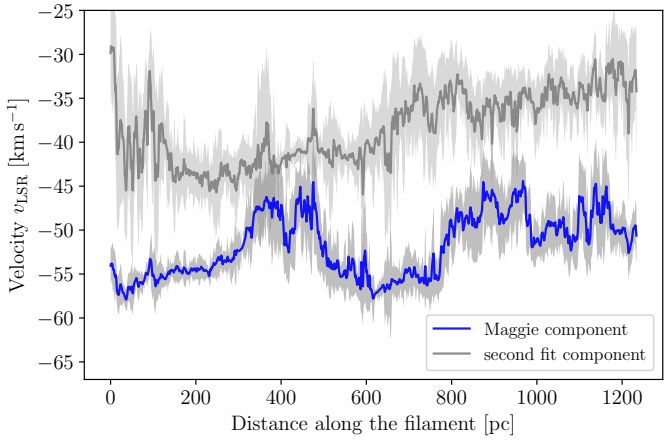

Fig. 9. Position-velocity plot along the filament. This plot shows the centroid velocity over distance along the filament marked by the red spine in the upper panel of Fig. 3. The centroid velocity is an intensity-weighted average over the 40 pc-wide slice perpendicular to every filament position. The blue curve represents the designated Maggie component, that is the first fit component of the GAUSSPY+ results. The gray curve shows the second fit component for comparison. The gray shaded area around both curves indicates the uncertainty estimated by the standard deviation over each slice.

structure in Sect. 4.2.2 to investigate signatures of characteristic spatial scales.

3.3. Column density and mass

Measuring H I absorption against strong continuum sources allows us a direct derivation of the optical depth. We therefore use the high angular resolution H I+continuum data, that consist of VLA C-array observations only, to filter out large-scale emission and measure H I absorption against discrete continuum sources. We measured H I+continuum spectra against seven continuum sources taken from Wang et al. (2018) that have brightness temperatures ≥200 K for a sufficient signal-to-noise ratio. The sources are located both within and outside the area of Maggie (i.e., within and outside the contour at 4.2 $M_\odot\,{\rm pc^{-2}}$, see later in this section). Since the synthesized beam of the native C-configuration data is ∼14″, we extracted an average H I+continuum spectrum from a $16'' \times 16''$ (4×4 pixels) area. To the first order, we can then estimate the H I optical depth by (see Bihr et al. 2015)

$$\tau_{\rm simplified} = -\ln\left(\frac{T_{\rm on}}{T_{\rm cont}}\right), \tag{2}$$

where $T_{\rm on}$ is the H I brightness temperature against the continuum source and $T_{\rm cont}$ is the brightness of the continuum source. Using the optical depth, we compute the H I spin temperature with

$$T_{\rm s} = \frac{T_{\rm off}}{1 - e^{-\tau}}. \tag{3}$$

We measured the H I brightness at an offset position $T_{\rm off}$ using the combined THOR H I data (see Sect. 2.1). We therefore selected an annulus around each source with inner and outer radii of 60″ and 120″, respectively, to measure an averaged $T_{\rm off}$.

All spectra measured against Galactic H II regions do not exhibit any absorption features in the velocity range of Maggie (Table 1). This gives additional support to the argument that

Table 1. Continuum sources.

Source ID	Glon [°]	Glat [°]	$T_{\rm cont}$ [K]	Type	L.o.s.	Absorption
G43.921-0.479	43.92	−0.48	304	extragal.	✓	✓
G43.890-0.783	43.89	−0.78	481	H II	✗	✗
G43.738-0.620	43.74	−0.62	503	extragal.	✓	✓
G43.177-0.519	43.18	−0.52	296	H II	✓	✗
G42.434-0.260	42.43	−0.26	277	H II	✗	✗
G42.028-0.605	42.03	−0.60	519	extragal.	✗	✗
G41.513-0.141	41.51	−0.14	205	H II	✗	✗

Notes. Column 1 gives the source name as listed in the continuum catalog by Wang et al. (2018). Columns 2 and 3 are the Galactic longitude and latitude, respectively. Column 4 denotes the average brightness temperature of the continuum source over an area of $16'' \times 16''$. Column 5 describes the physical nature of the continuum source. Column 6 marks if the continuum source is located along the line of sight (L.o.s), that is within the contour at 4.2 $M_\odot$ pc^{-2} (see Sect. 3.3). Column 7 indicates if absorption above 3σ is detected in the velocity range of Maggie.

the Maggie filament is in fact on the far side of the Galaxy as Galactic H II regions are located in the foreground and can therefore not be detected in absorption. Furthermore, we find that no absorption is evident in the spectra of both Galactic and extragalactic continuum sources offset from the line of sight of Maggie.

The extragalactic continuum sources G43.738-0.620 and G43.921-0.479[5] are the only sources along Maggie against which we detect H I in absorption. We show in Fig. 10 an overview of the optical depth and spin temperature measurement toward G43.738-0.620.

The top panel illustrates the averaged H I emission profile. In pressure equilibrium, the two stable phases of atomic hydrogen (CNM and WNM) both contribute to the emission. It is difficult to disentangle the properties of either phase from the emission alone, so the information from the absorption profile is critical to learn about the nature of the atomic gas. As the optical depth is proportional to $T_{\rm s}^{-1}$, we assume that any absorption measured toward the continuum is due to the CNM.

For simplicity we fit both the emission and absorption profile in this specific case (Fig. 10) with a single Gaussian component. The line width of the emission profile is 14.1 km s^{-1} and therefore significantly wider than the absorption line (8.8 km s^{-1}), suggesting that there is contribution to the emission from a broader component that is not evident in the absorption spectrum. However, the line width contributions are not able to pin down the column density fractions of either the CNM or WNM as the (kinetic) gas temperatures are unknown and the line widths can be significantly broadened by nonthermal effects. The total line width, however, can therefore set an upper limit on the kinetic temperature (Heiles & Troland 2003). We note that the nonthermal contribution to the CNM component, due to effects such as turbulent motion, should be dominant as the thermal line width of the CNM even with the loose constraint $T_{\rm k} < 500$ K (e.g., Heiles 2001; Heiles & Troland 2003) would account for a line width <4.8 km s^{-1}.

The optical depth above the 3σ limit, that we ascribe to the CNM, is plotted in the third panel of Fig. 10. The mean and velocity-integrated optical depth over the velocity range of Maggie is $\tau_{\rm mean} = 0.33$ and $\int \tau(v){\rm d}v = 2.5$ km s^{-1}, respectively. We note that the reported values in Wang et al. (2020a) are inferred from integrating over the whole velocity range of the THOR survey and are thus significantly larger. We use the optical depth and brightness temperature from the emission profile to compute the spin temperature (Eq. (3)) in the bottom panel of Fig. 10. The mean spin temperature is $\langle T_{\rm s} \rangle = 140$ K. Owing to the mixing of warm and cold H I along the line of sight, this spin temperature should not be interpreted as a gas kinetic temperature. Rather, it is a density-weighted harmonic mean of the kinetic temperature under the assumption that $T_{\rm s} = T_{\rm k}$ (see Dickey et al. 2000, 2003, 2009). If we assume that the warm H I does not contribute significantly to the optical depth, that the correction for H I self-absorption is negligible and that the CNM has a roughly constant temperature, then we can derive the CNM fraction via

$$f_{\rm CNM} \equiv \frac{N_{\rm CNM}}{N_{\rm CNM} + N_{\rm WNM}} \simeq \frac{T_{\rm CNM}}{\langle T_{\rm s} \rangle}, \tag{4}$$

where $N_{\rm CNM}$ and $N_{\rm WNM}$ are the column density of the CNM and WNM, respectively, and $T_{\rm CNM}$ is the CNM spin temperature. A representative value for the CNM spin temperature should lie within $40 < T_{\rm CNM} < 100$ K (Heiles & Troland 2003; Strasser & Taylor 2004; Dickey et al. 2009), which is also in good agreement with the theoretical models by Wolfire et al. (2003) at the Galactocentric distance $R_{\rm GC} = 11$ kpc. This conservative estimate gives a CNM fraction in the range $30\% < f_{\rm CNM} < 70\%$. Applying this finding to the whole cloud, Maggie should consequently be composed of a significant fraction of CNM.

We estimate the column density of the atomic hydrogen using (e.g., Wilson et al. 2013)

$$N_{\rm H\,I} = 1.8224 \times 10^{18} \int T_{\rm s}(v)\, \tau(v)\, {\rm d}v\,, \tag{5}$$

where $N_{\rm H\,I}$ is the total column density of H I in units of cm^{-2} as a function of the H I spin temperature $T_{\rm s}(v)$ and optical depth $\tau(v)$ integrated over the velocity v. The spin temperature is $T_{\rm s}(v) = T_{\rm B}/(1 - e^{-\tau(v)})$, where $T_{\rm B}$ is the brightness temperature of the H I emission. The optical depth correction of the column density scales as $\tau/[1 - \exp(-\tau)]$, resulting in a correction factor of 1.17 for $\tau_{\rm mean} = 0.33$. We integrated the column density from −57.5 to −48.5 km s^{-1}, taking into account the predominant velocities of Maggie. The column density then has values in the range $3{-}10 \times 10^{20}$ cm^{-2}. We used the observed THOR-H I emission data instead of the individual Maggie component of the decomposition to compute the column density although there might be additional contributions from components blending in at a similar velocity. We show in Appendix A that the integrated

[5] The optical depth measurement against G43.921-0.479 is close to the detection limit and will not be discussed further here. However, the mean optical depth $\tau_{\rm mean} = 0.30$ and velocity-integrated optical depth $\int \tau(v){\rm d}v = 2.2$ km s^{-1} are consistent with the other source G43.738-0.620.

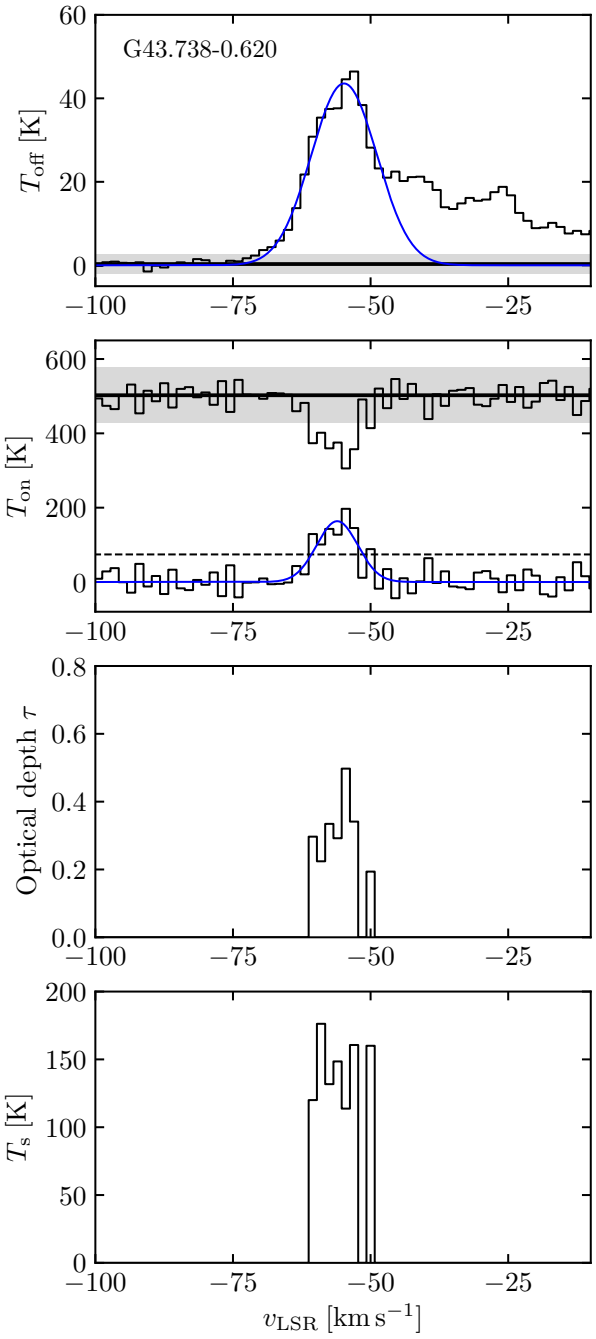

Fig. 10. H I emission and absorption spectrum of the extragalactic continuum source G43.738-0.620. *Top panel*: emission spectrum and is measured in an annulus around the source with inner and outer radii of 60″ and 120″, respectively (corresponding to 3−6 beams of the H I data). *Second panel*: absorption spectrum toward the point source. The lower spectrum indicates the inverted absorption spectrum along with the 3σ noise (dashed line). In the *first two panels*, the gray shaded area and blue curve indicate the 3σ noise and a single-component Gaussian fit, respectively. The *third panel* shows the optical depth computed using Eq. (2). In the *bottom panel* we present the spin temperature, which is computed using Eq. (3).

Maggie component is not spatially coherent in amplitude and shows patches of lower amplitude, likely owing to the existence of blended components that are difficult for GaussPy+ to disentangle. To estimate a maximum fraction of unrelated blended emission contributing to the column density, we compare the two column density maps. We find that the column density (and ultimately mass) inferred from the single Maggie component (Appendix A) is smaller by approximately 30%.

We note that choosing a single fit component as the "filament" can be problematic as the multiphase nature of H I emission may have multiple components at similar velocities contributing to Maggie. Here, we define Maggie by the centroid velocity of a single component and do not impose any restrictions on the H I phase of Maggie, which is reasonable in order to study the velocity structure. Yet, by choosing a single component we might include only one H I phase. GaussPy+ does not explicitly check for spatial coherence in amplitude or line width but for coherence in number of components and their centroid velocity. As we argue above, in a CNM-WNM mixture of the gas it is then difficult to characterize Maggie based on the centroid velocity of a single component alone. One way to fit and identify multiple phases in H I emission spectra, is to introduce spatial regularization terms to the loss function of the fitting that are able to cluster different phases (here: CNM and WNM) even if close in velocity (Marchal et al. 2019). This approach, however, is computationally expensive and will not be followed here.

The picture of a well-defined CNM phase layer within a WNM envelope assuming thermal pressure equilibrium might be overly simplified. Instead, we would expect a clumpy CNM embedded in a 'sea' of WNM, implying that the CNM has a lower volume and area filling factor than the WNM and a strong density contrast (Kalberla & Kerp 2009). The patchy structure seen in Fig. A.1 could be a reflection of the clumpy structure of the CNM. Lower intensity in small-scale structures could also hint at CNM that has already converted to molecular gas (see Sect. 4.4). If we want to infer the global multiphase column density of Maggie, however, it is beneficial to examine the properties of the filament by means of the integrated emission rather than using a single GaussPy+ fit component, in order to avoid missing mass.

Figure 11 shows the mass surface density map of Maggie derived from the observed emission data. The mass surface density reaches values of more than ~8.0 $M_\odot$ pc^{-2} (= 1.0 × 10^{21} cm^{-2}, $A_{\rm V} \sim 1$ mag)[6] toward the eastern hub and its distribution is constrained to within one order of magnitude (see Sect. 4.3). Wolfire et al. (2003) predict an average midplane thermal pressure at the Galactocentric radius $R_{\rm GC}$ = 12 kpc of $P_{\rm th,ave}/k \sim 1580$ K cm^{-3}. We can estimate the expected length scale of Maggie along the line of sight by $l_{\rm los} \sim N_{\rm H\,I}\, T_{\rm k}/(P_{\rm th,ave}/k)$, where we set the kinetic temperature to be close to the density-weighted mean spin temperature $T_{\rm k} \approx \langle T_{\rm s}\rangle$. Since Maggie is significantly below the midplane, the pressure is likely to be smaller than 1580 K cm^{-3}. Thus, we assume a canonical pressure of $P_{\rm th,ave}/k \sim 1000$ K cm^{-3}. At the mean column density of $\langle N_{\rm H\,I}\rangle = 4.8 \times 10^{20}$ cm^{-2} (see also Sect. 4.3), the line-of-sight length is estimated to be ~22 pc, which is roughly consistent with the on-sky width of 40 pc given the uncertainty in temperature and pressure.

Because of the diffuse nature of the WNM that is significantly contributing to the observed H I emission it is difficult to define a clear edge of the filament. We therefore estimate two different masses: a "diffuse mass" in which the dense filament is embedded (mass above 5σ), and a "dense filament mass" which resides within an approximately closed contour at the 9σ level

[6] We relate the total hydrogen column density $N_{\rm H} = N_{\rm HI} + 2\,N_{\rm H_2}$ to the visual extinction using $N_{\rm H} = 2.2 \times 10^{21}$ cm^{-2} mag^{-1} × $A_{\rm V}$ (Güver & Özel 2009). We take into account the molecular column density using the upper limit estimated in Sect. 4.1.

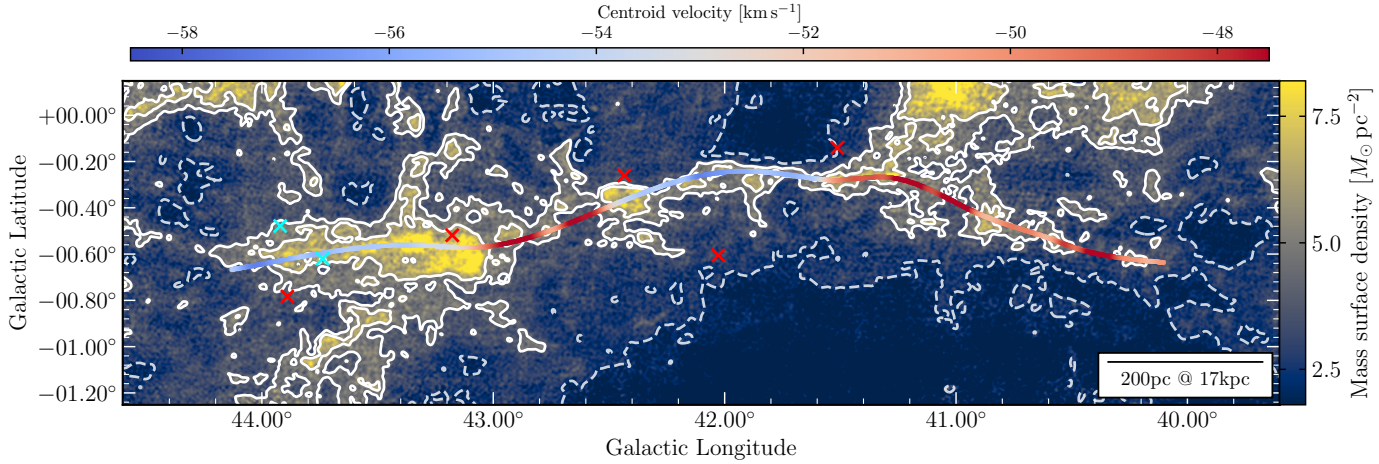

Fig. 11. Mass surface density map. The map shows the H I mass surface density integrated over the velocity interval $-57.5 < v_{\rm LSR} < -48.5$ km s^{-1}. The dashed contour gives the 5σ level at 2.3 $M_\odot$ pc^{-2} by which the diffuse mass is calculated. The white contours show the mass surface densities at the levels 4.2 and 5.6 $M_\odot$ pc^{-2}, respectively. The spine of Maggie is color-coded by the centroid velocity shown in Fig. 9. The crosses mark the positions of the continuum sources listed in Table 1. The cyan crosses indicate the positions of absorption detections.

(=4.2 $M_\odot$ pc^{-2}) that roughly corresponds to the filament width obtained in Sect. 3.1.

The total diffuse mass including the dense filament is 1.7×10^6 $M_\odot$. As mentioned above, due to its diffuse nature the mass derivation of the H I has a high uncertainty as it is dependent on the selection of filament regions taken into account. Varying the contour level by $\pm 1\sigma$ to estimate the uncertainty, the dense filament mass is $7.2^{+2.5}_{-1.9} \times 10^5$ $M_\odot$. The atomic mass of Maggie therefore is comparable to the highest masses of large-scale filaments identified in the Milky Way (Ragan et al. 2014; Abreu-Vicente et al. 2016; Wang et al. 2015, 2016; Zucker et al. 2015, 2018), providing a large atomic gas reservoir to potentially host molecular cloud formation.

If we take the column density we use to define the edge of the "dense" filament and assume that the width of the filament is the same along the line of sight as in the plane of the sky, we obtain a density of 4.2/40 $M_\odot$ pc^{-3} $\sim 7 \times 10^{-24}$ g cm^{-3}, which equates to an H nucleus number density of roughly 4 cm^{-3}. This is intermediate between the average densities we expect for the WNM and CNM at this Galactocentric radius (Wolfire et al. 2003), which is best explained by the filament being a mix of both components.

However, given a mean density of $\bar{n} = 4$ cm^{-3}, we point out that either the CNM mass fraction $f_{\rm CNM}$ must be higher or that a substantial fraction of the WNM must be in an unstable phase to be in agreement with the theoretical model given in Wolfire et al. (2003). This follows from the mass conservation:

$$\bar{n} = f_{\rm V,CNM}\, n_{\rm CNM} + (1 - f_{\rm V,CNM})\, \gamma\, n_{\rm CNM}, \quad (6)$$

where $f_{\rm V,CNM}$ is the volume filling fraction of the CNM, $n_{\rm CNM}$ is the CNM number density, and $\gamma = n_{\rm WNM}/n_{\rm CNM} = T_{\rm CNM}/T_{\rm WNM} \sim 0.01$ is the ratio of the CNM and WNM temperature in pressure equilibrium. The CNM mass fraction $f_{\rm CNM} = f_{\rm V,CNM}\, n_{\rm CNM}/\bar{n}$ can then be expressed as

$$f_{\rm CNM} = \frac{1}{1-\gamma}\left(1 - \gamma\, \frac{n_{\rm CNM}}{\bar{n}}\right). \quad (7)$$

Assuming the Wolfire et al. (2003) CNM-WNM model, with $\bar{n} = 4$ cm^{-3} the CNM mass fraction is higher than our assumed values. For example, even for a CNM density $n_{\rm CNM} = 40$ cm^{-3} (the maximum CNM density at $R_{\rm GC} = 11$ kpc; Wolfire et al. 2003), the CNM fraction is still $f_{\rm CNM} = 0.91$. The mean density can be brought into agreement with the previously derived CNM fraction of ~0.5 if the temperature ratio becomes $\gamma \sim 0.05$, which corresponds to a mean temperature in the unstable regime ~1000–2000 K, compared to the higher temperatures of the classical WNM around ~6000 K. Observational works have found the fraction of H I in the unstable phase to be at least 20–50% (Heiles & Troland 2003; Roy et al. 2013b; Murray et al. 2015, 2018; Nguyen et al. 2019), possibly even higher (Koley & Roy 2019).

4. Discussion

4.1. Molecular gas tracers, continuum, and dust

We investigate the molecular gas fraction of Maggie and possible signatures of molecular cloud formation by examining the molecular lines ^{12}CO, ^{13}CO, and C^{18}O (J = 1–0) using the Milky Way Imaging Scroll Painting (MWISP; Su et al. 2019) survey. The rms noise of the data is 0.5 K for ^{12}CO at the velocity resolution of 0.16 km s^{-1} and 0.3 K for ^{13}CO and C^{18}O at 0.17 km s^{-1} resolution. The spatial resolution in all three tracers is ~50″.

In all three tracers, we detect no molecular gas emission in the velocity range between −57.5 and −48.5 km s^{-1}. We show in Appendix B a map of the integrated ^{12}CO emission as an example. To estimate the molecular hydrogen content, we used the CO-H_2 conversion factor $X = 1.8 \times 10^{20}$ cm^{-2} K^{-1} km^{-1} s (Dame et al. 2001). We determined an upper limit at 5σ for the molecular surface density of $\Sigma_{\rm H_2} \sim 13$ $M_\odot$ pc^{-2}. Since the sensitivity limit is close to the observed H I column densities, we cannot entirely dismiss the presence of CO-bright H_2 gas.

Indeed, Nakanishi et al. (2020) identify molecular gas clumps toward Maggie within small-scale H I clouds that they infer from a dendrogram analysis of the VGPS H I data (Stil et al. 2006). The average size of the H I clouds is 5 pc. They derive the molecular mass by summing the brightness temperature of the FUGIN ^{12}CO data (Umemoto et al. 2017) within each H I cloud, including voxels with a brightness temperature lower than the noise. In doing that, diffuse emission can be recovered in the same manner as stacking emission spectra. The total H_2 mass of all 353 identified cloudlets that are within the dense filament

contour at 4.2 $M_\odot$ pc^{-2} (area filling factor ~0.21) is ~6×10^4 $M_\odot$, only a small fraction of the H I mass (0.08), with the cloudlets' average surface density being ~2.5 $M_\odot$ pc^{-2} that is consistent with our upper limit. The mean molecular mass fraction of the H I clouds is ~0.2 of the total gas mass (Nakanishi et al. 2020). Naturally, the H I "leaves" of the dendrogram analysis sample the highest-density parcels of the medium, hence a portion of the total molecular mass outside the H I clouds might be missed.

In addition, CO might not always be a good tracer of molecular hydrogen as "CO-dark" H_2 could account for a significant fraction of the total H_2 (Pineda et al. 2008; Goodman et al. 2009; Tang et al. 2016), particularly at low column densities (Goldsmith et al. 2008; Planck Collaboration XIX 2011). The CO-dark fraction has furthermore been observed to increase with Galactocentric distance (Pineda et al. 2013). Simulations of molecular gas in Milky Way type galaxies indicate that the CO-dark fraction is indeed higher in regions of lower surface density (Smith et al. 2014a). CO observations at visual extinctions $A_V \leq 3.5$ could underestimate the molecular hydrogen content by a significant fraction (Glover & Mac Low 2011).

As OH may be a better molecular gas tracer than CO in atomic-molecular transition regions at low column densities (e.g., Tang et al. 2017), we inspected the strongest OH line at 1667 MHz in the THOR OH data (see Rugel et al. 2018). The THOR OH data consist of VLA C-array observations only, so large-scale emission cannot be recovered due to the missing flux on short *uv* spacings. However, we can derive an OH abundance from OH absorption measurements against continuum sources. At the velocity range of Maggie, no absorption is found against any of the continuum sources listed in Table 1. We note that depending on the peak flux of the continuum source and opacity of the OH line, the detection limit could be above the expected OH column density. The presence of molecular gas traced by OH can therefore not be ruled out. Follow-up OH observations at higher sensitivity will address this in a future study.

We investigated the possibility of Maggie being a foreground structure by examining the extinction of stars up to a distance of 5 kpc using the early installment of the third *Gaia* data release (Gaia Collaboration 2021). The K_s-band extinctions are obtained using the Rayleigh-Jeans Colour Excess (RJCE, Majewski et al. 2011) method which combines near- and mid- infrared data from the Two Micron All-Sky Survey (2MASS; Skrutskie et al. 2006) and the Wide-Field Infrared Survey Explorer (WISE; Wright et al. 2010), respectively. The *Gaia* Archive[7] provides the crossmatch of *Gaia* sources with both the 2MASS and WISE catalogs.

In the presence of an obscuring medium, for example an atomic cloud, high extinction in the densest parts would prevent the detection of stars in the *Gaia* sample, yet extinction toward the outer portions of the cloud can still be detected in order to trace its silhouette. Figure C.1 presents the obtained K_s-band extinctions toward Maggie for sources up to a distance of 5 kpc from us. If Maggie were located within 5 kpc from the Sun, higher extinction stars would represent its shape in the ℓ-b plane. However, since no signature of Maggie emerges from the extinction map, this indicates that Maggie is not a foreground cloud with large deviation from the Galactic rotation velocity.

We furthermore detect neither extinction features in the Spitzer Galactic Legacy Infrared Mid-Plane Survey Extraordinaire (GLIMPSE; Churchwell et al. 2009) that might hint at the presence of IRDCs nor any excess infrared emission or signatures of stellar activity as observed in all bands of GLIMPSE and the Improved Reprocessing of the IRAS Survey (IRIS; Miville-Deschênes & Lagache 2005). There are two possible reasons for not detecting IRDC features: (1) Maggie has not formed cold, dense molecular regions on a large scale such that there exist no IRDCs within Maggie or (2) Maggie is too distant to be observed as an IRDC due to the lack of emission background.

We have also examined the high-sensitivity maps of *Planck* at 353, 545, and 857 GHz emission (Planck Collaboration I 2014) and the higher angular resolution *Herschel* Hi-GAL data (Marsh et al. 2017) to further search for signatures of Maggie. We do not detect Maggie in any of the continuum bands while noting that we are most likely dominated by foreground emission in the Galactic plane.

After taking a variety of surveys into consideration, Maggie appears to show no signs of stellar activity as it is mostly atomic, while molecular gas has formed only in high-density clumps on the smallest spatial scales (Nakanishi et al. 2020).

4.2. Kinematic signatures

4.2.1. Mach number distribution

In the following, we adopt a stable two-phase CNM-WNM model to derive the turbulent Mach number distributions. We estimate the three-dimensional scale-dependent Mach number of the filament assuming isotropic turbulence with $\mathcal{M} = \sqrt{3}\,\sigma_{\rm turb}/c_{\rm s}$, where $\sigma_{\rm turb}$ and $c_{\rm s}$ are the turbulent one-dimensional velocity dispersion and sound speed, respectively. The turbulent line width is calculated by subtracting the thermal line width contribution from the observed line width as

$$\sigma_{\rm turb} = \sqrt{\sigma^2_{\rm obs} - \sigma^2_{\rm th}}, \tag{8}$$

where $\sigma_{\rm obs}$, and $\sigma_{\rm th}$ are the observed, and thermal velocity dispersion, respectively. Since we expect the observed line width to be a combination of both CNM and WNM, and do not know the turbulent contribution to the observed line width of either H I phase, we compute a single Mach number for the two-phase medium using a weighted mean of the thermal line width and sound speed. The thermal velocity dispersion can be decomposed into a CNM and WNM component as

$$\sigma^2_{\rm th} = f_{\rm CNM}\,\sigma^2_{\rm th,CNM} + (1 - f_{\rm CNM})\,\sigma^2_{\rm th,WNM}, \tag{9}$$

where $\sigma_{\rm th,CNM}$ and $\sigma_{\rm th,WNM}$ refer to the thermal velocity dispersion of the CNM and WNM, respectively. The combined velocity dispersion and sound speed can then be computed using the weighted mean kinetic temperature

$$\langle T_{\rm k} \rangle = f_{\rm CNM}\,T_{\rm k,CNM} + (1 - f_{\rm CNM})\,T_{\rm k,WNM}, \tag{10}$$

where $T_{\rm k,CNM}$ and $T_{\rm k,WNM}$ are the kinetic temperatures of the CNM and WNM, respectively.

We assume that the kinetic temperature of the CNM is close to the spin temperature and set $T_{\rm s,CNM} = T_{\rm k,CNM}$. In pressure equilibrium, Wolfire et al. (2003) predict a range of WNM kinetic temperatures between 4100–8800 K. As the densities of the WNM are low, the gas can usually not be thermalized by collisional excitation only, such that $T_{\rm s,WNM} < T_{\rm k,WNM}$. Through excitation by resonant Lyα photon scattering, sufficient Lyα radiation can allow $T_{\rm s,WNM} \rightarrow T_{\rm k,WNM}$. The WNM spin temperature $T_{\rm s}$ has been observed to be as high as 7200 K (Murray et al. 2014). We therefore assume a mean WNM kinetic temperature of $T_{\rm k,WNM} = 6000$ K.

[7] https://gea.esac.esa.int/archive/

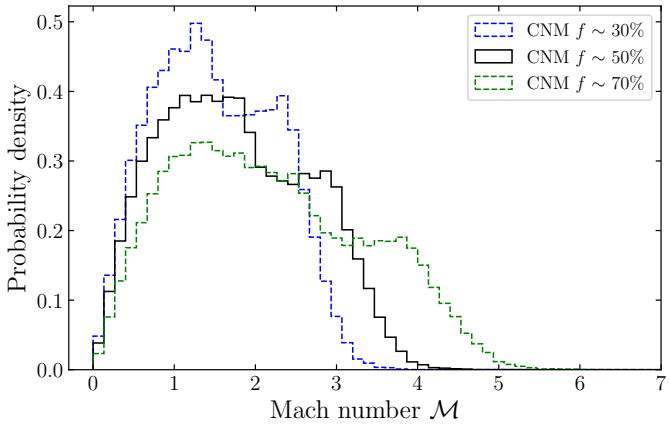

Fig. 12. Mach number distribution of Maggie. The blue, black, and green histograms show the Mach number distributions assuming a CNM fraction of 30, 50, and 70%, respectively. Each distribution is computed using a single thermal line width and sound speed inferred from a weighted mean kinetic temperature.

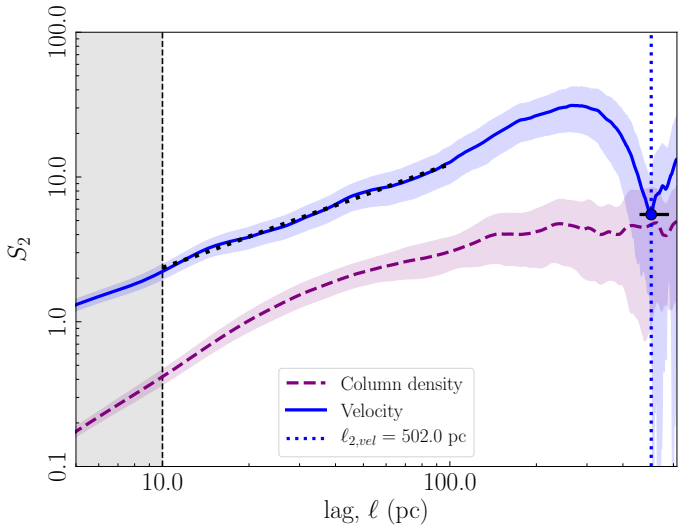

Fig. 13. Second order structure function. This plot shows the second order structure function of column density (dashed purple) and velocity (blue) computed from the respective average quantity along the Maggie filament. A 2D plane fit to the velocity field has been subtracted prior to computing the velocity structure function. The vertical blue dotted line indicates the characteristic spatial scale of the velocity. The black dotted line is a power-law fit to the velocity structure function. The gray shaded area indicates the range of scales below three beams for which the recovered structure function is unreliable.

Since the thermal line width and sound speed scale as $T_{\rm k}^{1/2}$, the variation is moderate and does not change the Mach number significantly. We calculate three different Mach number distributions for $f_{\rm CNM}$ = 30, 50, 70%, corresponding to $T_{\rm k,CNM}$ = 40, 70, 100 K.

Figure 12 shows the Mach number distributions of Maggie under different assumptions of the CNM fraction. The median Mach number is 1.5, 1.7, and 2.0 for the CNM fractions 30, 50, and 70%, respectively, which is generally in contrast to molecular filaments being driven by highly supersonic turbulence (e.g., Elmegreen & Scalo 2004). However, since we use a single mean kinetic temperature in each case, the Mach number distributions also reflect a slight bimodality that is already evident in the line width distribution (Fig. 8). While each Mach number distribution represents a different mixture of CNM-WNM gas, we note that the bump toward higher Mach numbers reflected in each distribution is due to the increased line widths associated with higher $v_{\rm LSR}$ velocities, likely contaminated by blending background gas (see Sect. 3.2).

Numerical studies show that the condensation process toward a stable CNM phase depends on the properties of turbulence in the WNM (Seifried et al. 2011; Saury et al. 2014; Bellomi et al. 2020). These studies suggest that highly supersonic turbulence that is driven by compressive modes regulates the formation of a stable CNM phase as turbulent compression and decompression of the gas will be much more rapid than the cooling timescale. Once the gas has moved toward an unstable phase, that is the cooling timescale $t_{\rm cool}$ is shorter than the dynamical time $t_{\rm dyn} = \ell/c_s$, the gas can cool down fast enough to reach a stable branch of the CNM before being decompressed again (Hennebelle & Pérault 1999). On the other hand, a too low degree of turbulence cannot produce sufficient density fluctuations to sustain the formation of a thermally stable CNM. The WNM is mostly found in a subsonic and transonic regime (e.g., Marchal & Miville-Deschênes 2021), therefore providing the dynamical conditions for thermal instability to grow efficiently to form a bistable H I phase.

Absorption studies show that a substantial fraction of H I in fact lies in a thermally unstable or transient phase (Heiles 2001; Heiles & Troland 2003; Roy et al. 2013a,b). Shifting the mean temperature of the WNM to an unstable regime around ~1500 K leads to Mach numbers in the range 2–4.

4.2.2. Structure function

In order to investigate the velocity and column density structure of Maggie in a statistical sense, we compute the 1D structure function (see Henshaw et al. 2020) of the data using

$$S_p(\ell) = \langle \delta x(r,\ell)^p \rangle = \langle |x(r) - x(r+\ell)|^p \rangle, \quad (11)$$

where $\delta x(r,\ell)$ is the absolute difference in the quantity x measured between two locations r and $r+\ell$ separated by ℓ. The brackets indicate the average of the quantity over all locations. The order p of the structure function can be directly related to a physical quantity. We use the second order structure function, which in the case of the velocity structure is proportional to the kinetic energy of the line-of-sight component of the flow. We show in Fig. 13 the second order structure function derived from the weighted centroid velocity (Fig. 9), after having subtracted a 2D plane fit to the velocity field in order to better examine the wave-like oscillation, and the column density along the 40 pc-wide spine of Maggie. A significant local minimum is evident at the spatial scale of $\ell_{2,\rm vel} \approx 500$ pc. Since the structure function is evaluated for pairs of positions that are separated by the lag ℓ, the maximum spatial scale at which the structure function can be fully sampled is given by half the total length of the filament. Therefore, the structure function is only computed over spatial scales $\ell \lesssim 600$ pc. Consequently, the full velocity 'dip' cannot be recovered and we estimate the uncertainty of the scale by the difference between $\ell_{\rm max}$ and $\ell_{2,\rm vel}$, giving $\ell_{2,\rm vel} = 502 \pm 113$ pc. This result confirms the wave-like pattern seen in Figs. 3 and 9.

The column density structure of Maggie does not show any characteristic length scale at which the structure function has a local minimum but roughly represents a power-law scaling, indicative of the scale-invariant behavior of a gas structure generated by turbulence (Elmegreen & Scalo 2004).

The phenomenological Kolmogorov relation for incompressible turbulence predicts a dissipationless cascade of energy in

the inertial subrange, such that the kinetic energy transfer rate $\mathcal{E} \approx \delta v^3/\ell$ stays constant (Kolmogorov 1941). Due to the nature of the turbulent cascade, the velocity dispersion then increases with scale as $\delta v \propto \ell^{q_{\rm theo}}$, where $q_{\rm theo} = 1/3$, which we can relate to the power-law scaling of the second order structure function $\langle \delta v^2 \rangle \propto \ell^{\gamma}$, where $\gamma = 2q_{\rm theo}$ (see e.g., Elmegreen & Scalo 2004; Henshaw et al. 2020). We therefore fit the velocity structure function with a power-law function of the form $S_2(\ell) \propto \ell^{\gamma}$. We set the lower limit of the fitting range to three beam sizes (10 pc), below which no reliable structure functions can be recovered, and the upper limit at the approximate scale above which the structure function begins to turn over (100 pc). For Maggie we find $q = 0.36 \pm 0.02$ in the case where we subtracted the 2D plane fit, which is in good agreement with the expected power-law scaling $q_{\rm theo} = 1/3$ in the subsonic and transonic regime (see also Wolfire et al. 2003; Elmegreen & Scalo 2004; McKee & Ostriker 2007; Saury et al. 2014; Marchal & Miville-Deschênes 2021). While the dip in the velocity structure function is less pronounced if we do not subtract a 2D plane fit, the power-law scaling remains the same with $q = 0.37 \pm 0.02$. These values are in agreement with the observational scaling exponents originally found by Larson (1979, 1981), who investigated the velocity dispersion–size relation in a sample of molecular clouds and clumps using different tracers. He utilized the radial velocity dispersion mainly in terms of the line width of optically thin lines, such that the relation between line width and individual cloud size can be interpreted as a correlation between velocity fluctuation and scale size (see also Scalo 1984). Larson (1979, 1981) found a velocity dispersion–size relation following a power law with an index $q_{\rm L} = 0.38$, which is, although slightly larger, similar to the Kolmogorov relation. Systematic studies of more homogeneous samples of molecular clouds traced in CO found velocity scaling exponents in a range between 0.24 and 0.79 (see Table 2 in Izquierdo et al. 2021, which summarizes more than 40 years of observational work).

Izquierdo et al. (2021) find in their simulations that the scaling exponents lean toward lower values when molecular clouds are dominated by the influence of the Galactic potential rather than the effects of clustered SN feedback and self-gravity (see also Chira et al. 2019). This suggests that low velocity scaling exponents are associated with diffuse regions of low star forming activity (e.g., Federrath et al. 2010), which also favors the development of long filamentary structures (Smith et al. 2020).

Maggie's velocity scaling is in better agreement with the lower end of the values found for molecular clouds (see above). This regime around 1/3 is indicative of Kolmogorov turbulence, and following Izquierdo et al. (2021), this turbulence may be caused by the Galactic potential rather than being driven by the effects of self-gravity or stellar feedback.

The existence of density enhancements that are spaced at a characteristic scale within a highly filamentary cloud would be in broad agreement with theoretical predictions of the fragmentation of a self-gravitating fluid cylinder due to the "sausage" instability (e.g., Chandrasekhar & Fermi 1953). As we find no characteristic scale in the column density structure, we argue that the effects of self-gravity have not yet become important and it is difficult to correlate the observed velocity signature with possible mechanisms driving the column density structure formation.

We caution, as pointed out in Sect. 3.3, that the column density is a product of both the CNM and WNM and is inferred from the integration over velocity. This could lead to a structure that smooths over an underlying higher-density CNM structure due to the diffuse nature of the WNM. The column density structure can furthermore not be described by single power-law alone. This could be due to the filament being a mix of both CNM and WNM and that the two phases have different turbulence scalings. With the WNM being more widespread, the structure function as probed by H I emission could be dominated by the warm diffuse gas on large scales while the CNM dominates the power-law on small scales (see e.g., Choudhuri & Roy 2019). The integrated emission might therefore not be the best estimator of the spatial distribution of the column density.

4.3. Column density PDF

To elaborate on the dominant physical processes acting within the cloud, we employed the commonly used probability density function (PDF) of the column density. The shape of column or volume density PDFs are used as a means to describe the underlying physical mechanisms of a cloud (e.g., Federrath & Klessen 2013; Kainulainen et al. 2014; Schneider et al. 2015). A log-normal shape of a PDF is considered a signature of turbulent motion dominating a cloud's structure (see e.g., Klessen 2000). The magnitude of turbulence can furthermore be reflected in the width of a log-normal PDF and can then be associated with the Mach number (e.g., Padoan et al. 1997; Passot & Vázquez-Semadeni 1998; Kritsuk et al. 2007; Federrath et al. 2008; Konstandin et al. 2012; Molina et al. 2012), while noting that the turbulence driving scale and CNM-WNM mass ratio also affect the width of the PDF (Bialy et al. 2017b).

Molecular clouds that are subject to the increasing effect of self-gravity develop high-density regions, producing a power-law tail in their PDF (e.g., Girichidis et al. 2014; Burkhart et al. 2017). Many star-forming molecular clouds have been confirmed to show such power-law tails (Kainulainen et al. 2009; Schneider et al. 2013, 2016). Even before the effects of gravity become dominant, gravitationally unbound clumps can exhibit power-law tails due to pressure confinement from the surrounding medium (Kainulainen et al. 2011).

We show in Fig. 14 the normalized column density PDF (N-PDF) of Maggie taken above the 5σ limit at $\sim 3 \times 10^{20}$ cm^{-2} (=2.3 $M_\odot$ pc^{-2}). The mean of the N-PDF is 4.8×10^{20} cm^{-2}. The distribution varies within one order of magnitude and can be well described by a log-normal function of width $\sigma = 0.28$. We caution that the shape and width of the resulting N-PDFs are sensitive to the region taken into account (Schneider et al. 2015; Lombardi et al. 2015; Ossenkopf-Okada et al. 2016; Körtgen et al. 2019). If we only consider column densities within an approximately last closed contour at $\sim 4 \times 10^{20}$ cm^{-2}, the N-PDF shows a narrower log-normal shape with a width smaller by $\Delta\sigma = 0.05$.

Figure 14 presents a comparison between the N-PDF of Maggie and the optical depth corrected N-PDFs derived from H I emission and self-absorption toward the giant molecular filament GMF20.0-17.9 (Syed et al. 2020). The width of Maggie's N-PDF comes close to the value found for the CNM gas traced by HISA and is larger than that of the atomic gas traced by H I emission toward the midplane (see also Wang et al. 2020c). However, the width of the CNM distribution is likely to be higher as the HISA N-PDF is limited by the column density range at which H I self-absorption can be detected.

Since we use the same data set and follow the same methodology for deriving the N-PDF as Wang et al. (2020c) and Syed et al. (2020), systematic differences should be minimized. Despite that, we note a slight cutoff in the PDF at $N_{\rm H\,I} \sim 1 \times 10^{21}$ cm^{-2}. As we use a uniform optical depth correction for the whole cloud (Sect. 3.3), the optical depth could be underestimated in regions of high column density. If we assume a

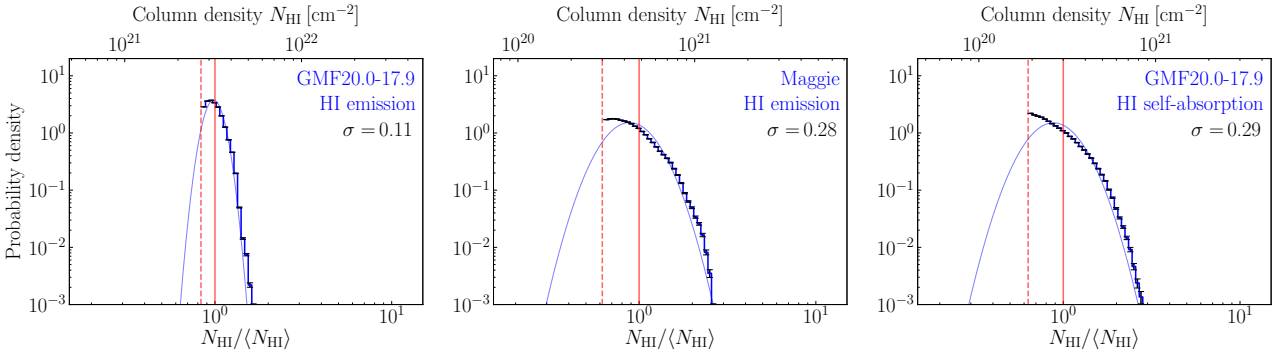

Fig. 14. Comparison of N-PDFs. The blue histograms show the column density distributions of Maggie (*middle panel*) in comparison with the N-PDFs taken from a different GMF region analyzed in Syed et al. (2020) that trace the midplane H I emission (*left*) and CNM by means of H I self-absorption (*right*). The blue curve in each panel indicates a log-normal fit to the distribution. The red vertical dashed and solid lines mark the column density thresholds taken into account and the mean column density, respectively.

column density $\sim 10^{21}$ cm^{-2} and a FWHM line width ~ 10 km s^{-1}, the peak optical depth becomes $\tau \sim 1$. The cutoff in the PDF could therefore be due to the cloud becoming optically thick and the width could then be underestimated.

Narrow log-normal shaped N-PDFs are commonly observed in the diffuse H I emission toward well-known molecular clouds (Burkhart et al. 2015; Imara & Burkhart 2016; Rebolledo et al. 2017). On the other hand, the HISA N-PDFs that trace the CNM show similar or somewhat broader distributions, indicative of the clumpy structure and higher degree of turbulence. The CNM-WNM mixture of Maggie, that is spatially more clearly defined than the material usually traced by H I emission, agrees with our previous findings and suggests that Maggie might be at the verge of becoming a dense filament forming out of the diffuse atomic ISM.

4.4. Molecular gas formation and timescale

We address Maggie's current evolutionary stage and evaluate the surface density threshold of H I required to efficiently form molecular hydrogen, based on the analytical models by Sternberg et al. (2014), and Bialy & Sternberg (2016). These authors derived a steady-state H I-to-H_2 transition model assuming a slab irradiated isotropically on two sides by FUV flux. The expected H I surface density at the transition then is

$$\Sigma_{\rm H\,I} = \frac{6.71}{\tilde{\sigma}_g} \ln\left(\frac{\alpha G}{3.2} + 1\right) M_\odot\,{\rm pc}^{-2} \tag{12}$$

(see Eq. (1) in Bialy et al. 2017a), where $\tilde{\sigma}_g$ is the dust absorption cross section per hydrogen nucleus in the Lyman-Werner (LW) dissociation band (11.2–13.6 eV) relative to the standard solar neighborhood value, α is the ratio of the unshielded H_2 dissociation rate to formation rate, and G is the average H_2 self-shielding factor in dusty clouds. The product αG is the H_2 formation-to-destruction rate ratio, accounting for H_2 shielding by H_2 line absorption and dust, and may be written as

$$\alpha G = 2.0\, I_{\rm UV} \left(\frac{30\,{\rm cm}^{-3}}{n}\right)\left(\frac{9.9}{1 + 8.9\,\tilde{\sigma}_g}\right)^{0.37} \tag{13}$$

(Sternberg et al. 2014; Bialy & Sternberg 2016), where $I_{\rm UV}$ is the free-space (unshielded) interstellar radiation intensity relative to the Draine (1978) field, and n is the number density. The UV intensity exponentially decreases with Galactocentric distance (Wolfire et al. 2003). For $R_{\rm GC} = 12$ kpc, Wolfire et al. (2003) suggest $I_{\rm UV} \sim 0.39$. We note that this approximation is derived for the midplane and should be roughly constant up to the scale height of OB stars and could be significantly lower at larger distances from the plane. Since Maggie's distance from the plane is an order of magnitude greater than the OB scale height in the solar neighborhood (Reed 2000), the UV intensity should be viewed as an upper limit with a relative uncertainty of at least 40% (see also Elias et al. 2006).

The dust absorption cross-section $\tilde{\sigma}_g$ can be estimated as the inverse of the gas-to-dust ratio at the position of Maggie relative to the value in the local ISM if we assume that the dust grain properties are the same in the outer Galaxy. Giannetti et al. (2017) derive a gas-to-dust ratio γ as a function of galactocentric distance and metallicity based on $C^{18}O$ observations. Their reported gas-to-dust gradient predicts $\gamma \sim 305$ for Maggie and $\sim$141 as a local value for $R_{\rm GC} = 8.15$ kpc (Reid et al. 2019). The relative dust absorption cross-section then becomes $\tilde{\sigma}_g = 0.46$. With the estimated number density of $\bar{n} = 4\,{\rm cm}^{-3}$ (see Sect. 3.3), the H I threshold for molecular hydrogen to form efficiently is expected at $\Sigma_{\rm H\,I} \approx 17.5\,M_\odot\,{\rm pc}^{-2}$. This exceeds the observed column densities by a factor of more than 2.

However, we note that the model by Sternberg et al. (2014) is highly idealized and the derived gas-to-dust relation in Giannetti et al. (2017) has large systematic uncertainties due to variations in the CO abundance and poorly constrained dust properties. In addition with the uncertainties in distance and UV intensity, the H I threshold is uncertain by at least $\sim$50%.

Furthermore, observational effects might produce a bias against the measurement of a clumpy multiphase medium. Small pockets of high-density CNM clumps could be hidden below the resolution of the telescope beam ($\approx$3.3 pc). Even if we assume that those unresolved CNM clumps have a small volume filling fraction $f_{\rm V,CNM}$, a substantial fraction of the mass $f_{\rm CNM} = (n_{\rm CNM} \cdot f_{\rm V,CNM})/\bar{n}$ could be allocated in the CNM phase although the observed mean density is low compared to that of the CNM. High-density CNM clumps are much more efficient at shielding and allow the onset of H_2 formation at lower surface densities. As we argue in Sect. 3.3, the CNM fraction should be considerable and could allow H_2 formation on a scale that is inaccessible to our observations.

In fact, Nakanishi et al. (2020) report a total molecular gas mass of $\sim 6 \times 10^4\,M_\odot$ that resides within small CNM cloudlets with an average size of 5 pc (see Sect. 4.1). These cloudlets have a mean H I density of $\sim$50 cm^{-3}. Inserting this density into Eqs. (13) and (12) gives a molecule formation threshold at $\Sigma_{\rm H\,I} \approx 2.5\,M_\odot\,{\rm pc}^{-2}$, which is below their observed H I surface density

of $\sim 7\,M_\odot\,\mathrm{pc}^{-2}$. Hence, we would expect molecule formation to take place within these high-density CNM cloudlets. Clearly, employing segmentation methods such as the dendrogram algorithm rather than using integrated quantities is important when aiming at identifying small-scale CNM structures in a clumpy multiphase medium.

We estimate the timescale of H_2 formation by the time after which the balance between H_2 destruction and formation rate has reached a steady state. If we consider the dissociation rate to be negligible compared to the formation rate, the H_2 formation timescale is given by

$$t_{\mathrm{H_2}} \sim \frac{1}{2\,R\,n} \tag{14}$$

(see Eq. (23) in Bialy & Sternberg 2016), where $R = 3 \times 10^{-17}\,\tilde{\sigma}_\mathrm{g}\,\mathrm{cm}^3\,\mathrm{s}^{-1}$ is the H_2 formation rate coefficient and n is the total hydrogen number density that is essentially given by our estimated H I number density in the absence of H_2. For a mean density of $\bar{n} = 4\,\mathrm{cm}^{-3}$ the formation timescale is $\sim$300 Myr, about an order of magnitude larger than those predicted by models of nearby molecular clouds (Goldsmith et al. 2007) and numerical simulations (Clark et al. 2012). As the formation of molecular hydrogen is most effective on the surface of dust grains (Gould & Salpeter 1963), the formation timescale at a given density is primarily limited by the dust-to-gas ratio $\propto \tilde{\sigma}_\mathrm{g}$, rendering the formation of molecular hydrogen less likely in the outer Galaxy. However, the effects of small-scale density fluctuations, either due to a clumpy CNM-WNM medium or due to transient structures produced by supersonic turbulence (Glover & Mac Low 2007), can shorten the timescale. For CNM clumps at several times the observed density, the timescale could decrease significantly. Again using the H I density reported by Nakanishi et al. (2020), the molecule formation timescale within the CNM-rich structures becomes $\sim$25 Myr.

In conclusion it is difficult to tightly constrain the formation of molecules within Maggie based on integrated H I and CO observations alone. Because of observational effects small-scale density fluctuations might be missed that allow the formation of molecular hydrogen even at the current stage. On the smallest scales within CNM-rich structures, the presence of molecular hydrogen traced by CO emission has been confirmed by Nakanishi et al. (2020). In addition, time delays between the formation of H_2 and CO could lead to Maggie not being identified as a molecular cloud, even though the formation of molecular hydrogen is already taking place on larger scale (Clark et al. 2012).

4.5. Comparison with literature

Maggie's origin remains unclear. While there are morphological and kinematic similarities, it is difficult to argue in favor of Maggie being an immediate counterpart to large-scale molecular filaments known in the literature.

The aspect ratio of Maggie is in rough agreement with large-scale filaments and Milky Way bones (Wang et al. 2015; Zucker et al. 2015), even though we point out that the determination of filament widths is in many cases subject to the tracer being used to identify the filament. Furthermore, Maggie has a smooth velocity gradient of less than $\pm 3\,\mathrm{km\,s}^{-1}\,(10\,\mathrm{pc})^{-1}$ (Sect. 3.2) and therefore satisfies the Milky Way bone criteria applied in Zucker et al. (2015).

Maggie does not display any close association with a spiral arm structure although the predicted location of the spiral arms is highly model-dependent (see e.g., Zucker et al. 2015). It is observed that all types of large-scale filaments identified in the literature commonly show a range of orientations with respect to spiral arm features in position-velocity (ℓ-v) space, both tracing out the spine of spiral arms as well as following ℓ-v tracks that are inclined to spiral arms (Zucker et al. 2018). The latter could potentially trace a spurious feature trailing off a spiral arm.

The orientation of Maggie's major axis is remarkably well aligned with the Galactic midplane, a fact that the majority of large-scale filaments have in common (Zucker et al. 2018). However, Maggie is observed at a large distance of $\approx$500 pc from the Galactic midplane while most filaments are found in spatial proximity to it. The fact that Maggie shows both a spatial (position-position) and kinematic (position-velocity) displacement from any spiral structure suggests a different formation path than those of known molecular filaments in the disk.

The Radcliffe wave (Alves et al. 2020) that consists of an association of molecular cloud complexes in the solar neighborhood has an aspect ratio of 20:1 and exhibits an undulating structure (in p-p-p space) below and above the midplane. At its peak, it reaches an offset of $\sim$200 pc from the plane. It is argued that the accretion of a tidally stretched gas cloud settling into the Galactic disk could in principle mimic the shape of the observed structure, but seems unlikely given that it would require the Radcliffe wave to be synchronized with the Galactic rotation over large scales.

Infalling gas, possibly of extragalactic origin, would interact with the gas in the disk and become compressed or even shocked along the edge closest to the disk if the relative velocities are large enough. In the case of Maggie, we then consider the chances of infalling gas becoming plane-parallel to be small. Additionally, in this scenario it seems unlikely for the gas to keep a coherent velocity structure.

Conversely, gas that is being expelled out of the plane as a result of SN feedback tends to have a randomized orientation with respect to the Galactic disk (Smith et al. 2020). Observations of the preferentially vertical H I filaments seen in the Galactic plane can be associated with enhanced SN feedback (Soler et al. 2020) and indicate that Maggie is unlikely to have originated in SN feedback events.

As Maggie could be linked to the global spiral arm structure in ℓ-v space, we speculate that the filament originated in the midplane and might have been brought to large distance by vertical oscillations. Levine et al. (2006) decomposed the structure of the Galactic H I disk into Fourier modes. They found that the Galactic warp (e.g., Burton 1988) can be described sufficiently well by the first three Fourier modes. Locally, significant high-frequency modes at large Galactocentric distances are also evident within the disk. This suggests that deviations from the midplane generally do occur but do not pervade the Galactic disk as they could be triggered by local perturbations.

Wobbles and disk corrugations could be induced by merger events or satellite galaxies plunging through the Galactic plane as suggested by simulations of the stellar disk of galaxies (e.g., Edelsohn & Elmegreen 1997; Widrow et al. 2014; Chequers & Widrow 2017). This hypothesis is supported by stellar populations that have formed in the stellar disk but are found in the Galactic halo at kiloparsec distances from the plane (Bergemann et al. 2018). Unsupervised machine learning measurements of stellar populations in the solar neighborhood reveal filamentary or string-like clusters that are oriented parallel to the Galactic disk, partly tens of parsecs above and below the plane (Kounkel & Covey 2019). Their filamentary shape is argued to be primordial and not an effect of tidal interaction, suggesting that the structure might be inherited from parental molecular clouds.

5. Conclusions

We have studied the atomic gas properties within the giant atomic filament Maggie. Kinematic information was obtained from the spectral decomposition of H I emission spectra using the automated Gaussian decomposition routine GAUSSPY+. The main results are summarized as follows:

1. Maggie is one of the largest coherent filaments identified in the Milky Way, detected solely in atomic gas. At a kinematic distance of ∼17 kpc, the projected length of 1.2 ± 0.1 kpc exceeds those of the largest molecular filaments identified to date. Maggie has an aspect ratio of 30:1, thus making its filamentary shape comparable to smaller-scale molecular filaments. Despite being parallel to the Galactic disk, Maggie has a distance of ≈500 pc from the midplane.
2. The centroid velocity information obtained from the spectral decomposition reveals an undulating velocity structure with a smooth gradient of less than $3\,\mathrm{km\,s^{-1}}\,(10\,\mathrm{pc})^{-1}$. Based on the kinematic information, Maggie could be remotely linked to the global spiral structure of the Milky Way as it is trailing the Outer Arm by $5-10\,\mathrm{km\,s^{-1}}$. There is a slight bimodality in the line width distribution that is attributed to the phase degeneracy of the H I emission and broad components blending in at similar v_{LSR} velocities.
3. Optical depth measurements against strong continuum sources indicate that Maggie is composed of a significant fraction of CNM. Detections and nondetections of absorption toward extragalactic sources and Galactic H II regions, respectively, suggest that the distance inference based on the kinematic information is reasonable and that Maggie is located on the far side of the Galaxy. K_{S}-band extinction obtained from 2MASS and WISE combined with *Gaia* parallaxes confirm that Maggie is not located within a distance of 5 kpc from us. After correcting for optical depth effects, the column density has a mean of $\langle N_{\mathrm{H\,I}}\rangle = 4.8\times10^{20}\,\mathrm{cm^{-2}}$. We estimate a "dense" filament mass of $M = 7.2^{+2.5}_{-1.9}\times10^{5}\,M_{\odot}$, providing a large atomic gas reservoir for molecular gas to form. If we assume radial symmetry about the major axis, the column densities equate to an average hydrogen number density of $\sim4\,\mathrm{cm^{-3}}$.
4. We find no molecular counterpart to Maggie as traced by integrated CO emission or dust. However, by summing diffuse CO emission voxels within H I cloudlets obtained with a dendrogram analysis, Nakanishi et al. (2020) confirm the presence of molecular gas on the smallest spatial scales. But as the molecular gas mass is only a few percent of the total mass budget of the filament, Maggie still appears to be in a predominantly atomic phase. Yet we note that CO might not be a good probe at low column densities and early evolutionary stages as the onset of molecular filaments becoming CO-bright could lag behind the formation of H_2. Although the formation timescale of molecular hydrogen in CNM-rich structures can be on the order of a few tens of Myr and Maggie exposes the onset of molecule formation in the highest-density pockets of H I, the formation of H_2 is not yet likely to be efficient in the more diffuse parts of the filament.
5. Kinematic signatures are investigated by means of the Mach number distribution and reveal transonic and moderately supersonic velocities. The velocity structure function exhibits a modest power-law scaling that is still consistent with the Kolmogorov scaling for subsonic turbulence and is at the lower end of commonly observed scaling exponents in molecular clouds. The column density PDF can be represented by a log-normal distribution with a moderate width of $\sigma = 0.28$, indicative of an intermediate CNM-WNM mix that is not dominated by the effects of gravitational contraction.
6. Given the large distance from the Galactic midplane while remaining well aligned with the disk, we speculate that Maggie could be the signature of Galactic disk oscillations in the vertical direction. Future simulations of atomic and molecular structure formation in spiral galaxies under the influence of vertical perturbations could help constrain this possible formation pathway.

Clearly, further work is needed to better understand the origins of Maggie. Higher-sensitivity CO observations as well as CO-dark molecular gas tracers will facilitate our understanding of this unique filament. While Maggie so far constitutes a case study, future studies of the entire THOR survey as well as other H I surveys will investigate the more general nature of such large atomic filamentary structures.

Acknowledgements. J.S., H.B., S.C.O.G., and R.S.K. acknowledge support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 138713538 – SFB 881 ("The Milky Way System", subprojects A01, B01, B02, and B08). H.B. and J.D.S. further acknowledge funding from the European Research Council under the Horizon 2020 Framework Program via the ERC Consolidator Grant CSF-648505. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 851435). S.B. acknowledges the Institute for Theory and Computations (ITC) at the Harvard-Smithsonian Center for Astrophysics for financial support. This research was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology, under contact with the National Aeronautics and Space Administration (80NM0018D0004). S.C.O.G. and R.S.K. also acknowledge financial support from the Heidelberg Cluster of Excellence STRUCTURES in the framework of Germany's Excellence Strategy (grant EXC-2181/1 - 390900948) and funding from the European Research Council via the ERC Synergy Grant ECOGAL (grant 855130). R.J.S. is funded by an STFC ERF (grant ST/N00485X/1). This work has made use of data from the European Space Agency (ESA) mission *Gaia* (`https://www.cosmos.esa.int/gaia`), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, `https://www.cosmos.esa.int/web/gaia/dpac/consortium`). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. This research has made use of Astropy and affiliated packages, a community-developed core Python package for Astronomy (Astropy Collaboration 2018), Python package SciPy (`https://www.scipy.org`), and APLpy, an open-source plotting package for Python (Robitaille & Bressert 2012). We thank the anonymous referee for the detailed comments and appreciate the effort that has clearly gone into reviewing our work. J.S. also thanks Mattia Sormani for the fruitful discussions that helped improve this paper.

References

Abreu-Vicente, J., Ragan, S., Kainulainen, J., et al. 2016, A&A, 590, A131
Alves, J., Zucker, C., Goodman, A. A., et al. 2020, Nature, 578, 237
Arzoumanian, D., André, P., Didelon, P., et al. 2011, A&A, 529, L6
Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, AJ, 156, 123
Bellomi, E., Godard, B., Hennebelle, P., et al. 2020, A&A, 643, A36
Bergemann, M., Sesar, B., Cohen, J. G., et al. 2018, Nature, 555, 334
Beuther, H., Bihr, S., Rugel, M., et al. 2016, A&A, 595, A32
Bialy, S., & Sternberg, A. 2016, ApJ, 822, 83
Bialy, S., & Sternberg, A. 2019, ApJ, 881, 160
Bialy, S., Bihr, S., Beuther, H., Henning, T., & Sternberg, A. 2017a, ApJ, 835, 126
Bialy, S., Burkhart, B., & Sternberg, A. 2017b, ApJ, 843, 92
Bihr, S., Beuther, H., Ott, J., et al. 2015, A&A, 580, A112
Brand, J., & Blitz, L. 1993, A&A, 275, 67
Burkhart, B., Lee, M.-Y., Murray, C. E., & Stanimirović, S. 2015, ApJ, 811, L28
Burkhart, B., Stalpes, K., & Collins, D. C. 2017, ApJ, 834, L1
Burton, W. B. 1988, The Structure of Our Galaxy Derived from Observations of Neutral Hydrogen, eds. K. I. Kellermann, & G. L. Verschuur (New York: Springer), 295
Chandrasekhar, S., & Fermi, E. 1953, ApJ, 118, 116
Chequers, M. H., & Widrow, L. M. 2017, MNRAS, 472, 2751
Chira, R. A., Ibáñez-Mejía, J. C., Mac Low, M. M., & Henning, T. 2019, A&A, 630, A97

Choudhuri, S., & Roy, N. 2019, MNRAS, 483, 3437
Churchwell, E., Babler, B. L., Meade, M. R., et al. 2009, PASP, 121, 213
Clark, P. C., Glover, S. C. O., Klessen, R. S., & Bonnell, I. A. 2012, MNRAS, 424, 2599
Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, ApJ, 547, 792
Dickey, J. M., & Lockman, F. J. 1990, ARA&A, 28, 215
Dickey, J. M., Mebold, U., Stanimirovic, S., & Staveley-Smith, L. 2000, ApJ, 536, 756
Dickey, J. M., McClure-Griffiths, N. M., Gaensler, B. M., & Green, A. J. 2003, ApJ, 585, 801
Dickey, J. M., Strasser, S., Gaensler, B. M., et al. 2009, ApJ, 693, 1250
Dobbs, C. L. 2015, MNRAS, 447, 3390
Draine, B. T. 1978, ApJS, 36, 595
Draine, B. T. 2011, Physics of the Interstellar and Intergalactic Medium (Kassel: Princeton University Press)
Duarte-Cabral, A., & Dobbs, C. L. 2016, MNRAS, 458, 3667
Duarte-Cabral, A., & Dobbs, C. L. 2017, MNRAS, 470, 4261
Edelsohn, D. J., & Elmegreen, B. G. 1997, MNRAS, 287, 947
Elias, F., Cabrera-Caño, J., & Alfaro, E. J. 2006, AJ, 131, 2700
Elmegreen, B. G., & Scalo, J. 2004, ARA&A, 42, 211
Federrath, C., & Klessen, R. S. 2013, ApJ, 763, 51
Federrath, C., Glover, S. C. O., Klessen, R. S., & Schmidt, W. 2008, Phys. Scr., 132, 014025
Federrath, C., Roman-Duval, J., Klessen, R. S., Schmidt, W., & Mac Low, M. M. 2010, A&A, 512, A81
Ferrière, K. M. 2001, Rev. Mod. Phys., 73, 1031
Field, G. B. 1965, ApJ, 142, 531
Field, G. B., Goldsmith, D. W., & Habing, H. J. 1969, ApJ, 155, L149
Gaia Collaboration (Brown, A. G. A., et al.) 2021, A&A, 649, A1
Giannetti, A., Leurini, S., König, C., et al. 2017, A&A, 606, L12
Gibson, S. J., Taylor, A. R., Higgs, L. A., & Dewdney, P. E. 2000, ApJ, 540, 851
Girichidis, P., Konstandin, L., Whitworth, A. P., & Klessen, R. S. 2014, ApJ, 781, 91
Glover, S. C. O., & Mac Low, M.-M. 2007, ApJ, 659, 1317
Glover, S. C. O., & Mac Low, M. M. 2011, MNRAS, 412, 337
Goldsmith, P. F., Li, D., & Krčo, M. 2007, ApJ, 654, 273
Goldsmith, P. F., Heyer, M., Narayanan, G., et al. 2008, ApJ, 680, 428
Goodman, A. A., Pineda, J. E., & Schnee, S. L. 2009, ApJ, 692, 91
Goodman, A. A., Alves, J., Beaumont, C. N., et al. 2014, ApJ, 797, 53
Gould, R. J., & Salpeter, E. E. 1963, ApJ, 138, 393
Güver, T., & Özel, F. 2009, MNRAS, 400, 2050
Heiles, C. 2001, ApJ, 551, L105
Heiles, C., & Troland, T. H. 2003, ApJ, 586, 1067
Hennebelle, P., & Pérault, M. 1999, A&A, 351, 309
Henshaw, J. D., Kruijssen, J. M. D., Longmore, S. N., et al. 2020, Nat. Astron., 4, 1064
Imara, N., & Burkhart, B. 2016, ApJ, 829, 102
Izquierdo, A. F., Smith, R. J., Glover, S. C. O., et al. 2021, MNRAS, 500, 5268
Jackson, J. M., Finn, S. C., Chambers, E. T., Rathborne, J. M., & Simon, R. 2010, ApJ, 719, L185
Kainulainen, J., Beuther, H., Henning, T., & Plume, R. 2009, A&A, 508, L35
Kainulainen, J., Beuther, H., Banerjee, R., Federrath, C., & Henning, T. 2011, A&A, 530, A64
Kainulainen, J., Federrath, C., & Henning, T. 2014, Science, 344, 183
Kalberla, P. M. W., & Kerp, J. 2009, ARA&A, 47, 27
Kalberla, P. M. W., Kerp, J., Haud, U., et al. 2016, ApJ, 821, 117
Klessen, R. S. 2000, ApJ, 535, 869
Klessen, R. S., & Glover, S. C. O. 2016, Saas-Fee Advanced Course, 43, 85
Koley, A., & Roy, N. 2019, MNRAS, 483, 593
Kolmogorov, A. 1941, Akademiia Nauk SSSR Doklady, 30, 301
Konstandin, L., Girichidis, P., Federrath, C., & Klessen, R. S. 2012, ApJ, 761, 149
Körtgen, B., Federrath, C., & Banerjee, R. 2019, MNRAS, 482, 5233
Kounkel, M., & Covey, K. 2019, AJ, 158, 122
Kritsuk, A. G., Norman, M. L., Padoan, P., & Wagner, R. 2007, ApJ, 665, 416
Larson, R. B. 1979, MNRAS, 186, 479
Larson, R. B. 1981, MNRAS, 194, 809
Levine, E. S., Blitz, L., & Heiles, C. 2006, ApJ, 643, 881
Li, D., & Goldsmith, P. F. 2003, ApJ, 585, 823
Lindner, R. R., Vera-Ciro, C., Murray, C. E., et al. 2015, AJ, 149, 138
Lombardi, M., Alves, J., & Lada, C. J. 2015, A&A, 576, L1
Majewski, S. R., Zasowski, G., & Nidever, D. L. 2011, ApJ, 739, 25
Marchal, A., & Miville-Deschênes, M.-A. 2021, ApJ, 908, 186
Marchal, A., Miville-Deschênes, M.-A., Orieux, F., et al. 2019, A&A, 626, A101
Marsh, K. A., Whitworth, A. P., Lomax, O., et al. 2017, MNRAS, 471, 2730
McKee, C. F., & Ostriker, J. P. 1977, ApJ, 218, 148
McKee, C. F., & Ostriker, E. C. 2007, ARA&A, 45, 565
Miville-Deschênes, M.-A., & Lagache, G. 2005, ApJS, 157, 302
Molina, F. Z., Glover, S. C. O., Federrath, C., & Klessen, R. S. 2012, MNRAS, 423, 2680
Murray, C. E., Lindner, R. R., Stanimirović, S., et al. 2014, ApJ, 781, L41
Murray, C. E., Stanimirović, S., Goss, W. M., et al. 2015, ApJ, 804, 89
Murray, C. E., Stanimirović, S., Goss, W. M., et al. 2018, ApJS, 238, 14
Nakanishi, H., Fujita, S., Tachihara, K., et al. 2020, PASJ, 72, 43
Nguyen, H., Dawson, J. R., Lee, M.-Y., et al. 2019, ApJ, 880, 141
Ossenkopf-Okada, V., Csengeri, T., Schneider, N., Federrath, C., & Klessen, R. S. 2016, A&A, 590, A104
Ostriker, J. 1964, ApJ, 140, 1056
Padoan, P., Nordlund, A., & Jones, B. J. T. 1997, MNRAS, 288, 145
Passot, T., & Vázquez-Semadeni, E. 1998, Phys. Rev. E, 58, 4501
Pineda, J. E., Caselli, P., & Goodman, A. A. 2008, ApJ, 679, 481
Pineda, J. L., Langer, W. D., Velusamy, T., & Goldsmith, P. F. 2013, A&A, 554, A103
Planck Collaboration I. 2014, A&A, 571, A1
Planck Collaboration XIX. 2011, A&A, 536, A19
Ragan, S. E., Henning, T., Tackenberg, J., et al. 2014, A&A, 568, A73
Rebolledo, D., Green, A. J., Burton, M., et al. 2017, MNRAS, 472, 1685
Reed, B. C. 2000, AJ, 120, 314
Reid, M. J., Dame, T. M., Menten, K. M., & Brunthaler, A. 2016, ApJ, 823, 77
Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2019, ApJ, 885, 131
Riener, M., Kainulainen, J., Henshaw, J. D., et al. 2019, A&A, 628, A78
Robitaille, T., & Bressert, E. 2012, Astrophysics Source Code Library [record ascl:1208.017]
Roy, N., Kanekar, N., Braun, R., & Chengalur, J. N. 2013a, MNRAS, 436, 2352
Roy, N., Kanekar, N., & Chengalur, J. N. 2013b, MNRAS, 436, 2366
Rugel, M. R., Beuther, H., Bihr, S., et al. 2018, A&A, 618, A159
Saury, E., Miville-Deschênes, M. A., Hennebelle, P., Audit, E., & Schmidt, W. 2014, A&A, 567, A16
Scalo, J. M. 1984, ApJ, 277, 556
Schneider, N., André, P., Könyves, V., et al. 2013, ApJ, 766, L17
Schneider, N., Ossenkopf, V., Csengeri, T., et al. 2015, A&A, 575, A79
Schneider, N., Bontemps, S., Motte, F., et al. 2016, A&A, 587, A74
Seifried, D., Schmidt, W., & Niemeyer, J. C. 2011, A&A, 526, A14
Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163
Smith, R. J., Glover, S. C. O., Clark, P. C., Klessen, R. S., & Springel, V. 2014a, MNRAS, 441, 1628
Smith, R. J., Glover, S. C. O., & Klessen, R. S. 2014b, MNRAS, 445, 2900
Smith, R. J., Treß, R. G., Sormani, M. C., et al. 2020, MNRAS, 492, 1594
Soler, J. D., Beuther, H., Syed, J., et al. 2020, A&A, 642, A163
Sparke, L. S. 1993, AIP Conf. Ser., 278, 447
Sternberg, A., Le Petit, F., Roueff, E., & Le Bourlot, J. 2014, ApJ, 790, 10
Stil, J. M., Taylor, A. R., Dickey, J. M., et al. 2006, AJ, 132, 1158
Strasser, S., & Taylor, A. R. 2004, ApJ, 603, 560
Su, Y., Yang, J., Zhang, S., et al. 2019, ApJS, 240, 9
Suri, S., Sánchez-Monge, Á., Schilke, P., et al. 2019, A&A, 623, A142
Syed, J., Wang, Y., Beuther, H., et al. 2020, A&A, 642, A68
Tang, N., Li, D., Heiles, C., et al. 2016, A&A, 593, A42
Tang, N., Li, D., Heiles, C., et al. 2017, ApJ, 839, 8
Umemoto, T., Minamidani, T., Kuno, N., et al. 2017, PASJ, 69, 78
Wakker, B. P., & van Woerden, H. 1997, ARA&A, 35, 217
Wang, K., Testi, L., Ginsburg, A., et al. 2015, MNRAS, 450, 4043
Wang, K., Testi, L., Burkert, A., et al. 2016, ApJS, 226, 9
Wang, Y., Bihr, S., Rugel, M., et al. 2018, A&A, 619, A124
Wang, Y., Beuther, H., Rugel, M. R., et al. 2020a, A&A, 634, A83
Wang, Y., Beuther, H., Schneider, N., et al. 2020b, A&A, 641, A53
Wang, Y., Bihr, S., Beuther, H., et al. 2020c, A&A, 634, A139
Westmeier, T. 2018, MNRAS, 474, 289
Widrow, L. M., Barber, J., Chequers, M. H., & Cheng, E. 2014, MNRAS, 440, 1971
Wilson, T. L., Rohlfs, K., & Hüttemeister, S. 2013, Tools of Radio Astronomy (Berlin Heidelberg: Springer Science & Business Media)
Wolfire, M. G., McKee, C. F., Hollenbach, D., & Tielens, A. G. G. M. 2003, ApJ, 587, 278
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868
Zucker, C., Battersby, C., & Goodman, A. 2015, ApJ, 815, 23
Zucker, C., Battersby, C., & Goodman, A. 2018, ApJ, 864, 153

Appendix A: Column density of Maggie component

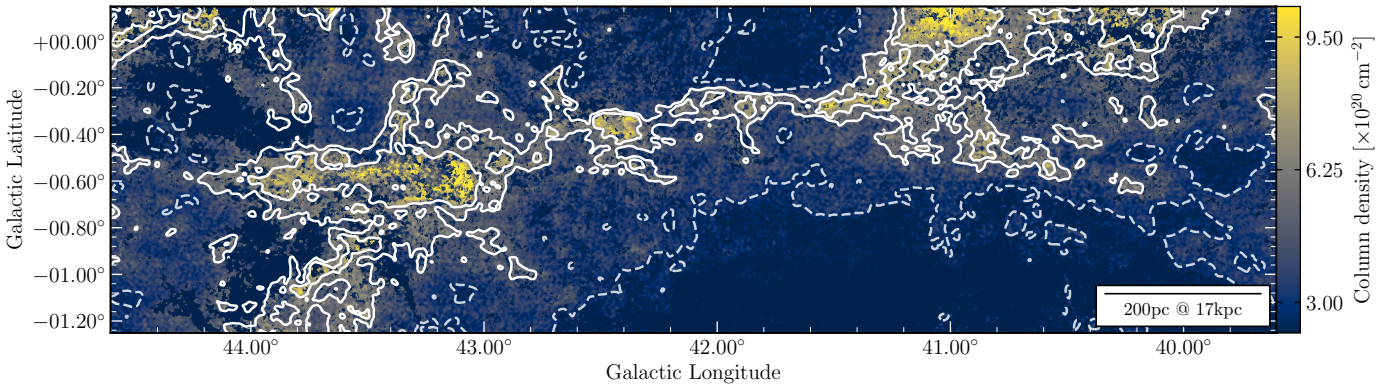

Fig. A.1. Integrated column density of the fit Maggie component. In a thermally bistable medium, two components (i.e., CNM and WNM), which are close in velocity, are encoded in the emission profile. We identify the Maggie component based on the mean position in the velocity domain, thus picking up both CNM and WNM features. The contours are the same as in Fig. 11.

Appendix B: MWISP CO observations

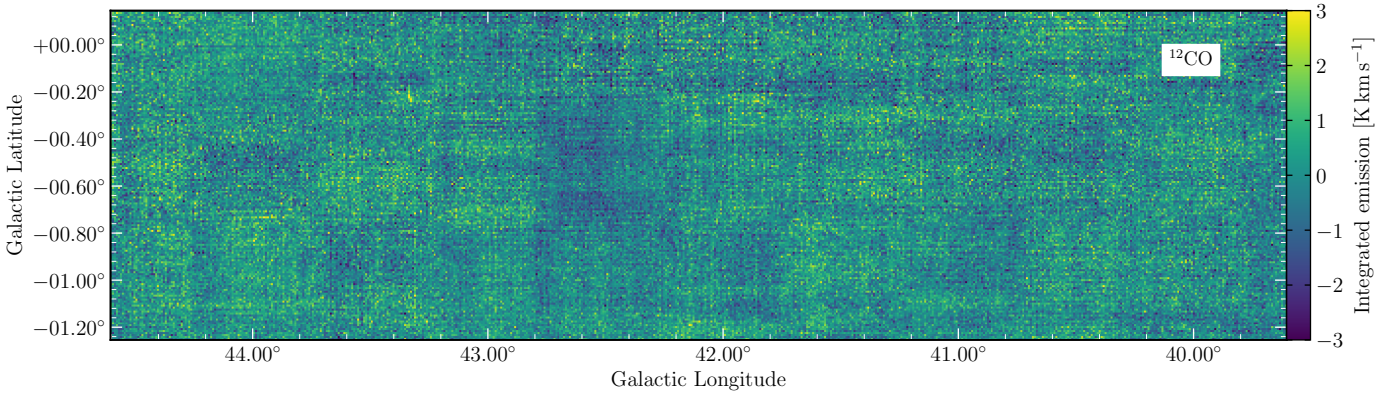

Fig. B.1. MWISP ^{12}CO integrated emission. The ^{12}CO (J =1–0) emission data taken from the MWISP survey are integrated over the velocity interval between −57.5 and −48.5 km s^{-1}.

Appendix C: Stellar extinctions using Gaia, 2MASS, and WISE

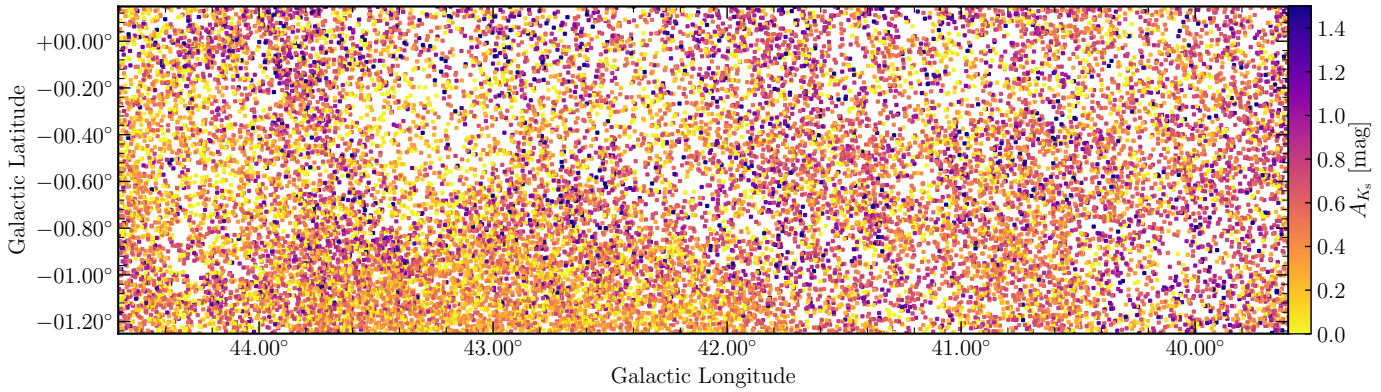

Fig. C.1. Extinctions up to 5 kpc distance using *Gaia* in combination with near- and mid-infrared photometry data from 2MASS and WISE. The blank spaces are regions containing obscuring material that has blocked out stars from our *Gaia* sample and left some highly extinguished stars in their surroundings.