

Structure and stability in TMC-1: Analysis of NH₃ molecular line and *Herschel* continuum data

O. Fehér^{1,2}, L. V. Tóth¹, D. Ward-Thompson³, J. Kirk³, A. Kraus⁴, V.-M. Pelkonen⁵, S. Pintér¹, and S. Zahorecz^{1,6}

¹ Eötvös Loránd University, Department of Astronomy, Pázmány Péter sétány 1/A, 1117 Budapest, Hungary
e-mail: feher.orsolya@cssf.k.mta.hu

² Konkoly Observatory, Research Centre for Astronomy and Earth Sciences, Hungarian Academy of Sciences, Konkoly Thege Miklós út 15-17, 1121 Budapest, Hungary

³ Jeremiah Horrocks Institute, University of Central Lancashire, Preston PR1 2HE, UK

⁴ Max Planck Institute for Radioastronomy, Auf dem Hügel 69, 53121 Bonn, Germany

⁵ Department of Physics, PO Box 64, University of Helsinki, 00014 Helsinki, Finland

⁶ European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching bei München, Germany

Received 12 June 2014 / Accepted 11 March 2016

ABSTRACT

Aims. We examined the velocity, density, and temperature structure of Taurus molecular cloud-1 (TMC-1), a filamentary cloud in a nearby quiescent star forming area, to understand its morphology and evolution.

Methods. We observed high signal-to-noise (S/N), high velocity resolution NH₃(1,1), and (2, 2) emission on an extended map. By fitting multiple hyperfine-split line profiles to the NH₃(1, 1) spectra, we derived the velocity distribution of the line components and calculated gas parameters on several positions. *Herschel* SPIRE far-infrared continuum observations were reduced and used to calculate the physical parameters of the *Planck* Galactic Cold Clumps (PGCCs) in the region, including the two in TMC-1. The morphology of TMC-1 was investigated with several types of clustering methods in the parameter space consisting of position, velocity, and column density.

Results. Our *Herschel*-based column density map shows a main ridge with two local maxima and a separated peak to the south-west. The H₂ column densities and dust colour temperatures are in the range of $0.5\text{--}3.3 \times 10^{22} \text{ cm}^{-2}$ and 10.5–12 K, respectively. The NH₃ column densities and H₂ volume densities are in the range of $2.8\text{--}14.2 \times 10^{14} \text{ cm}^{-2}$ and $0.4\text{--}2.8 \times 10^4 \text{ cm}^{-3}$. Kinetic temperatures are typically very low with a minimum of 9 K at the maximum NH₃ and H₂ column density region. The kinetic temperature maximum was found at the protostar IRAS 04381+2540 with a value of 13.7 K. The kinetic temperatures vary similarly to the colour temperatures in spite of the fact that densities are lower than the critical density for coupling between the gas and dust phase. The *k*-means clustering method separated four sub-filaments in TMC-1 with masses of 32.5, 19.6, 28.9, and $45.9 M_{\odot}$ and low turbulent velocity dispersion in the range of $0.13\text{--}0.2 \text{ km s}^{-1}$.

Conclusions. The main ridge of TMC-1 is composed of four sub-filaments that are close to gravitational equilibrium. We label these TMC-1F1 through F4. The sub-filaments TMC-1F1, TMC-1F2, and TMC-1F4 are very elongated, dense, and cold. TMC-1F3 is a little less elongated and somewhat warmer, and probably heated by the Class I protostar, IRAS 04381+2540, which is embedded in it. TMC-1F3 is approximately 0.1 pc behind TMC-1F1. Because of its structure, TMC-1 is a good target to test filament evolution scenarios.

Key words. molecular data – ISM: clouds – ISM: molecules – ISM: abundances – ISM: kinematics and dynamics – radio lines: ISM

1. Introduction

Large-scale filaments have long been recognized as fundamental environments of the star forming process. Several studies revealed that filaments are common structures in interstellar clouds that are populated by young stars in different stages of formation (Schneider & Elmegreen 1979; Bally et al. 1987; Hatchell et al. 2005; Goldsmith et al. 2008). Recent results based on *Herschel* far-infrared (FIR) observations of nearby star forming regions has once again directed attention to the connection between filaments, dense cores, and stars (see e.g. André et al. 2010). Since star formation occurs mostly in prominent filaments, characterizing the physical properties of these regions is the key to understand the process of star formation.

The Taurus molecular cloud (TMC) is one of the closest, low-mass star forming regions at 140 pc (Elias 1978; Onishi et al. 2002). It was a target of several cloud evolution

and star formation studies (Ungerechts & Thaddeus 1987; Mizuno et al. 1995; Goldsmith et al. 2008) and was mapped extensively in CO (Ungerechts & Thaddeus 1987; Onishi et al. 1996; Narayanan et al. 2008) and extinction (Cambrésy 1999; Padoan et al. 2002; Dobashi et al. 2005). The most massive molecular cloud in Taurus is the Heiles cloud 2 (HCL 2; Heiles 1968; Onishi et al. 1996; Tóth et al. 2004), which is the second brightest source on the 353, 545, and 857 GHz *Planck* maps (Tauber et al. 2010; Planck Collaboration I 2011) in Taurus with a flux density peak in HCL 2B (Heiles 1968).

There are 406 *Planck* Galactic Cold Clumps (PGCCs) in the Taurus region (Planck Collaboration XVIII 2016). These are clustered and form 76 groups as was found by Tóth et al. (2016) via the Minimum Spanning Tree method (Planck Collaboration XXIII 2011). One of the largest PGCC groups is HCL 2 with 14 cold clumps; their parameters are listed in Table 1. The clump temperatures and densities were derived

Table 1. Properties of the PGCCs in HCL 2.

N	PGCC clump	$\alpha(2000)$ [deg]	$\delta(2000)$ [deg]	$T_{\text{dust,min}}$ [K]	$T_{\text{dust,ave}}$ [K]	$N(\text{H}_2)_{\text{peak}}$ $[10^{22} \text{ cm}^{-2}]$	$N(\text{H}_2)_{\text{ave}}$ $[10^{22} \text{ cm}^{-2}]$	Other IDs
1	G174.09-13.24	70.4039	26.0315	9.9	12.4	4.1	0.7	TMC-1C ⁽¹⁾
2	G174.20-13.44	70.3034	25.8207	10.4	12.7	3.3	0.7	TMC-1 ⁽²⁾ , TMC-1(NH ₃) ⁽³⁾
3	G174.40-13.45	70.4373	25.6680	10.8	13.0	2.8	0.7	TMC-1 ⁽²⁾ , TMC-1 CP ^(3,4)
4	G173.95-13.76	69.8655	25.8070	11.2	12.5	3.0	1.1	HCL 2B ⁽²⁾
5	G173.44-13.34	69.8673	26.4553	10.5	12.7	3.1	0.5	TGU H1211 P6 ⁽⁵⁾ , TMC-1N ⁽⁶⁾
6	G173.14-13.32	69.6767	26.6971	11.8	12.9	0.9	0.3	TMC-1N ⁽⁶⁾
7	G174.57-13.68	70.3566	25.3955	12.3	13.3	1.1	0.6	HCL 2ES ⁽⁴⁾
8	G173.73-13.49	69.9406	26.1503	11.3	12.7	2.7	0.6	TGU H1211 P5 ⁽⁵⁾
9	G173.40-13.76	69.4780	26.2173	11.5	12.7	1.2	0.4	TGU H1211 P4 ⁽⁵⁾
10	G173.25-13.76	69.3752	26.3275	11.3	12.5	1.3	0.4	...
11	G173.07-13.69	69.3107	26.5089	11.8	12.5	0.8	0.4	[WMD94] Tau A 11 ⁽⁷⁾
12	G173.58-13.73	69.6301	26.1011	11.0	12.8	1.7	0.5	...
13	G173.64-13.14	70.1768	26.4379	11.3	12.9	0.5	0.2	...
14	G173.71-13.91	69.5612	25.8895	12.4	13.1	0.7	0.3	...

Notes. See Sect. 2.5 for clump temperatures and densities, and Fig. 1 as a finding chart.

References. (1) Myers et al. (1983); (2) Heiles (1968); (3) Little et al. (1979); (4) Tóth et al. (2004); (5) Dobashi et al. (2005); (6) Malinen et al. (2012); (7) Wood et al. (1994).

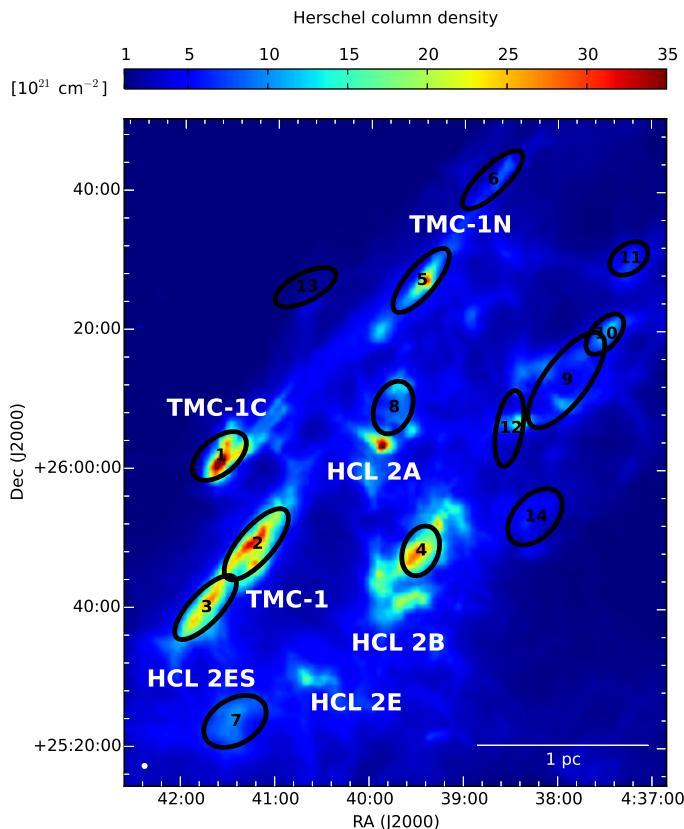


Fig. 1. *Herschel* $N(\text{H}_2)$ column density map of HCL 2 and the main parts of the cloud. PGCCs with flux quality of 1 and 2 are plotted with black ellipses and the numbers refer to Table 1, which lists the properties of the PGCCs.

from *Herschel* maps (see Sect. 2.4). The structure of the PGCC group is shown in Fig. 1, which also serves as a finding chart for the parts of HCL 2.

Molecular emission maps show the south-eastern region of HCL 2 as a ring-like structure called the Taurus molecular ring

(TMR; Schloerb & Snell 1984). TMC-1 appears as a dense, narrow ridge at the eastern edge of TMR. Two PGCCs are found in this region almost parallel with the galactic plane. Radio spectroscopic surveys indicated a considerable variation in the relative abundance of gas phase molecules, e.g. well-separated cyanopolyyne, ammonia (Little et al. 1979), and SO (Pratap et al. 1997) peaks (see also Table 1). Studies of NH₃, HC_nN ($n = 3, 5, 7$), C₄H, CS, C³⁴S, HCS⁺ (Tölle et al. 1981; Gaida et al. 1984; Olano et al. 1988; Hirahara et al. 1992), CH (Suutarinen et al. 2011; Sakai et al. 2012), and other carbon-chain and sulfur-containing molecules (Langer et al. 1995; Pratap et al. 1997) show complex velocity and chemical structure in TMC-1. Hirahara et al. (1992) suggested that the displacement of the column density peaks of carbon-chain and nitrogen bearing molecules is due to different evolutionary states of the regions. The molecular abundances in TMC-1 were modelled by McElroy et al. (2013) using the fifth release of the UMIST Database for Astrochemistry.

The morphologic and evolutionary connection between TMC-1 and HCL 2 was described by Schloerb et al. (1983) and Schloerb & Snell (1984) as TMC-1 being a collapsing, fragmented part of a torus shaped cloud structure (TMR). Cernicharo & Guelin (1987) considered TMC-1 a clump inside a thick, bent filamentary complex projecting itself as a loop on the plane of the sky. The existence and relative motion of fragments in TMC-1 were discussed in various papers (Snell et al. 1982; Schloerb et al. 1983; Olano et al. 1988; Hirahara et al. 1992; Langer et al. 1995; Pratap et al. 1997). Nutter et al. (2008) concluded that in sub-millimeter and FIR, TMR is composed of a dense filament on one side (TMC-1) and a series of point-like sources on the other. High spatial resolution *Herschel* (Pilbratt et al. 2010) observations resolved the cold clumps found by *Planck* and in those observations the physical characteristics of the clumps could be determined. Malinen et al. (2012) identified two long filaments in HCL 2 based on near-IR (NIR) extinction and *Herschel* data; one of these filaments was TMC-1. These authors fitted the column density perpendicular to the filament with a Plummer-like profile and derived an average width of 0.1 pc, which is typical of star forming filaments.

In this paper we discuss the substructures of TMC-1 using our high signal-to-noise ratio (S/N) $\text{NH}_3(1, 1)$, $\text{NH}_3(2, 2)$ line observations, and high resolution *Herschel* FIR maps. We examine the possibility that overlapping sub-filaments are present inside the cloud by partitioning it using various clustering methods in the derived parameter space. We reveal the morphology of TMC-1 and assess the physical state and stability of the different cloud parts.

2. Observations and data reduction

2.1. NH_3 observations

The observations of the $\text{NH}_3(J, K) = (1, 1)$ and $(2, 2)$ inversion transitions were carried out with the Effelsberg 100-m telescope¹ during November 21–30, 2008, April 8–13, 2015 and September 10, 2015. The map was obtained in raster mode, which was a number of targeted observations on a grid with a spacing of $40''$ (0.027 pc at the distance of TMC-1) with a velocity resolution of 0.038 km s^{-1} . The $(0, 0)$ position of the spectral grid was $\alpha(2000) = 04:41:42.5$, $\delta(2000) = 25:41:27$. The spectra from 2008 were observed in both position switching (PSW) and frequency switching (FSW) modes; both lines and polarizations were measured simultaneously. In the dataset from 2008, 44% of the spectra were PSW and 56% FSW, and the frequency throw of the FSW observations was 5 MHz. After folding and baseline subtraction of the FSW spectra, the FSW and PSW observations were averaged together weighted by noise. The spectra from 2015 were measured only in FSW mode with a frequency throw of 6 MHz. The pointing errors are within $5\text{--}10''$ for both datasets. We observed a total of 258 positions with a typical integration time of 4 minutes; a sample spectrum is shown in Fig. 2.

The observing sessions started by measuring the continuum emission of calibration sources (NGC 7027, W3OH). Pointing measurements towards a source close to the target were repeated at least once per hour. For the observations from 2008 NGC 7027 was used for flux calibration with an assumed flux density of 5.51 Jy (Zijlstra et al. 2008). With the average beam size of $37.1''$ of the receiver this is equivalent to a main beam brightness temperature $T_{\text{MB}} = 8.7 \text{ K}$. We calibrated the dependent atmospheric and instrumental gain and efficiency to the elevation with a quadratic function, which was calculated from observing NGC 7027 at different elevations during the observing sessions. The calibration of the 2008 dataset is estimated to be accurate to $\pm 15\%$.

The data from 2015 was observed with a different receiver (the secondary focus system instead of the prime focus K -band receiver used in 2008), which had an average beam size of $36.6''$. The flux calibrator was NGC 7027 again, but since a planetary nebula expands adiabatically, its flux density decreases. According to the model by Zijlstra et al. (2008), the flux density of NGC 7027 was $\approx 5.46 \text{ Jy}$ in 2015. This gives $T_{\text{MB}} = 8.8 \text{ K}$, which was used to calibrate the 2015 dataset. The spectra on the positions observed in both 2008 and 2015 have an $\approx 10\%$ difference in intensity, resulting in a 15% accuracy of the merged dataset. The data were reduced and further analysed with the software packages CLASS and GREG². The calibrated spectra have a typical rms noise of 0.15 K . $\text{NH}_3(1, 1)$ was detected with $S/N > 3$ on 171 positions and the $\text{NH}_3(2, 2)$ line was detected

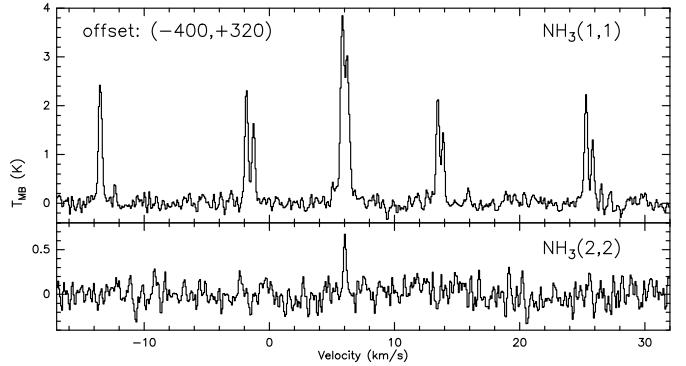


Fig. 2. Typical $\text{NH}_3(1, 1)$ and $\text{NH}_3(2, 2)$ spectra on our NH_3 grid. This position has the highest $\text{NH}_3(1, 1)$ optical depth from the spectra where we observed the $\text{NH}_3(2, 2)$ line as well (position #5 in Table 2).

with $S/N > 3$ on 17 positions. The given line temperatures in this paper are in T_{MB} .

2.2. Separating $\text{NH}_3(1, 1)$ velocity components

In order to discern the different velocity components in the $\text{NH}_3(1, 1)$ data, instead of fitting one hyperfine-split (HFS) line profile, we analysed the spectra as follows. We first fitted a single HFS line profile, subtracted it, and then searched the residual for a second component with $S/N > 3$. This line was also fitted with a HFS line profile. Since in some cases the HFS structure of the secondary components were not detected and the $\tau(1, 1)$ optical depth could not be determined well, the v_{LSR} line velocity in the local standard rest frame and the peak T_{MB} values from this fit were verified with a Gaussian profile fit. We then simultaneously fitted both the primary and secondary components with HFS line profiles on the original spectra. During this final fit the software CLASS required initial line parameters for both lines: $T_{\text{MB}}(1, 1)$, v_{LSR} , $\tau(1, 1)$, and Δv linewidth. For the primary component the initial parameters were the results of the single HFS line profile fit, and for the secondary component v_{LSR} and Δv were taken from the HFS line profile fit of the line on the residual, $T_{\text{MB}}(1, 1)$ from the Gaussian fit, and the initial optical depth was 1. None of these parameters were fixed during the simultaneous HFS profile fit. See Fig. 4, which demonstrates this process. We note that line components were not discernible on any of the $\text{NH}_3(2, 2)$ spectra.

Fitting a single HFS line profile to potentially multiple line components during the first step can be misleading, since to minimize the residual, the fitting process broadens the resulting line. In case the second component can be found on the residual, we can fit the original spectra again, using two components. However, if the secondary component is blended sufficiently and does not appear on the residual, a single HFS line profile fit is able to reproduce the line but gives false line parameters. This appears as an uncertainty in our calculations, however, the results are more precise where we identify the secondary component. This way we avoid the detection of a fake $\text{NH}_3(1, 1)$ maximum and get a more detailed picture of the velocity structure.

We examined the minimum separation of the two line components where it is still possible to identify the weaker line component on the residual. The $\text{NH}_3(1, 1)$ line with the smallest observed linewidth was added to an example spectra while its intensity and their channel separation varied, and then we proceeded to search these spectra for the secondary component with our method. With a channel width of 0.08 km s^{-1} and

¹ The 100-m telescope at Effelsberg/Germany is operated by the Max-Planck-Institut für Radioastronomie on behalf of the Max-Planck-Gesellschaft (MPG).

² <http://www.iram.fr/IRAMFR/GILDAS>

rms noise in the range of 0.1–0.3 K, we can only identify the secondary line component with a channel separation of three channels (0.24 km s^{-1}) if its intensity is greater than 50% of the primary component. With a channel separation of four channels (0.32 km s^{-1}), the secondary line with 30% of the primary line component intensity can already be identified. The intensities of the detected secondary components were between 10% and 40% of the primary component with velocity separations of about 0.08 km s^{-1} . Since we fitted both components with HFS line profiles and the separation of the two lines is always greater than the HFS linewidths, it is possible to detect close lines in a few cases.

2.3. Gas parameter calculations from NH_3 observations

When calculating gas temperatures and densities from the spectra, the primary (or if no secondary component was identified, the only) $\text{NH}_3(1, 1)$ line component was used along with the $\text{NH}_3(2, 2)$ line on the positions where the $(2, 2)$ line was detected with $S/N > 3$. We fitted the primary $\text{NH}_3(1, 1)$ components with Gaussian functions to obtain $T_{\text{MB}}(1, 1)$ and with HFS line profiles to get $\tau(1, 1)$, v_{LSR} and Δv . The $\text{NH}_3(2, 2)$ lines were fitted with Gaussian functions to obtain their peak main beam brightness temperatures, $T_{\text{MB}}(2, 2)$.

The calculations were carried out following Ho & Townes (1983), Ungerechts et al. (1986), and Harju et al. (1993). The T_{rot} rotation temperature of NH_3 was calculated with

$$T_{12} = \frac{-41.5}{\ln\left(\frac{-0.282}{\tau(1,1)} \ln\left(1 - \frac{T_{\text{MB}}(2,2)}{T_{\text{MB}}(1,1)} (1 - e^{-\tau(1,1)})\right)\right)}. \quad (1)$$

The T_{ex} excitation temperature of the $\text{NH}_3(1, 1)$ transition was derived from

$$T_{\text{MB}}(1, 1) = \frac{h\nu_{11}}{k} [F(T_{\text{ex}}) - F(T_{\text{bg}})] (1 - e^{-\tau(1,1)}), \quad (2)$$

where ν_{11} is the frequency of the $(1, 1)$ transition, T_{bg} is the cosmic background temperature (2.726 K), $F(T) = 1/(e^{h\nu_{11}/kT} - 1)$, and we assume a beam filling factor of 1 since the source is a nearby, extended cloud. We calculated the column density of the ammonia molecules in the upper transition level with

$$N_u = \frac{8\pi\nu_{11}^3}{c^3 A} F(T_{\text{ex}}) \int \tau(1, 1)(v) dv, \quad (3)$$

where A is the Einstein coefficient of the $\text{NH}_3(1, 1)$ transition ($1.7 \times 10^{-7} \text{ s}^{-1}$). Assuming a Gaussian line profile and integrating τ over the main line group we can write the integral as

$$\int \tau(1, 1)(v) dv = \frac{\sqrt{\pi}}{2\sqrt{\ln 2}} \Delta v \tau(1, 1). \quad (4)$$

The $N(\text{NH}_3(1, 1))$ total column density of the NH_3 molecules in the $(1, 1)$ state is

$$N(\text{NH}_3((1, 1))) = N_u(1, 1) + N_l(1, 1) = N_u(1, 1) (1 + e^{h\nu_{11}/kT_{\text{ex}}}). \quad (5)$$

Ammonia has two distinct species, the ortho- NH_3 ($K = 3n$) and the para- NH_3 ($K \neq 3n$), which arise from different relative orientations of the three hydrogen spins. Since the higher energy levels are by orders of magnitude less populated on temperatures around 10 K , we may estimate the $N_p(\text{NH}_3)$ para- NH_3 column density from the observed $(1, 1)$ and $(2, 2)$ transitions with

$$N_p(\text{NH}_3) = N(\text{NH}_3(1, 1)) \left(1 + \frac{5}{3} e^{-41.5/T_{12}}\right). \quad (6)$$

The T_{kin} kinetic temperature of the gas was calculated with the half-empiric equation from Tafalla et al. (2004) as follows:

$$T_{\text{kin}} = \frac{T_{12}}{1 - \frac{T_{12}}{42} \ln(1 + 1.1e^{-16/T_{12}})}. \quad (7)$$

The $n(\text{H}_2)$ local molecular hydrogen volume density was estimated using the equation from Ho & Townes (1983) as follows:

$$n(\text{H}_2) = \frac{A}{C} \frac{F(T_{\text{ex}}) - F(T_{\text{bg}})}{F(T_{\text{kin}}) - F(T_{\text{ex}})} [1 + F(T_{\text{kin}})], \quad (8)$$

where C is the collisional de-excitation rate ($8.5 \times 10^{-11} \text{ cm}^3 \text{ s}^{-1}$) from Danby et al. (1988).

We calculated $N_p(\text{NH}_3)$ upper limits from the secondary $\text{NH}_3(1, 1)$ line components at each of the 31 positions where those were detected (see Fig. 5 and Appendix B). Based on the HFS line profile fits, on 82 positions the $\tau(1, 1)$ optical depth had a relative error of less than 50% (median relative error was 17%), so the $N(\text{NH}_3(1, 1))$ column density could be estimated from these high S/N spectra.

2.4. Herschel observations and data reduction

The Taurus region was observed as part of the *Herschel* Gould Belt Survey (André et al. 2010; Kirk et al. 2013). The SPIRE instrument (Spectral and Photometric Imaging Receiver) used three arrays of bolometers to observe wavelengths centred at 250, 350, and $500 \mu\text{m}$ with bandwidths of 33% each; this effectively covered the $208\text{--}583 \mu\text{m}$ range (Griffin et al. 2010). The FWHM beam sizes were $17.6''$, $23.9''$, and $35.2''$ for the three bands, respectively, and the field-of-view was $4 \times 8'$. The Photodetector Array Camera and Spectrometer (PACS) observed wavelengths from 60 to $210 \mu\text{m}$ with two bolometer arrays, simultaneously with the $125\text{--}210 \mu\text{m}$ band (red) and with either the $60\text{--}85$ or the $85\text{--}125 \mu\text{m}$ band (blue; Poglitsch et al. 2010). The blue channels had 32×64 pixel arrays and the red channel had a 16×32 pixels array. Both channels covered a field of view of $1.75 \times 3.5'$ with full beam spacing in each band.

As a part of HCL 2, TMC-1 was mapped using the fast-scanning SPIRE/PACS parallel mode in two orthogonal scan directions. The *Herschel* observation IDs of the data are 1342202252 and 1342202253. The data presented in this paper were processed with HIPE³ 12.1.0. The level-0.5 timelines were calibrated and converted to physical units separately and joined in the level-1 stage. The destripper module was applied to remove the observation baseline. Then the sourceExtractorTimeline and the sourceExtractorSExtractor tasks were used to search for point sources with a given FWHM and to determine their fluxes. The point sources were then subtracted with the sourceSubtractor task from the timeline data of all three wavelengths. See the HIPE Help System⁴ for the description of the applied tasks. We located four point sources on the $350 \mu\text{m}$ timeline, one of them, IRAS 04381+2540 is also seen on 250 and $500 \mu\text{m}$. After point source subtraction the maps were combined using the naive map-making method. Absolute calibrated maps were produced with the ZeroPointCorrection task, which calculates the absolute offsets by cross-correlating with *Planck* HFI 545 and 857 GHz data. The calibration accuracy of the SPIRE data is estimated to be 7%. After convolution to $40''$ common resolution the maps were co-aligned on the $500 \mu\text{m}$ pixel-grid with $14''$ pixel width and each pixel was colour-corrected.

³ *Herschel* interactive processing environment, <http://hereschel.esac.esa.int/hipe/>

⁴ <http://hereschel.esac.esa.int/hcss-doc-14.0/>

2.5. Calculations from Herschel SPIRE data

We used the method by [Juvela et al. \(2012\)](#) to derive hydrogen column density from the SPIRE maps. We fitted the spectral energy distribution (SED) of each pixel with $B_\nu(T_{\text{dust}})^\beta$, where $B_\nu(T)$ is the Planck-function for a black body with colour temperature T_{dust} at ν frequency. The value of β varies between 1.8 to 2.2 in the cold dense interstellar medium (ISM) anti-correlates with temperature and correlates with column density and galactic position ([Juvela et al. 2015b](#)); see Appendix A.2 for further information. In our calculations the SED was fitted to all three SPIRE data points with a fixed $\beta=2$ spectral index and returned T_{dust} for each pixel.

The $N(\text{H}_2)$ molecular hydrogen column density was calculated as

$$N(\text{H}_2) = \frac{I_\nu}{B_\nu(T_{\text{dust}})\kappa m_{\text{H}}\mu}, \quad (9)$$

where $\mu=2.33$ the particle mass per hydrogen molecule, m_{H} is the atomic mass unit, κ is the dust opacity from the formula $0.1 \text{ cm}^2/\text{g}(\nu/1000 \text{ GHz})^\beta$ for high-density environments ([Beckwith et al. 1990](#); [Planck Collaboration XXIII 2011](#)) and I_ν is the intensity at $\nu=856 \text{ GHz}$ ($\lambda=350 \mu\text{m}$). The $350 \mu\text{m}$ *Herschel* band overlaps with the *Planck* 857 GHz band, which was used for zero-point correction. This way we minimize the error that originates from the absolute calibration.

The $N(\text{H}_2)$ distribution calculated this way does not exclude the contribution of structures in the foreground or background of HCL 2. In order to correct for this, the extent of HCL 2 was indicated with the $A_V=1$ magnitude contour ($9.4 \times 10^{20} \text{ cm}^{-2}$, where interstellar hydrogen starts to become molecular; [Bohlin et al. 1978](#)) on the $N(\text{H}_2)$ map derived as written above. Then the average 250, 350, and $500 \mu\text{m}$ intensity at this contour was subtracted from the corresponding SPIRE 250, 350, and $500 \mu\text{m}$ maps. The final T_{dust} and $N(\text{H}_2)$ maps were calculated from these background subtracted intensity images. Other methods of background subtraction are evaluated in Appendix A.1.

[Pagani et al. \(2015\)](#) concluded that a single-temperature SED fitting we use here may miss some of the cold dust in dense cores, thus we probably slightly underestimate the total column density. With only three SPIRE data points, we cannot fit multi-temperature SEDs to model an even colder component but this can be checked with independent methods, such as star-reddening. [Malinen et al. \(2012\)](#) compared FIR emission-based and NIR reddening-based $N(\text{H}_2)$ values in HCL 2, and they found that the two methods are in good agreement. A significant source of uncertainty in the $N(\text{H}_2)$ calculation is the error in the surface brightness, as shown by [Juvela et al. \(2012\)](#) using Monte Carlo simulations. Assuming a 13% uncertainty in the *Herschel* intensities an error of less than 1.5 K appears in regions with temperatures below 16 K. This corresponds to an error in $N(\text{H}_2)$ of a few times 10^{21} cm^{-2} in the outer region of the filament and around $1-2 \times 10^{22} \text{ cm}^{-2}$ in the densest regions.

3. Results

3.1. T_{dust} and $N(\text{H}_2)$ distribution from *Herschel* data

The *Herschel*-based $N(\text{H}_2)$ map of HCL 2 is shown in Fig. 1, in which the PGCCs are represented by numbered ellipses. All PGCCs with flux quality of 1 (flux density estimates $S/N > 1$ in all bands) and 2 (flux density estimates $S/N > 1$ only in the 857, 545, and 353 GHz *Planck* bands) are plotted on

the figure. Table 1 contains the calculated parameters of these PGCCs, where the columns are: (1) number, also in Fig. 1; (2) PGCC ID; (3, 4) equatorial coordinates; (5, 6) $T_{\text{dust,min}}$ minimum and $T_{\text{dust,ave}}$ average colour temperature in the clump; (7, 8) $N(\text{H}_2)_{\text{max}}$ maximum and $N(\text{H}_2)_{\text{ave}}$ average column density in the clump; and (9) other IDs of the clump. The PGCCs in TMC-1, TMC-1C, and HCL 2B are similarly cold with $T_{\text{dust,min}} \leq 11 \text{ K}$ and dense with $N(\text{H}_2)_{\text{ave}} \geq 7 \times 10^{21} \text{ cm}^{-2}$.

The *Herschel* T_{dust} map of TMC-1 is shown in Fig. 3a with the positions of high $\text{NH}_3(2, 2)$ S/N spectra, where temperatures and densities from the NH_3 data could be calculated (see Sect. 3.2). T_{dust} has a minimum of 10.4 K in the northern, densest region and is 11 K at its southern end. The lower density outskirts are a few K warmer.

The *Herschel* $N(\text{H}_2)$ map of TMC-1 is shown in Fig. 3b with the $\text{NH}_3(1, 1)$ main group integrated intensity contours overplotted (the secondary line components were already subtracted from the spectra). Both PGCC G174.20-13.44 and PGCC G174.40-13.45 are indicated by ellipses and the positions of TMC-1 CP, TMC-1(NH3), the SO peak, and IRAS 04381+2540 are also indicated. The map shows TMC-1 as a narrow ridge with an average diameter of 3–4' (0.12–0.17 pc). Three local $N(\text{H}_2)$ maxima appear in TMC-1. Two very elongated maxima are located along the ridge with peaks 2' north-west from TMC-1(NH3) with $N(\text{H}_2) = 3.3 \times 10^{22} \text{ cm}^{-2}$ and $\approx 30''$ south-west from TMC-1 CP with $N(\text{H}_2) = 2.7 \times 10^{22} \text{ cm}^{-2}$. The third local $N(\text{H}_2)$ maximum is at $\approx 18''$ south-east from the position of the Class I protostar IRAS 04381+2540 with $N(\text{H}_2) = 2.8 \times 10^{22} \text{ cm}^{-2}$. The PGCCs ellipses coincide with the northern and the southern maxima. The cloud regions cannot be simply separated by the half power contours, since the column density remains above 10^{22} cm^{-2} in the central ridge along the filament. We cannot achieve a robust separation of substructures inside TMC-1 (see Appendix C.1) using the $N(\text{H}_2)$ distribution only.

3.2. Results from NH_3 measurements

The $\text{NH}_3(1, 1)$ integrated intensity contours coincide well with the $N(\text{H}_2)$ distribution, as seen in Fig. 3b. The integrated intensity has a peak at the densest, coldest region of the filament, $\approx 1'$ to the south-east of TMC-1(NH3), and then the line intensities decrease gradually towards the southern edge, while $N(\text{H}_2)$ remains over 10^{22} cm^{-2} even at the southern end of the ridge.

The measured $\text{NH}_3(1, 1)$ primary component HFS linewidths are in the range of $0.15-1.0 \text{ km s}^{-1}$ with an average of $\Delta\nu$ of 0.39 km s^{-1} . As a result of the small linewidths and our high velocity resolution, we were able to separate subregions by their velocities and identify multiple velocity components. The spatial distribution of velocities of the $\text{NH}_3(1, 1)$ lines with $S/N > 3$ on 171 positions are plotted in Fig. 5a. The area indicated by these positions is above the *Herschel*-based 10^{22} cm^{-2} column density contour. At the northern region of the ridge the velocities are around $5.9-6.0 \text{ km s}^{-1}$ and most spectra show a secondary line component with velocities around $5.2-5.5 \text{ km s}^{-1}$. South-west of the main ridge, near the Class I young stellar source ([Terebey & Baud 1980](#)) IRAS 04381+2540, the primary line velocities are lower ($v_{\text{LSR}} \approx 5.2 \text{ km s}^{-1}$). The secondary line components in this same region have $v_{\text{LSR}} \approx 5.9 \text{ km s}^{-1}$, except for one component that has 5.6 km s^{-1} . There is a $4.9 \text{ km s}^{-1} \text{ pc}^{-1}$ velocity gradient seen in the middle section ($25:41:00 < \delta(2000) < 25:47:00$). The primary line component velocities decrease southwards from the mid-region of the ridge from 5.9 to 5.5 km s^{-1} , while the secondary component velocities increase.

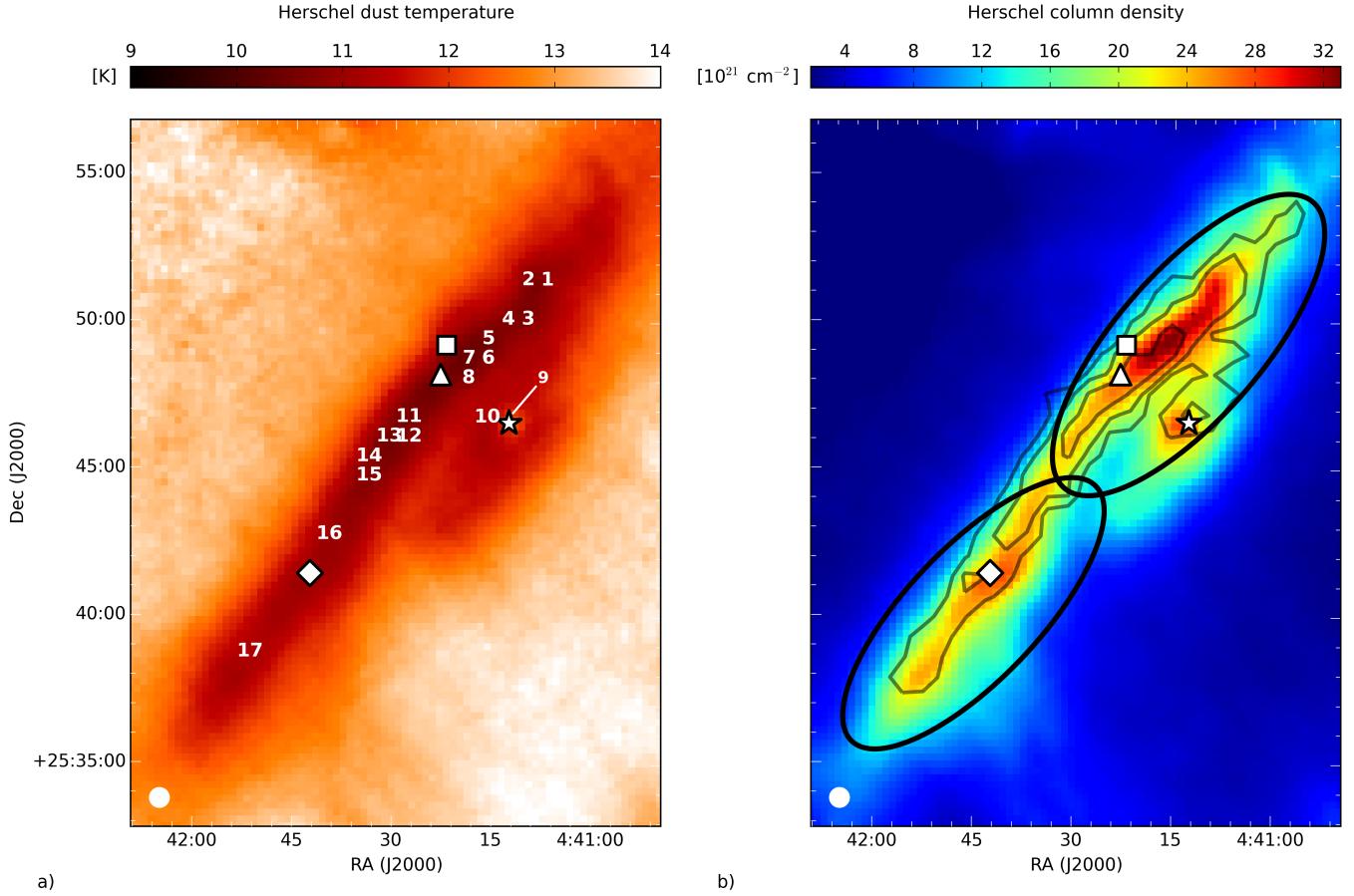


Fig. 3. **a)**: *Herschel*-based T_{dust} map of TMC-1. The numbers indicate the positions where high S/N $\text{NH}_3(1, 1)$ and $(2, 2)$ spectra were taken and used to calculate the physical parameters of the gas (see Table 2). **b)** *Herschel*-based $N(\text{H}_2)$ map with the $\text{NH}_3(1, 1)$ main group integrated intensity contours: 0.6, 1.0, and 2.0 K km s^{-1} (20, 35, and 65% of the maximum). The secondary line components were subtracted from the spectra. Black ellipses indicate the two PGCCs in TMC-1. The positions of TMC-1 CP, TMC-1(NH₃), the SO peak, and IRAS 04381+2540 are indicated with a diamond, a triangle, a square, and an asterisk, respectively. HPBW of $40''$ is shown in the left bottom corner.

The $N(\text{NH}_3(1, 1))$ was calculated on 82 positions that are all above the $1.3 \times 10^{22} \text{ cm}^{-2}$ *Herschel* contour; see Fig. 5b. The $N(\text{NH}_3(1, 1))$ distribution follows well the *Herschel*-based $N(\text{H}_2)$ and reaches $1.5 \times 10^{14} \text{ cm}^{-2}$ even around the CP peak. Its maximum ($1.4 \times 10^{15} \text{ cm}^{-2}$) is at the offset of $(-320, 440)$, with absolute coordinates of $\alpha(2000) = 04:41:13$, $\delta(2000) = 25:48:47$ (position 7 in Fig. 3b). That is $\approx 1'$ to north-west of the position which was formerly known as the ammonia peak, i.e. TMC-1(NH₃).

The $\text{NH}_3(2, 2)$ line was detected with $S/N > 3$ in as many as 17 spectra, we indicate the corresponding locations in TMC-1 as position #s where the number s runs from 1 to 17 (see Fig. 3a). Parameters at these positions were calculated according to Sect. 2.3 and the results are listed in Table 2. The derived T_{kin} values vary between 9 and 13 K with the exception of position #10 with 13.7 K. The NH_3 column density is $2.8 \times 10^{14} < N_p(\text{NH}_3) < 1.4 \times 10^{15} \text{ cm}^{-2}$ and the H_2 volume density is around $0.4 \times 10^4 < n(\text{H}_2) < 2.8 \times 10^4 \text{ cm}^{-3}$. The $N_p(\text{NH}_3)$ distribution generally follows the *Herschel*-based $N(\text{H}_2)$ distribution along the ridge, although their linear correlation is not very high (the correlation coefficient is $p = 0.4$). The $N(\text{NH}_3(1, 1))$ maximum coincides with the $N_p(\text{NH}_3) = N_p(\text{NH}_3)_{\text{prim}} + N_p(\text{NH}_3)_{\text{sec}}$ maximum on position #7 and this should be considered as the new NH_3 maximum of TMC-1.

We calculated the NH_3 relative abundance at our NH_3 peak and close to TMC-1 CP using the derived $N(\text{H}_2)$ and $N_p(\text{NH}_3)$ values in Table 2. This gives 5×10^{-8} and 1.4×10^{-8} for the NH_3 peak and TMC-1 CP, respectively, which are consistent with the estimates of Harju et al. (1993) for dense cores in Orion and Cepheus ($1-5 \times 10^{-8}$) and with the chemical modelling of Suzuki et al. (1992) (around 10^{-8}). Generally we find higher abundances in the northern and mid-region of TMC-1 and lower values around the YSO and in the southern end.

The total velocity dispersion was calculated from the primary $\text{NH}_3(1, 1)$ component of the 17 high S/N spectra with

$$\sigma_{\text{total}} = \frac{\Delta v}{\sqrt{8 \ln(2)}}, \quad (10)$$

where Δv is the HFS linewidth. The thermal component of the velocity dispersion is

$$\sigma_{\text{th}} = \sqrt{\frac{k_B T_{\text{kin}}}{\mu_{\text{NH}_3} m_H}}, \quad (11)$$

where $\mu_{\text{NH}_3} = 17$ is the ammonia molar mass and m_H is the atomic mass unit. The non-thermal (turbulent) component was then derived by

$$\sigma_{\text{nth}}^2 = \sigma_{\text{total}}^2 - \sigma_{\text{th}}^2. \quad (12)$$

Table 2. Results from our NH₃ and the *Herschel* observations.

NH ₃ observations										<i>Herschel</i> observations		
<i>N</i>	$\Delta\alpha$ ["]	$\Delta\delta$ ["]	Δv [km s ⁻¹]	τ	T_{kin} [K]	T_{ex} [K]	T_{rot} [K]	$N_p(\text{NH}_3)_{\text{prim}}$ [10 ¹⁴ cm ⁻²]	$N_p(\text{NH}_3)_{\text{sec}}$ [10 ¹⁴ cm ⁻²]	$n(\text{H}_2)$ [10 ⁴ cm ⁻³]	$N(\text{H}_2)$ [10 ²² cm ⁻²]	T_{dust} [K]
1*	-480	600	0.251 ± 0.006	2.5 ± 0.2	12.9 ± 1.2	5.1 ± 0.2	11.9 ± 0.9	4.73 ± 0.5	≤0.5	0.7 ± 0.1	2.4	11.0
2*	-440	600	0.253 ± 0.004	2.9 ± 0.2	10.1 ± 0.8	5.1 ± 0.2	9.7 ± 0.7	5.29 ± 0.4	≤0.5	0.8 ± 0.2	1.4	11.7
3*	-440	520	0.255 ± 0.006	2.9 ± 0.2	11.3 ± 0.8	5.3 ± 0.2	10.7 ± 0.7	5.64 ± 0.5	≤0.5	0.9 ± 0.2	3.1	10.7
4*	-400	520	0.253 ± 0.002	3.7 ± 0.1	9.0 ± 0.5	5.8 ± 0.3	8.7 ± 0.4	7.63 ± 0.5	≤0.5	1.6 ± 0.4	2.7	10.8
5*	-360	480	0.251 ± 0.004	5.5 ± 0.2	9.5 ± 0.6	6.3 ± 0.4	9.1 ± 0.5	12.41 ± 1.0	≤0.2	1.9 ± 0.5	2.9	10.6
6	-360	440	0.295 ± 0.008	2.8 ± 0.2	11.4 ± 0.9	6.2 ± 0.3	10.8 ± 0.8	7.45 ± 0.8	...	1.4 ± 0.3	3.1	10.4
7*	-320	440	0.332 ± 0.003	4.5 ± 0.1	9.0 ± 0.4	6.7 ± 0.4	8.7 ± 0.3	14.20 ± 1.0	≤0.5	2.8 ± 0.9	2.9	10.5
8	-320	400	0.391 ± 0.015	3.4 ± 0.3	10.6 ± 1.1	5.4 ± 0.3	10.1 ± 0.9	10.46 ± 1.2	...	1.0 ± 0.3	2.7	10.6
9*	-400	320	0.278 ± 0.004	1.6 ± 0.1	12.0 ± 0.8	5.5 ± 0.2	11.2 ± 0.6	3.47 ± 0.3	≤0.6	0.9 ± 0.1	2.4	11.3
10*	-360	320	0.221 ± 0.012	2.8 ± 0.6	13.7 ± 1.8	4.2 ± 0.2	12.6 ± 1.5	3.95 ± 0.9	≤0.7	0.4 ± 0.1	1.8	11.4
11	-200	320	0.249 ± 0.009	4.3 ± 0.4	9.2 ± 0.6	6.3 ± 0.3	8.8 ± 0.4	8.88 ± 0.6	...	2.1 ± 0.5	2.3	10.7
12	-200	280	0.308 ± 0.007	2.4 ± 0.2	10.4 ± 0.9	5.0 ± 0.2	9.9 ± 0.8	5.24 ± 0.5	...	0.8 ± 0.2	2.4	10.6
13	-160	280	0.356 ± 0.006	3.3 ± 0.1	9.7 ± 0.7	6.0 ± 0.3	9.3 ± 0.5	9.93 ± 0.7	...	1.5 ± 0.4	2.1	10.7
14*	-120	240	0.249 ± 0.009	4.3 ± 0.4	10.5 ± 1.1	5.0 ± 0.2	10.0 ± 0.9	7.70 ± 0.9	≤0.3	0.8 ± 0.2	2.0	10.9
15	-120	200	0.288 ± 0.005	2.7 ± 0.2	11.0 ± 0.7	5.7 ± 0.2	10.5 ± 0.5	6.45 ± 0.5	...	1.1 ± 0.2	2.4	10.8
16	-40	80	0.359 ± 0.006	1.0 ± 0.1	11.0 ± 1.0	5.9 ± 0.2	10.4 ± 0.8	2.93 ± 0.2	...	1.2 ± 0.3	2.1	11.1
17*	120	-160	0.262 ± 0.005	1.6 ± 0.2	10.1 ± 0.9	4.7 ± 0.1	9.7 ± 0.8	2.84 ± 0.3	≤0.3	0.7 ± 0.1	2.4	11.3

Notes. The columns are: (1) number of position in Fig. 3, asterisk marks where a secondary line component was found and subtracted before the calculations; (2, 3) right ascension and declination offsets; (4) linewidth from HFS line profile fit; (5) optical depth; (6) kinetic temperature; (7) excitation temperature; (8) rotation temperature; (9, 10) NH₃ column density from the primary and the secondary components respectively; (11) H₂ volume density; and (12, 13) *Herschel* H₂ column density and colour temperature. Numbers in the first column refer to the marked positions in Fig. 3a and asterisk indicates where a secondary line component was found.

The thermal component is around 0.066–0.08 km s⁻¹ in the 17 spectra and the non-thermal component is higher than this in almost every case, except on positions #1 and #10. The highest turbulence is observed on positions #8, #13, and #16.

4. Discussion

4.1. Density, temperature, and velocity structure

We obtained the most extensive NH₃ map of TMC-1 so far, in which both the (1, 1) and (2, 2) transitions were measured with a low noise level and high velocity resolution. Our observations confirm the variation in the NH₃(1, 1) line intensity in TMC-1 that was described by several studies (Little et al. 1979; Olano et al. 1988; Hirahara et al. 1992; Pratap et al. 1997). Our average HFS linewidths agree with the values of Tölle et al. (1981), although we mapped a significantly larger area.

Olano et al. (1988) measured NH₃(1, 1) with similar beam size, velocity resolution, and S/N in the high-density region of TMC-1. These authors also observed the velocity difference between the middle section of the ridge and its southern end and a considerably different velocity for the region around IRAS 04381+2540 (see their Fig. 9.). They pointed out a velocity gradient across the ridge as well. We were able to detect further velocity variations because our mapping extends further along and across the ridge compared to theirs.

The T_{ex} values we derived agree well with the 4–7 K range given by Tölle et al. (1981). Their observations do not completely overlap with ours because they measured positions with a spacing of 45", but the difference between their values and ours on positions of less than 35" apart is less than 25%, relative to their values. These authors calculated T_{rot} on two positions that are relatively close ($\approx 40''$ and $\approx 5''$) to our positions #11 and #15: their values are 9.5 ± 1 K and 10.3 ± 2 K and we derived 8.8 K and 10.5 K, respectively. Gaida et al. (1984) repeated NH₃(1, 1) and (2, 2) measurements for the positions of Tölle et al. (1981) along the main ridge of TMC-1 and made two cuts perpendicular to the major axis as well. They calculated

both NH₃ column densities and kinetic temperatures on three positions that are not farther than 40" from our position #11, #13, and #15. Their T_{kin} and T_{rot} results are in agreement with ours, but their $N(\text{NH}_3(1, 1))$ values are consistently 4–5 times higher than ours and $N(\text{NH}_3)$ only agrees well on the higher density positions. Their v_{LSR} and $n(\text{H}_2)$ distribution along the main ridge also shows similar values to those that we measured. Our T_{ex} values are somewhat lower than those by Olano et al. (1988) (6–9 K), which they show on a plot along the TMC-1 major axis. Their sample does not overlap with our sample either because they used different (0, 0) offset coordinates.

Pratap et al. (1997) and Hirahara et al. (1992) derived a factor of 5–6 higher than our results of H₂ volume densities from HC₃N and C³⁴S spectra. Our NH₃ measurements may not trace the densest gas in TMC-1 because, as shown by Bergin et al. (2002) and Pagani et al. (2005), NH₃ is best used as a tracer at lower than 5×10^5 cm⁻³ densities. Also, because of the low critical density (around 2×10^3 cm⁻³) of the (1, 1) line, it may have a substantial contribution from external, warmer layers as well (Pagani et al. 2007). Our results do not show significant differences in volume densities along the ridge, while the observations by Hirahara et al. (1992) imply a gradient of a factor of 10 in $n(\text{H}_2)$ between TMC-1(NH3) and TMC-1 CP. The NH₃ densities calculated by Tölle et al. (1981) at TMC-1(NH3) are consistent with our results, but their values are a factor of 2–3 higher at TMC-1CP. The results by Pratap et al. (1997) are consistent with our results at TMC-1 CP but a factor of 2–3 lower than our values in the northern region.

Nutter et al. (2008) presented temperature and density profile modelling for TMC-1 based on SCUBA, IRAS, and *Spitzer* continuum data. Their best fit was given by an inner colder filament (8 K) and an outer layer (12 K) and resulted in $n(\text{H}_2)$ values in the range of 0.7 – 1.75×10^5 cm⁻³. They also calculated $N(\text{H}_2)$ from the 850 μm SCUBA map, which yielded 3.8×10^{22} and 4.4×10^{21} cm⁻² for the central region and the outer layer, respectively. Malinen et al. (2012) used 2MASS and *Herschel* maps to derive the properties for TMC-1. Their results are consistent with the findings of Nutter et al. (2008) and their $N(\text{H}_2)$

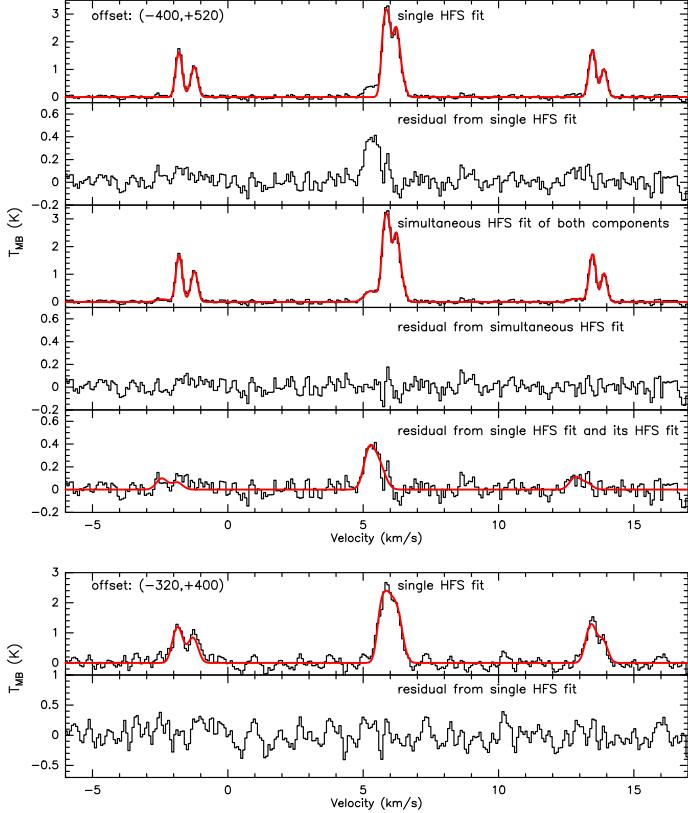


Fig. 4. Top: multiple HFS line profile fitting of $\text{NH}_3(1, 1)$ on the $(-400, +520)$ offset. The spectrum shows one of the brightest secondary line components and gives the lowest T_{kin} value; see position #4 in Table 2. Box 1: single HFS line profile fit; box 2: residual from the single HFS fit; box 3: simultaneous HFS profile fitting of both components; box 4: the residual from the simultaneous HFS fit; box 5: the residual from the single HFS fit with the HFS fit of the secondary component shown in box 4. Bottom: single $\text{NH}_3(1, 1)$ HFS line profile fit on position #8 and the residual after subtracting the fit. This position has the highest HFS linewidth from the spectra in Table 2. We note the different scales for the $(1, 1)$ and $(2, 2)$ lines.

peak is around $3.4 \times 10^{22} \text{ cm}^{-2}$. This is our derived $N(\text{H}_2)$ maximum as well. We note that our calculations are based on point source subtracted *Herschel* images. The T_{dust} values that we derived for the outer region of TMC-1 are consistent with the 12 K from Nutter et al. (2008), but we measure 10 K in the inner region instead of 8 K. The T_{kin} derived from NH_3 and T_{dust} derived from the *Herschel* measurements vary similarly but do not agree well, since the cooling-heating processes are different for gas and dust and the coupling of the two phases are not guaranteed in regions with $n(\text{H}_2) = 10^4 \text{ cm}^{-3}$ (Goldsmith 2001).

4.2. Substructures in TMC-1

4.2.1. Clustering methods and their results

The complex velocity structure and the appearance of secondary velocity components on several spectra suggest that different cloud parts are present in TMC-1. Since the filament is not necessarily one coherent structure and it could be interpreted as a chain of clumps, or even a complex superposition of filaments containing smaller scale cores, we apply several types of clustering methods to reveal its morphology.

Clump-finding in the NH_3 PPV datacube with methods like the 3D version of CLUMPFIND, GAUSSCLUMP, or the dendrogram technique (see Appendix C.1) is problematic in the case of TMC-1, since NH_3 abundance can be position-dependent in the cloud. Moreover, NH_3 alone is not a good tracer for clumps because of high optical depth effects such as self-absorption, which affects the line profiles. Searching for clumps in the *Herschel*-based $N(\text{H}_2)$ distribution is not enough either, since $N(\text{H}_2)$ changes smoothly along the ridge. As we can see from a quick decomposition of the filament with CLUMPFIND in Appendix C.1, we do not get a robust result; the extent and shape of the derived clumps strongly depend on the initial parameters given to the clump finding algorithm. Also, this way we ignore the velocity information derived from molecular emission that traces the relative motions of the gas and can help identify structures in the line of sight.

We expect that the presence of a clump changes the $N(\text{H}_2)$ distribution in its direction. We assume that inside a clump $N(\text{H}_2)$ changes continuously, velocities scatter around a certain value and the clump appears as an object roughly in the same direction on the plane of the sky. Thus, we combine the v_{LSR} derived from the HFS line profile fit of the $\text{NH}_3(1, 1)$ primary components and the *Herschel*-based $N(\text{H}_2)$ values on every position where the $(1, 1)$ line was observed with $S/N > 3$. Then we search for clusters in this datacube consisting of position, position, velocity, and column density. The importance of the secondary line components is discussed later (see Sect. 4.2.2).

The number of clusters in the cloud was not known and we did not want to decide which clustering method to use for the proportioning beforehand. Because of this we used the *NbClust* algorithm of the *R* statistical computing environment (Charrad et al. 2014), which is able to run several types of clustering methods while varying the number of clusters to be derived. It computes so-called cluster validity indices to inspect the results of the different runs, decides what is the most appropriate number of clusters in a dataset, and also performs the clustering itself. The *NbClust* algorithm is able to run the k -means method (see Appendix C.2) and hierarchical agglomerative clustering (HAC) methods with eight different agglomeration criteria (see Appendix C.3). A more detailed description of the *NbClust* package and the indices used can be found in Appendix C.4.

Since our parameters in the dataset are measured on different scales, standardization is necessary. We converted the position offsets to galactic offsets in arcminutes, velocity to tenth of km s^{-1} and column density to 10^{21} cm^{-2} to bring the parameter ranges closer to each other. Then we used the min-max scaling where the normalized value of parameter p_i of the n th point in the dataset is defined by

$$p_{i,n,\text{norm}} = \frac{p_{i,n} - p_{i,\min}}{p_{i,\max} - p_{i,\min}}, \quad (13)$$

where p_i are the four parameters (Δl , Δb , v_{LSR} , $N(\text{H}_2)$) of an $x = [p_1, p_2, p_3, p_4]$ point in the 4D parameter space with n points. The min-max scaling or scaling by the range of the variables was shown to give more accurate results for the most known clustering techniques than results found with traditional standardization methods, such as z -score (Steinley 2004).

The *NbClust* algorithm also requires the definition of a distance metric. We used Euclidean distance metric where the distance between two points (x_1 and x_2) in the $i = 4$ dimensional parameter space is

$$d(x_1, x_2) = \sqrt{\sum_{i=1}^4 (x_{1,i} - x_{2,i})^2}. \quad (14)$$

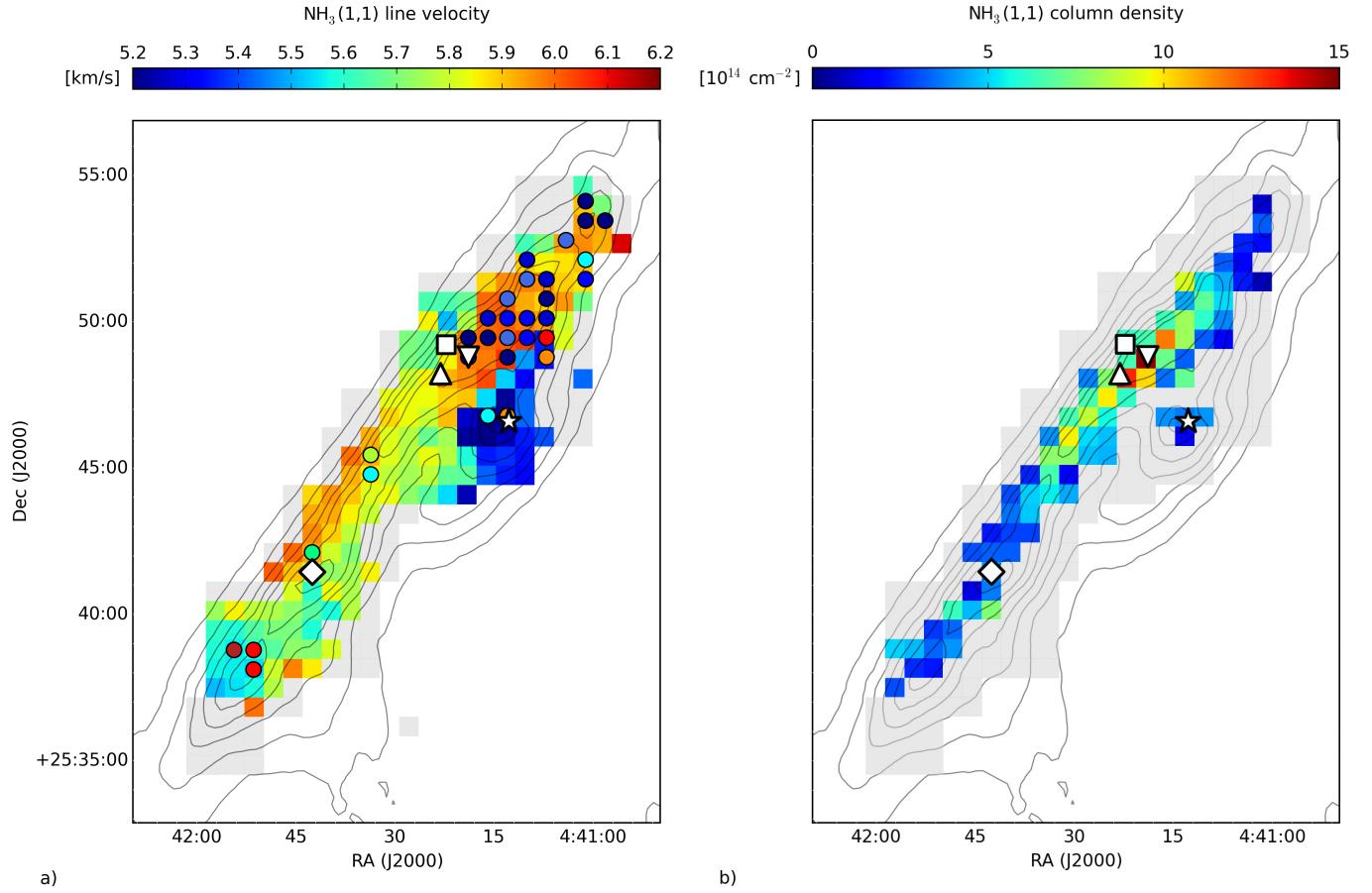


Fig. 5. a): v_{LSR} distribution of the main (or only) line component of the $\text{NH}_3(1, 1)$ lines with $S/N > 3$. Circles indicate the positions where a secondary component was found and their color corresponds to the velocity of this component. b): $N(\text{NH}_3(1, 1))$ distribution map calculated on positions where the τ optical depth from single HFS profile fit had an error of less than 50%. The grey area shows the total extension of the observed NH_3 map and the grey contours are the *Herschel* $N(\text{H}_2)$: 0.7, 1.0, 1.3, 1.6, 2, 2.3, $2.6 \times 10^{22} \text{ cm}^{-2}$ (from 20 to 80% of the maximum by 10%). The positions of TMC-1 CP, TMC-1(NH3), the SO peak, and IRAS 04381+2540 are indicated as before and the position of our NH_3 peak is indicated with an upside down triangle.

We can then define within-cluster and between-cluster distances for points inside one cluster and of two different clusters, respectively.

The nine available clustering methods in *NbClust* were run on the scaled and normalized TMC-1 joined dataset. When determining the ideal number of clusters with a certain clustering method, the number of clusters that was allowed to be derived varied from 2 to 15. The *NbClust* algorithm computed 30 indices to measure the validity of the derived clusters (see Appendix C.4) in each case. Each index suggested the ideal number of clusters that should be derived by the method and the final decision was made by the majority rule. After this the clustering was performed. The final results of the nine clustering methods are plotted on Fig. C.2.

Four clustering methods provided similar division of TMC-1 into four parts; see Figs. C.2b, c, d and 6. The HAC method with the average agglomeration criterion derived three clusters (Fig. C.2e), McQuitty's criterion derived 15 parts (Fig. C.2a), and the HAC method with the median, centroid and single agglomeration criteria did not part the filament (Fig. C.2f–h).

In order to decide between the results of the nine clustering methods, the sum and the weighted sum of the within-cluster variances were determined for each method. We first calculated the variance of the distances of all cluster members from the cluster centres in every cluster resulted by each clustering

method with

$$S_k = (\langle x \rangle - x_k)^2 \quad (15)$$

$$\sigma^2 = \frac{\sum_k (S_k - S_{\bar{k}})^2}{k}, \quad (16)$$

where k is the index of points inside a cluster, S_k are the distances of each cluster member (x_k) from the cluster centre ($\langle x \rangle$) in each cluster, and σ^2 is the variance of these distances for each cluster. Then we calculated a sum (C_m) and a weighted sum ($C_{m,w}$), adding together the cluster variances of the clusters belonging to each clustering method with

$$C_m = \sum \sigma^2 \quad (17)$$

$$C_{m,w} = \sum w \sigma^2, \quad (18)$$

where w weight is the number of pixels in a cluster. The most appropriate number of clusters and the number of indices suggesting that number are found in Table 3 for each clustering method, along with the derived sum and weighted sum of variances for the methods.

The HAC method with McQuitty's agglomeration criterion derived the clusters with the minimal sum and weighted sum of variances in the 4D-parameter space. However, when the clusters are projected back to the sky we find one fragment that is

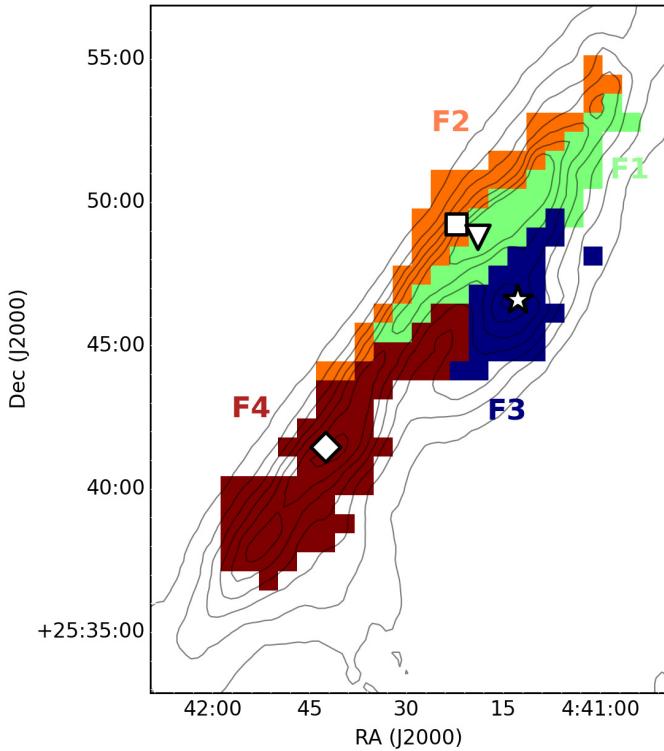


Fig. 6. Resulting clusters from the k -means clustering method projected back to the plane of the sky. The grey contours are the *Herschel* $N(\text{H}_2)$: 0.7, 1.0, 1.3, 1.6, 2, 2.3, $2.6 \times 10^{22} \text{ cm}^{-2}$ (from 20 to 80% of the maximum by 10%). The positions of TMC-1 CP, IRAS 04381+2540, the SO peak, and our NH_3 maximum are indicated as before.

Table 3. Results from the *NbClust* method.

m	N	I	C_m	$C_{m,w}$	Figure
<i>k</i> -means	4	11	0.0054	0.25	Fig. 6
McQuitty*	15	6	0.0029	0.048	Fig. C.2a
Ward	4	7	0.0059	0.38	Fig. C.2b
Ward2	4	8	0.0069	0.38	Fig. C.2c
Complete	4	9	0.0086	0.44	Fig. C.2d
Average	3	7	0.0123	0.78	Fig. C.2e
Median	2	7	0.0126	2.16	Fig. C.2f
Centroid	2	9	0.0129	2.22	Fig. C.2g
Single	2	8	0.0131	2.26	Fig. C.2h

Notes. The columns are: (1) name of the clustering method; (2) ideal number of clusters; (3) number of indices proposing N ; (4, 5) sum (Eq. (17)) and weighted sum (Eq. (18)) of within-cluster variances; and (6) reference to the figure showing the clusters projected to the plane of the sky. (*) fragments that are too small relative to the beam size.

too small (only 3 pixels), and three fragments that are not continuous, but each of these fragments is broken into two or three isolated parts. The existence of several very small fragments may be proved measuring several chemical species and using high spatial resolution and small spacing, but in the current study we accept the results of the k -means method as best, which had the second minimal sum of within-cluster variances.

The four clusters resulted by the k -means method were projected back to the plane of the sky, defining four objects, each of which is contiguous (see Fig. 6). Three of these sub-filaments are very elongated (TMC-1F1, TMC-1F2, TMC-1F4) and one of them (TMC-1F3) is less so. TMC-1F1 appears with the highest

velocities and column densities, while TMC-1F2 on the north-eastern edge has the lowest kinetic temperature. The environment of the YSO makes up TMC-1F3. TMC-1F4 encompasses the middle and southern regions with a range of velocities and the highest average linewidth.

The secondary $\text{NH}_3(1, 1)$ line components in TMC-1F3 appear at a velocity of the neighbouring TMC-1F1. Similarly, the secondary $\text{NH}_3(1, 1)$ line components in TMC-1F1 appear at $5.2\text{--}5.4 \text{ km s}^{-1}$, i.e. the velocity of TMC-1F3. The kinematics thus suggests that the clumps are slightly overlapped. Few spectra with secondary components in TMC-1F4 indicate small fragments with velocities similar to that of TMC-1F1 and TMC-1F2.

Despite the scaling and normalization of the parameters of the joined dataset, the Euclidean distance measured in one dimension cannot be unequivocally compared with the distance measured in another, since the parameters are independent and not interchangeable. For example, the min-max normalization only works well if the dispersion in each dimension is small and there are no extreme points that distort the calculations. The careful selection of the 82 spectral positions with high S/N NH_3 measurements and smoothly changing $N(\text{H}_2)$ ensures that this is not the case in this dataset. However, to reveal the discrepancies originating from using these statistical methods with a dataset like this, in a later study our methods should be carried out side by side with other clump-finding methods on a more extended dataset, where the decomposition of the molecular cloud is known, can be verified, and the results can be compared.

4.2.2. The parameters and stability of the sub-filaments

Averaged *Herschel*-based $T_{\text{dust},\text{ave}}$ were calculated inside the sub-filament contours and were found to be the same within 0.5 K for each sub-filament. The $T_{\text{kin},\text{ave}}$ kinetic temperatures of the sub-filaments were calculated as an average of the corresponding T_{kin} values on the positions of the 17 high S/N $\text{NH}_3(2, 2)$ spectra (see Table 2) inside the given sub-filament. Apparently, T_{kin} shows a higher variation than T_{dust} , which is also seen in the averaged values. Background corrected column densities were calculated for each sub-filament subtracting an average background $N(\text{H}_2)$ of HCL 2. Then the total hydrogen mass was calculated integrating inside the sub-filament boundary. In our simplified assumptions the sub-filaments are considered cylinders with L length and r radius and they are roughly perpendicular to the line of sight. Their maximum diameter along the galactic plane is L and r is half of their diameter perpendicular to this. The linear mass of the sub-filaments was calculated dividing their mass with the L length. TMC-1F1 has the highest $N(\text{H}_2)_{\text{ave}}$ value and TMC-F4 has the highest mass.

The greatest uncertainty of the mass calculation is the error in the $N(\text{H}_2)$ distribution caused by the ≈ 1.5 K uncertainty of the T_{dust} map. According to this, the calculated linear masses can have an error up to 56%. We note here that the uncertainty of the T_{dust} values was a conservative estimate but still this error makes our virial stability consideration uncertain. Additionally, because our sub-filaments overlap, $N(\text{H}_2)$ derived from the dust emission originates from more than one structure. The ratio of $N_p(\text{NH}_3)$ derived from the secondary components and the total $N_p(\text{NH}_3)$ is around 10%. If we assume that the relative abundance of NH_3 does not change much inside a sub-filament, this means that the ratio of $N_p(\text{NH}_3)$ from the main and secondary line components reflects the ratio of $N(\text{H}_2)$ between the two overlapping cloud parts. Thus we can say that the contribution of the overlapping part of a sub-filament to the mass of the other one is negligible.

Table 4. Physical parameters of the sub-filaments; see Fig. 6.

Sub-filament	$\alpha(2000)$	$\delta(2000)$	$v_{\text{LSR,ave}}$	$N(\text{H}_2)_{\text{ave}}$	$T_{\text{dust,ave}}$	$T_{\text{kin,ave}}$	Δv_{ave}	L	r	$M_{\text{lin,obs}}$	$M_{\text{lin,vir}}$	$\sigma_{\text{th,ave}}$	$\sigma_{\text{nth,ave}}$
	[deg]	[deg]	[km s $^{-1}$]	[10 22 cm $^{-2}$]	[K]	[K]	[km s $^{-1}$]	[pc]	[pc]	[M_{\odot}/pc]	[M_{\odot}/pc]	[km s $^{-1}$]	[km s $^{-1}$]
TMC-1F1	70.3048	25.8235	5.93	2.2	11.1	10.3	0.359	0.5	0.14	32.5	25.4	0.071	0.135
TMC-1F2	70.3341	25.8302	5.90	1.8	11.4	9.9	0.383	0.6	0.13	19.6	26.3	0.069	0.147
TMC-1F3	70.3026	25.7748	5.28	1.7	11.6	12.9	0.407	0.3	0.15	28.9	32.2	0.079	0.154
TMC-1F4	70.4220	25.6879	5.77	1.8	11.6	10.4	0.527	0.5	0.19	45.9	38.1	0.071	0.212

Notes. The columns are: (1) number in Fig. 6; (2, 3) coordinates of the sub-filament centre; (4) average v_{LSR} ; (5, 6) average *Herschel* H₂ column density and colour temperature in the sub-filament; (7, 8) average kinetic temperature and linewidth in the sub-filament; (9, 10) length and radius of the clump; (11, 12) linear mass and virial mass per unit length of the sub-filament; and (13, 14) average thermal and non-thermal velocity dispersion in the sub-filament.

We calculated the total ($\sigma_{\text{total,ave}}$), the thermal ($\sigma_{\text{th,ave}}$) and non-thermal velocity dispersion ($\sigma_{\text{nth,ave}}$) inside each sub-filament from the Δv HFS linewidth of their averaged spectra with Eqs. (10)–(12). We note that the total velocity dispersion includes the velocity gradients inside the sub-filaments; these are 0.18, 0.15, 0.03, and 0.35 km s $^{-1}$ pc $^{-1}$ for TMC-1F1 to F4, respectively. Then the virial mass per unit length of the sub-filaments could be derived with

$$M_{\text{lin,vir}} = \frac{2\sigma_{\text{H}_2,\text{total,ave}}^2}{G}, \quad (19)$$

where $\sigma_{\text{H}_2,\text{total,ave}}$ is the total velocity dispersion of the hydrogen (Ostriker 1964). This is calculated as

$$\sigma_{\text{H}_2,\text{total,ave}} = \sqrt{\sigma_{\text{nth,ave}}^2 + \sigma_{\text{H}_2,\text{th,ave}}^2}. \quad (20)$$

The physical parameters of the sub-filaments are listed in Table 4. The mass of the sub-filaments are slightly above (TMC-1F1 and TMC-1F4) or below (TMC-1F2 and TMC-1F3) their virial mass per unit length but with uncertainties around 50% in the observed linear masses we can only safely say that all of them are close to equilibrium. The non-thermal support of the sub-filaments is generally 2–2.5 times higher than the thermal. We note that there is no significant difference between the parameters of the clusters (size, shape, masses) from the k -means and the other clustering methods that also derived four clusters.

4.2.3. Sub-filaments TMC-1F1, TMC-1F2, and TMC-1F4

Our results show that the main ridge of TMC-1 can be resolved into three sub-filaments. A minimal overlap of TMC-1F4 and TMC-1F2 or the presence of other small fragments is seen in the middle of the ridge in the presence of secondary line components. Otherwise these sub-filaments are very well separated. TMC-1F4 has the highest average turbulent velocity dispersion. The $N(\text{NH}_3(1, 1))$ distribution peaks inside TMC-1F1 (see the upside down triangle in Fig. 5b), which is also the position of the maxima of $N_p(\text{NH}_3)$ and $n(\text{H}_2)$ (see Table 2).

Snell et al. (1982) observed TMC-1 in C³⁴S(1–0) and (2–1) and divided it into six fragments ([SLF82] A to F), many of which were overlapping. The position, extent, and velocity of [SLF82] B and [SLF82] C roughly correspond to TMC-1F4. Their [SLF82] A, [SLF82] D, and [SLF82] F coincide with the positions of TMC-1F1 and TMC-1F2, and the velocities of [SLF82] D and [SLF82] F are also similar to that of TMC-1F1 and TMC-1F2.

Hirahara et al. (1992) found six cores on CCS emission maps of TMC-1. In the northern region, their core A and B (TMC-1B) coincide with the northern region of our TMC-1F1 and

Table 5. Our sub-filaments and the coinciding objects.

Sub-filament	Other IDs
TMC-1F1	[SLF82] A ¹ , [SLF82] D ¹ , [SLF82] F ¹ , core A ² , TMC-1B ²
TMC-1F2	[SLF82] A ¹ , [SLF82] D ¹ , [SLF82] F ¹
TMC-1F3	[SLF82] A ¹ , TMC-1X ²
TMC-1F4	[SLF82] B ¹ , [SLF82] C ¹ , core C ² , TMC-1D ²

References. ⁽¹⁾ Snell et al. (1982); ⁽²⁾ Hirahara et al. (1992).

the overlapping section of TMC-1F3, which we detected with the secondary line components. They divide our TMC-1F4 into two parts, namely core C and D (TMC-1D) based on their CCS channel maps. We note however that they appear as one object with some structure on their CCS position-velocity and HC₃N integrated intensity maps. Their core E (TMC-1E) is positioned south-east of the southern edge of TMC-1F4, where we did not detect significant NH₃(1, 1) emission. Associations of the sub-filaments and these formally known parts of TMC-1 are in Table 5.

Pratap et al. (1997) mapped TMC-1 in SO(3₂–2₁), CS(2–1), and HC₃N(4–3) and (10–9). Their SO peak is $\approx 1'$ north-east from our NH₃ peak in TMC-1F2, which we also recognize in their SO channel maps. TMC-1F4 appears as an elongated maximum of the HC₃N(4–3) and (10–9) integrated intensity (see Fig. 7 in Pratap et al. 1997). A secondary maximum is seen in HC₃N integrated intensity maps inside our TMC-1F1. The relative density maxima of sulfur and nitrogen-bearing molecules are in the northern sub-filaments, whereas TMC-1F4 is the relative density maximum region of the carbon-chain molecules. The cyanopolyyne peak TMC-1 CP is roughly in the middle of TMC-1F4.

TMC-1F1 and TMC-1F4 are associated with two and one 350 μm *Herschel* point sources, respectively. We cannot exclude that the higher T_{kin} measured at position #1 (see Table 2) is related to the nearby *Herschel* point source, although we do not measure an increased temperature at position #2. The nature of the *Herschel* FIR point sources, other NIR point sources in the region (see e.g. Tóth et al. 2004) and their relation to the ISM will be discussed in future work.

We note that the HAC method with McQuitty's agglomeration criterion defines 12 sub-filaments in the main ridge (M1 to M12 in Fig. C.2a). None of those have the same position, velocity, and extent as any of the Snell et al. (1982) or Hirahara et al. (1992) fragments, but the fragment M4 that

coincides with the $N_p(\text{NH}_3)$ and $N(\text{H}_2)$ peak region, is located close to fragment A by Hirahara et al. (1992), although it has a greater extent.

4.2.4. Sub-filament TMC-1F3 and the Class I protostar IRAS 04381+2540

TMC-1F3 is separated from the main filament both in position and velocity. The velocity and position of a CS local maximum, Snell et al. (1982) fragment A, roughly corresponds to TMC-1F3. Their Fig. 8 also indicates that TMC-1F3 partly overlaps with the main ridge. The displacement of the CS peak may be partly due to their low resolution ($HPBW \approx 2'$) or a chemistry effect. TMC-1F3 corresponds to TMC-1X (identified as core X in the CCS(4–3) position-velocity maps) of Hirahara et al. (1992). TMC-1F3 also appears at similar velocities but with different extent in their $\text{HC}_3\text{N}(5–4)$ integrated intensity and $\text{NH}_3(1, 1)$ channel maps. The overlap with the main ridge is also apparent in their figures. One may recognize TMC-1F3 in the CS and SO integrated intensity maps of Pratap et al. (1997).

TMC-1F3 was also located by Yıldız et al. (2015) in ^{12}CO with a velocity of 5.2 km s^{-1} . The position of TMC-1F3 coincides with that of the protostar IRAS 04381+2540 (Beichman et al. 1986) that is known to drive a bipolar outflow detected in ^{12}CO (Bontemps et al. 1996) and H_2 (Gomez et al. 1997). Its disk was resolved and a YSO mass of $0.54 M_\odot$ was estimated by Harsono et al. (2014). Apai et al. (2005) described it as a binary in formation with a jet and substellar companion and showed that the northern outflow cavity is opening towards us. According to Yıldız et al. (2013), the velocity of the YSO is $v_{\text{LSR}} = 5.2 \text{ km s}^{-1}$, which agrees with that of our NH_3 clump ($5.2–5.4 \text{ km s}^{-1}$). The non-thermal velocity dispersion of TMC-1F3 is the second highest and its thermal velocity dispersion is the highest of all four sub-filaments.

From the positional coincidence and velocity one may assume that the Class I protostar is related to the sub-filament. The relatively high average kinetic temperature (13 K, see Table 4) of TMC-1F3 we interpret as a possible signature of the heating by IRAS 04381+2540. In fact the highest kinetic and dust temperatures were measured at the YSO position (see position #10 with $T_{\text{kin}} \approx 14 \text{ K}$ and $T_{\text{dust}} = 11.4 \text{ K}$ in Table 2).

The highest HFS linewidth of the $\text{NH}_3(1, 1)$ primary line components is at position #8 and around it ($\Delta v = 0.39 \text{ km s}^{-1}$, $T_{\text{kin}} = 10.6 \text{ K}$, see Table 2) in TMC-1F1. The appearance of increased linewidths could be explained by the outflow of IRAS 04381+2540 (120'' south-west of #8) releasing mechanical energy towards the main ridge. The outflow is very compact, however, with a lobe extent of only around 40'' to the north and south according to Hogerheijde et al. (1998) and Yıldız et al. (2015). The northern outflow lobe is blue-shifted, opening towards the observer, thus the protostar has to be positioned behind the main ridge and the outflow should extend to at least 120'' (projected distance of position #8) to affect the main ridge. Hogerheijde et al. (1998) calculated a dynamical age of 2100 yr from a $^{12}\text{CO}(3–2)$ outflow extent of 0.02 pc and outflow velocity of 9.3 km s^{-1} . One may however assume that a jet with about an order of magnitude higher speed propagated to more than 0.1 pc in the same time interval. We note here that velocities up to 70 km s^{-1} were also observed at very low-mass YSOs (see e.g. ISO 143 in Joergens et al. 2012). Our assumption from the above geometry is that the protostar IRAS 04381+2540 is embedded in TMC-1F3, which may be located around 0.1 pc behind the main ridge.

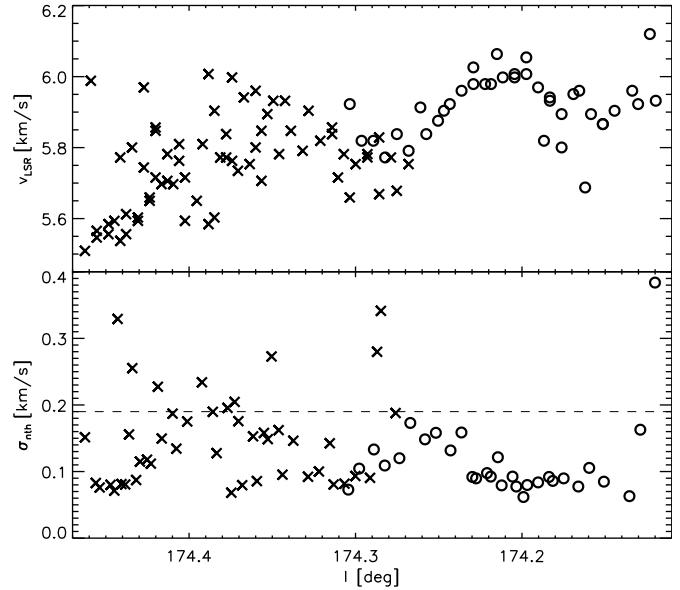


Fig. 7. Central velocity and non-thermal velocity dispersion of the $\text{NH}_3(1, 1)$ primary components along TMC-1F4 (crosses) and TMC-1F1 (circles) from left to right. The sound speed in the two clumps are indicated with a dashed line in the lower panel.

4.2.5. Fray and fragment?

The primary $\text{NH}_3(1, 1)$ line component v_{LSR} velocity distribution in TMC-1F1 and TMC-1F4 is continuous, as seen in Fig. 7. A velocity gradient of $\approx 0.45 \text{ km s}^{-1} \text{ pc}^{-1}$ is seen from SE to NW. An oscillating behaviour can be observed above the gradient in both sub-filaments, but with a higher dispersion in TMC-1F4, and the wavelength of the periodical change in TMC-1F1 from a sinusoidal approximation is $\approx 0.3 \text{ pc}$. Similar velocity oscillations were observed in several filaments in L1517 (Hacar & Tafalla 2011) together with sinusoidal column density oscillations as evidence of core formation. The non-thermal velocity dispersions also show oscillations and are mostly subsonic, since the typical sound speed, $c_s = \sqrt{\frac{k_B T_{\text{kin}}}{\mu m_H}} = 0.19 \text{ km s}^{-1}$ in the two sub-filaments. Supersonic points could be explained by confusion of multiple velocity components or outflows from embedded sources inside the sub-filaments. The pattern of velocity dispersions may indicate that the sub-filaments are decoupled from the large-scale turbulent velocity field and the turbulent dissipation in the cloud occurred at the scale of the sub-filaments (Tafalla & Hacar 2015).

Comparing the position-velocity diagram of TMC-1F1 and TMC-1F4, the latter looks less regular. Analysing a dust continuum map of TMC-1, Suutarinen et al. (2011) suggested that the southern half of the ridge has an extensive envelope, in contrast to the northern, i.e. the NH_3 peak region. They explain this as a difference in dynamical age: the northern region is more compressed. Based on time dependent modelling of the observed variation of chemistry they also concluded that the south-eastern end of the cloud is also chemically less evolved than the north-eastern end. The difference in the chemical ages of TMC-1F1 and TMC-1F4 was also indicated by Berczik et al. (2015).

Peng et al. (1998) studied the fragmentation in a $8' \times 8'$ region around TMC-1 CP in TMC-1F4 with CCS measurements. They found 45 very small fragments with sizes of $0.02–0.04 \text{ pc}$ and masses of $0.04–0.6 M_\odot$; five of those fragments were gravitationally bound. The northern end of TMC-1F4 was resolved

into cores with sizes of 0.03–0.06 pc and masses of 0.03–2 M_{\odot} by Langer et al. (1995), based on high resolution single-dish and interferometric observations of CCS and CS.

Hierarchical fragmentation was found by Hacar et al. (2013) and Tafalla & Hacar (2015) in the nearby L1495/B2013 complex. The fray and fragment filament evolution scenario proposed by these authors is the following: The filamentary cloud first fragments into intertwined, velocity-coherent filaments, so-called fibers. If the fiber can accumulate enough mass to become gravitationally unstable, it can further fragment into chains of closely spaced cores in an almost quasi-static way. As seen in Table 4, the length of TMC-1F1, TMC-1F2, and TMC-1F4 is similar that of the fibers in L1495/B2013. The typical length of the fibers of Hacar et al. (2013) is 0.5 pc (ranging from 0.2 pc to 1.2 pc). TMC-1F2 even has a similar linear mass as the fiber [HTK2013] 32 of Hacar et al. (2013). However their typical 0.5 pc-long fibers have a factor of two lower mass than the sub-filaments of TMC-1. Some of their fibers are unstable and fragmented into chains of cores.

The velocity and linewidth distributions in TMC-1F1 in the northern region of TMC-1 is very similar to that of the fibers described by Tafalla & Hacar (2015). The southern section of TMC-1 with TMC-1F4 is said to be structurally and chemically less evolved than the northern region of TMC-1, also the velocity dispersion is much higher in TMC-1F4 than in TMC-1F1. However, there are already small, gravitationally bound cores inside TMC-1F4. It is not clear whether fray and fragment is the process shaping TMC-1 because we cannot simply say that the filament first collapsed into velocity-coherent fibers and then formed cores inside them. One has to see the distribution of small-scale structures all along TMC-1 and discuss again the fray and fragment scheme.

5. Conclusions

Our work contributes to the study of HCL 2, focussing on the TMC-1 region based on *Herschel* SPIRE FIR and high S/N NH₃ 1.3 cm line mapping. We found TMC-1 as one of the coldest and densest regions of HCL 2 (along with TMC-1C and HCL 2B) from our $N(\text{H}_2)$ and T_{dust} calculations.

Our *Herschel*-based $N(\text{H}_2)$ shows values above 10^{22} cm^{-2} with two local maxima along a narrow ridge. The northern maximum is the peak $N(\text{H}_2)$ position with $3.3 \times 10^{22} \text{ cm}^{-2}$. There is a third, more separated peak south-west from this. T_{dust} is almost constant along the higher density inner region of the ridge.

Fitting multiple HFS line profiles to the NH₃(1, 1) spectra we identified multiple velocity components in the northern region and calculated low turbulent velocity dispersions (0.1 – 0.2 km s^{-1}). The derived $N(\text{NH}_3(1, 1))$ distribution follows the *Herschel*-based $N(\text{H}_2)$ distribution well. We define a new NH₃-peak in an offset of $\approx 1'$ relative from TMC-1(NH3). The NH₃-based $n(\text{H}_2)$ are lower than former C³⁴S and HC₃N-based values by a factor of 5–6.

Based on the observed velocity and column density variations we partitioned TMC-1 using the *NbClust* package of R. The four derived sub-filaments have masses of $20– $40 M_{\odot}$, three of these sub-filaments are very elongated (TMC-1F1, TMC-1F2, TMC-1F4) and all of them close to gravitational equilibrium. TMC-1F1 has the NH₃ peak and TMC-1F4 the cyanopolyyne peak inside. TMC-1F3 partly overlaps with TMC-1F1, and the Class I protostar IRAS 04381+2540 is apparently embedded into this sub-filament. The protostellar outflow is probably affecting TMC-1F1. The fray and fragment process is one possible$

scenario for the structuring of TMC-1, but the small-scale structure of TMC-1F1 should be investigated to ascertain this scenario.

A more detailed description and analysis could be given with a high resolution survey of the whole TMC-1 in carbonous and nitrogen bearing molecules (e.g. CCS and NH₃ lines) and, subsequently, using chemical and radiative transfer models in the interpretation.

Acknowledgements. This research was partly supported by the OTKA grants K101393 and NN-111016 and it was supported by the Momentum grant of the MTA CSFK Lendület Disk Research Group. The research leading to these results has received funding from the European Commission Seventh Framework Programme (FP/2007-2013) under grant agreement No 283393 (RadioNet3). V.-M.P. acknowledges the support of Academy of Finland grant 250741. Valuable discussions with Jorma Harju, Malcolm Walmsley, and Dimitris Stamatellos are acknowledged. We thank the anonymous referee for the careful reading of our manuscript and the valuable comments and suggestions. *Herschel* was an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with participation from NASA. SPIRE was developed by a consortium of institutes led by Cardiff Univ. (UK) and including Univ. Lethbridge (Canada); NAOC (China); CEA, LAM (France); IFSI, Univ. Padua (Italy); IAC (Spain); Stockholm Observatory (Sweden); Imperial College London; RAL, UCL-MSSL, UKATC, Univ. Sussex (UK); Caltech, JPL, NHSC, Univ. Colorado (USA). This development was supported by national funding agencies: CSA (Canada); NAOC (China); CEA, CNES, CNRS (France); ASI (Italy); MCINN (Spain); SNSB (Sweden); STFC (UK); and NASA (USA). PACS was developed by a consortium of institutes led by MPE (Germany) and including UVIE (Austria); KUL, CSL, IMEC (Belgium); CEA, OAMP (France); MPIA (Germany); IFSI, OAP/AOT, OAA/CAISMI, LENS, SISSA (Italy); IAC (Spain). This development has been supported by the funding agencies BMVIT (Austria), ESA-PRODEX (Belgium), CEA/CNES (France), DLR (Germany), ASI (Italy), and CICT/MCT (Spain). Based on observations with the 100-m telescope of the MPIfR (Max-Planck-Institut für Radioastronomie) at Effelsberg, the authors thank their high level of technical support and assistance.

References

- André, P., Men'shchikov, A., Bontemps, S., et al. 2010, *A&A*, **518**, L102
- Apai, D., Tóth, L. V., Henning, T., et al. 2005, *A&A*, **433**, L33
- Bally, J., Langer, W. D., Stark, A. A., & Wilson, R. W. 1987, *ApJ*, **312**, L45
- Beckwith, S. V. W., Sargent, A. I., Chini, R. S., & Guesten, R. 1990, *AJ*, **99**, 924
- Beichman, C. A., Myers, P. C., Emerson, J. P., et al. 1986, *ApJ*, **307**, 337
- Berczik, P., Bertsyk, P., Toth, V., & Baranyai, A. 2015, *IAU General Assembly, Meeting 29*, 2252605
- Bergin, E. A., Alves, J., Huard, T., & Lada, C. J. 2002, *ApJ*, **570**, L101
- Bohlin, R. C., Savage, B. D., & Drake, J. F. 1978, *ApJ*, **224**, 132
- Bontemps, S., Andre, P., Terebey, S., & Cabrit, S. 1996, *A&A*, **311**, 858
- Cambrésy, L. 1999, *A&A*, **345**, 965
- Cernicharo, J., & Guélin, M. 1987, *A&A*, **176**, 299
- Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. 2014, *J. Statistical Software*, **61**, 1
- Danby, G., Flower, D. R., Valiron, P., Schilke, P., & Walmsley, C. M. 1988, *MNRAS*, **235**, 229
- Dobashi, K., Uehara, H., Kandori, R., et al. 2005, *PASJ*, **57**, 1
- Dunn, J. C. 1974, *J. Cybernetics*, **4**, 95
- Elias, J. H. 1978, *ApJ*, **224**, 857
- Gaida, M., Ungerechts, H., & Winnewisser, G. 1984, *A&A*, **137**, 17
- Goldsmith, P. F. 2001, *ApJ*, **557**, 736
- Goldsmith, P. F., Heyer, M., Narayanan, G., et al. 2008, *ApJ*, **680**, 428
- Gomez, M., Whitney, B. A., & Kenyon, S. J. 1997, *AJ*, **114**, 1138
- Gower, J. C. 1966, *Biometrics*, **23**, 123
- Griffin, M. J., Abergel, A., Abreu, A., et al. 2010, *A&A*, **518**, L3
- Hacar, A., & Tafalla, M. 2011, *A&A*, **533**, A34
- Hacar, A., Tafalla, M., Kauffmann, J., & Kovács, A. 2013, *A&A*, **554**, A55
- Halkidi, M., Vazirgiannis, M., & Batistakis, Y. 2000, in *Principles of Data Mining and Knowledge Discovery*, eds. D. Zighed, J. Komorowski, & J. Zytkow (Berlin, Heidelberg: Springer), *Lect. Not. Comput. Sci.*, **1910**, 265
- Harju, J., Walmsley, C. M., & Wouterloot, J. G. A. 1993, *A&AS*, **98**, 51
- Harsono, D., Jørgensen, J. K., van Dishoeck, E. F., et al. 2014, *A&A*, **562**, A77
- Hatchell, J., Richer, J. S., Fuller, G. A., et al. 2005, *A&A*, **440**, 151
- Heiles, C. E. 1968, *ApJ*, **151**, 919
- Hirahara, Y., Suzuki, H., Yamamoto, S., et al. 1992, *ApJ*, **394**, 539
- Ho, P. T. P., & Townes, C. H. 1983, *ARA&A*, **21**, 239

- Hogerheijde, M. R., van Dishoeck, E. F., Blake, G. A., & van Langevelde, H. J. 1998, [ApJ](#), **502**, 315
- Jain, A. K., Murty, M. N., & Flynn, P. J. 1999, [ACM Comput. Surv.](#), **31**, 264
- Joergens, V., Kopytova, T., & Pohl, A. 2012, [A&A](#), **548**, A124
- Juvela, M., Ristorcelli, I., Pagani, L., et al. 2012, [A&A](#), **541**, A12
- Juvela, M., Demyk, K., Doi, Y., et al. 2015a, [A&A](#), **584**, A94
- Juvela, M., Ristorcelli, I., Marshall, D. J., et al. 2015b, [A&A](#), **584**, A93
- Kirk, J. M., Ward-Thompson, D., Palmeirim, P., et al. 2013, [MNRAS](#), **432**, 1424
- Kramer, C., Stutzki, J., Rohrig, R., & Cornelissen, U. 1998, [A&A](#), **329**, 249
- Langer, W. D., Velusamy, T., Kuiper, T. B. H., et al. 1995, [ApJ](#), **453**, 293
- Little, L. T., MacDonald, G. H., Riley, P. W., & Matheson, D. N. 1979, [MNRAS](#), **189**, 539
- MacQueen, J. 1967, in Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1: Statistics (Berkeley, Calif.: University of California Press), 281
- Malinen, J., Juvela, M., Rawlings, M. G., et al. 2012, [A&A](#), **544**, A50
- McElroy, D., Walsh, C., Markwick, A. J., et al. 2013, [A&A](#), **550**, A36
- McQuitty, L. L. 1966, [Educational and Psychological Measurement](#), **26**, 825
- Milligan, G., & Cooper, M. 1985, [Psychometrika](#), **50**, 159
- Mizuno, A., Onishi, T., Yonekura, Y., et al. 1995, [ApJ](#), **445**, L161
- Murtagh, F., & Heck, A. 1987, Multivariate Data Analysis, [Astrophys. Space Sci. Libr.](#), **131**
- Murtagh, F., & Legendre, P. 2014, [J. Classification](#), **31**, 274
- Myers, P. C., Linke, R. A., & Benson, P. J. 1983, [ApJ](#), **264**, 517
- Narayanan, G., Heyer, M. H., Brunt, C., et al. 2008, [ApJS](#), **177**, 341
- Nutter, D., Kirk, J. M., Stamatellos, D., & Ward-Thompson, D. 2008, [MNRAS](#), **384**, 755
- Olano, C. A., Walmsley, C. M., & Wilson, T. L. 1988, [A&A](#), **196**, 194
- Onishi, T., Mizuno, A., Kawamura, A., Ogawa, H., & Fukui, Y. 1996, [ApJ](#), **465**, 815
- Onishi, T., Mizuno, A., Kawamura, A., Tachihara, K., & Fukui, Y. 2002, [ApJ](#), **575**, 950
- Ostriker, J. 1964, [ApJ](#), **140**, 1056
- Padoan, P., Cambrésy, L., & Langer, W. 2002, [ApJ](#), **580**, L57
- Pagani, L., Pardo, J.-R., Apponi, A. J., Bacmann, A., & Cabrit, S. 2005, [A&A](#), **429**, 181
- Pagani, L., Bacmann, A., Cabrit, S., & Vastel, C. 2007, [A&A](#), **467**, 179
- Pagani, L., Lefèvre, C., Juvela, M., Pelkonen, V.-M., & Schuller, F. 2015, [A&A](#), **574**, L5
- Peng, R., Langer, W. D., Velusamy, T., Kuiper, T. B. H., & Levin, S. 1998, [ApJ](#), **497**, 842
- Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, [A&A](#), **518**, L1
- Planck Collaboration I. 2011, [A&A](#), **536**, A1
- Planck Collaboration XIX. 2011, [A&A](#), **536**, A19
- Planck Collaboration XXIII. 2011, [A&A](#), **536**, A23
- Planck Collaboration XXV. 2011, [A&A](#), **536**, A25
- Planck Collaboration XXVIII. 2016, [A&A](#), in press, DOI: [10.1051/0004-6361/201525819](#)
- Poglitsch, A., Waelkens, C., Geis, N., et al. 2010, [A&A](#), **518**, L2
- Pratap, P., Dickens, J. E., Snell, R. L., et al. 1997, [ApJ](#), **486**, 862
- Rosolowsky, E. W., Pineda, J. E., Kauffmann, J., & Goodman, A. A. 2008, [ApJ](#), **679**, 1338
- Rousseauw, P. J. 1987, [J. Comput. Appl. Math.](#), **20**, 53
- Sakai, N., Maezawa, H., Sakai, T., Menten, K. M., & Yamamoto, S. 2012, [A&A](#), **546**, A103
- Schloerb, F. P., & Snell, R. L. 1984, [ApJ](#), **283**, 129
- Schloerb, F. P., Snell, R. L., & Young, J. S. 1983, [ApJ](#), **267**, 163
- Schneider, S., & Elmegreen, B. G. 1979, [ApJS](#), **41**, 87
- Seber, G. 2009, Multivariate Observations, Wiley Series in Probability and Statistics (Wiley)
- Snell, R. L., Langer, W. D., & Frerking, M. A. 1982, [ApJ](#), **255**, 149
- Sokal, R. R., & Michener, C. D. 1958, [University of Kansas Scientific Bulletin](#), **28**, 1409
- Sorensen, T. 1948, [Biologiske Skrifter](#), **5**, 1
- Steinley, D. 2004, in Classification, Clustering, and Data Mining Applications, eds. D. Banks, F. McMorris, P. Arabie, & W. Gaul, Studies in Classification, Data Analysis, and Knowledge Organisation (Berlin, Heidelberg: Springer), 53
- Stutzki, J., & Guesten, R. 1990, [ApJ](#), **356**, 513
- Suutarien, A., Geppert, W. D., Harju, J., et al. 2011, [A&A](#), **531**, A121
- Suzuki, H., Yamamoto, S., Ohishi, M., et al. 1992, [ApJ](#), **392**, 551
- Tafalla, M., & Hacar, A. 2015, [A&A](#), **574**, A104
- Tafalla, M., Myers, P. C., Caselli, P., & Walmsley, C. M. 2004, [A&A](#), **416**, 191
- Tauber, J. A., Mandolesi, N., Puget, J.-L., et al. 2010, [A&A](#), **520**, A1
- Terebey, S., & Baud, B. 1980, [BAAS](#), **12**, 861
- Theodoridis, S., & Koutroumbas, K. 2006, Pattern Recognition, Pattern Recognition Series (Elsevier Science)
- Tölle, F., Ungerechts, H., Walmsley, C. M., Winnewisser, G., & Churchwell, E. 1981, [A&A](#), **95**, 143
- Tóth, L. V., Haas, M., Lemke, D., Mattila, K., & Onishi, T. 2004, [A&A](#), **420**, 533
- Tóth, L. V., Zahorecz, S., Marton, G., et al. 2016, Proc. IAU Symp. 315, eds. P. Jabolonka, F. Van der Tak, & P. Andre, submitted
- Ungerechts, H., & Thaddeus, P. 1987, [ApJS](#), **63**, 645
- Ungerechts, H., Winnewisser, G., & Walmsley, C. M. 1986, [A&A](#), **157**, 207
- Ward, J. H. 1963, [J. Am. Stat. Ass.](#), **58**, 236
- Williams, J. P., de Geus, E. J., & Blitz, L. 1994, [ApJ](#), **428**, 693
- Wood, D. O. S., Myers, P. C., & Daugherty, D. A. 1994, [ApJS](#), **95**, 457
- Yıldız, U. A., Kristensen, L. E., van Dishoeck, E. F., et al. 2013, [A&A](#), **556**, A89
- Yıldız, U. A., Kristensen, L. E., van Dishoeck, E. F., et al. 2015, [A&A](#), **576**, A109
- Zijlstra, A. A., van Hoof, P. A. M., & Perley, R. A. 2008, [ApJ](#), **681**, 1296

Appendix A: Column density calculation from *Herschel* SPIRE measurements

A.1. Determination of the background level

The background/foreground subtraction used on the *Herschel* SPIRE images after calibration with HIPE is described in Sect. 2.5. The difference between T_{dust} without and with this correction is less than 0.5 K, which means a 1–3% difference in $N(\text{H}_2)$ in the dense regions of TMC-1, increasing to 20–30% in the outer regions. The contour value used to determine the background/foreground level is very similar to the baseline fitted by Malinen et al. (2012) when deriving radial density profile for TMC-1 and it was mentioned by them that a background subtraction may be needed when determining filament properties.

Another way of determining the background level is calculating $N(\text{H}_2)$ distribution without any correction, then taking the contour above which we were able to detect NH_3 . This contour is around $0.7 \times 10^{22} \text{ cm}^{-2}$ in our case and it marks the boundaries of the dense ISM. We created T_{dust} and $N(\text{H}_2)$ with subtracting the average intensities found at this contour from the SPIRE images. The difference between the T_{dust} values calculated with the method in the paper and with the method described above is around 1 K. It means a 10–20% difference in $N(\text{H}_2)$ at the dense region of the ridge, quickly increasing to more than half a magnitude difference close to the marked boundary.

A.2. The effect of the spectral index

Early results from *Planck* showed that the dust opacity spectral index, β is around 1.8 on the all-sky map observed by *Planck* HFI (Planck Collaboration XIX 2011). Later the statistical analysis of the properties of the Cold Clump Catalogue of *Planck* Objects using *Planck* and IRAS observations determined that in the clumps β varies from 1.4 to 2.8 with a median value around 2.1 and anti-correlates with temperature (Planck Collaboration XXIII 2011). Planck Collaboration XXV (2011) examined the Taurus region and found $\beta > 1.8$ in structures with temperatures below 14 K. Juvela et al. (2012) used an average value of 2 when determining the T_{dust} and $N(\text{H}_2)$ on the fields observed during the *Herschel* open time key programme Galactic Cold Cores. Recently Juvela et al. (2015b) examined the sub-millimeter dust opacity and gave the average value of β close to 1.9 and they indicated that it can be even higher in the coldest regions. The variations of β were described in detail by Juvela et al. (2015a). They derived β with different combinations of observations and concluded that it anti-correlates with T_{dust} , correlates with $N(\text{H}_2)$ and galactic latitude and its value depends on the wavelength of the observations used in the analyses. Using *Herschel* measurements $\beta = 2$ is appropriate for dense clumps and can reach 2.2 in the coldest regions.

We used $\beta = 2$ in the calculations in this paper but estimated the effect of changing the spectral index between 1.8 and 2.2. The differences in T_{dust} originating from this do not reach 10%, which gives a <30% deviation in the derived $N(\text{H}_2)$ distribution.

Appendix B: Estimating $N_p(\text{NH}_3)$ from the secondary $\text{NH}_3(1, 1)$ velocity components

The physical parameters of the gas were calculated using the primary component of the $\text{NH}_3(1, 1)$ line, however, the secondary components also contribute to the total $N_p(\text{NH}_3)$ and

$N(\text{NH}_3(1, 1))$. This was estimated by first calculating the fraction of $N(\text{NH}_3(2, 2))$ compared to the total $N_p(\text{NH}_3)$ from the primary velocity components with Eqs. (5) and (6). This gives typically 3–6% for the fraction of $N(\text{NH}_3(2, 2))$ in the total $N_p(\text{NH}_3)$. We then derived $N(\text{NH}_3(1, 1))$ from the secondary components with Eq. (5). Since the HFS of the secondary components could not be fitted well (the relative error of τ was 50% even in the best HFS line profile fit), we assumed an optical depth of $\tau = 0.1$ as upper limit, which was derived from the HFS line profile fit of the noise-weighted average of all the secondary line spectra in the northern region of TMC-1. Then assuming that the contribution of the molecules in the $\text{NH}_3(2, 2)$ state is 10% of this (as a conservative estimate), we computed the total $N_p(\text{NH}_3)$ from the secondary velocity components (see Table 2 in Col. 10).

Appendix C: Clustering methods

Clustering methods are categorized based on cluster models, i.e. how one defines a cluster. Typical models are, for example connectivity models, where clusters are based on distance connectivity (hierarchical clustering), centroid models, where each cluster is represented by a single mean vector (k -means clustering), distribution, density, or graph-based models. For each of the types there are several subtypes and different algorithms to find clusters in a dataset (Jain et al. 1999; Halkidi et al. 2000).

C.1. Clustering of molecular clouds

A well-known routine for identifying local maxima as peaks of clumps and following them to lower levels is CLUMPFIND (Williams et al. 1994). CLUMPFIND works by searching through the data for distinct objects identified by closed contours from the highest contour levels to the lowest. Contours that surround only one peak are assigned to the corresponding clump, blended contours that surround more than one peak are split up using a friends-of-friends algorithm. The conservative variant of CLUMPFIND, CSAR (Cardiff Source-finding AlgoRithm; Kirk et al. 2013) can be used in a hierarchical mode. GAUSS-CLUMP (Stutzki & Guesten 1990; Kramer et al. 1998) iteratively fits 3D Gaussian functions to a datacube to derive sources. The dendrogram technique (Rosolowsky et al. 2008) combines the statistical approach with segmenting the data into physically relevant structures while preserving and characterizing the hierarchy of emission isosurfaces.

We included a decomposition of TMC-1 with CLUMPFIND in Fig. C.1 to show that the resulting clumps strongly depend on the contouring threshold and stepsize used as an input. We applied CLUMPFIND with a contouring threshold of $3\sigma_N$ and stepsizes of 1, 0.5, and $0.2\sigma_N$ (where σ_N is the standard deviation measured on the $N(\text{H}_2)$ map) and depending on the stepsize, resolved the filament into 4, 5, and 7 clumps. This analysis shows that by decreasing the stepsize, we can effectively derive clumps around every local $N(\text{H}_2)$ maxima and so clump-finding with CLUMPFIND and similar methods on the $N(\text{H}_2)$ map is not sufficient in the case of TMC-1.

C.2. k -means clustering

The k -means clustering (MacQueen 1967; Murtagh & Heck 1987) uses the centroid cluster model. It is an iterative method that minimizes the within-cluster sum of squares for a given number of clusters, starting with an initial guess for cluster centres. Each observation is placed in the closest cluster, the cluster

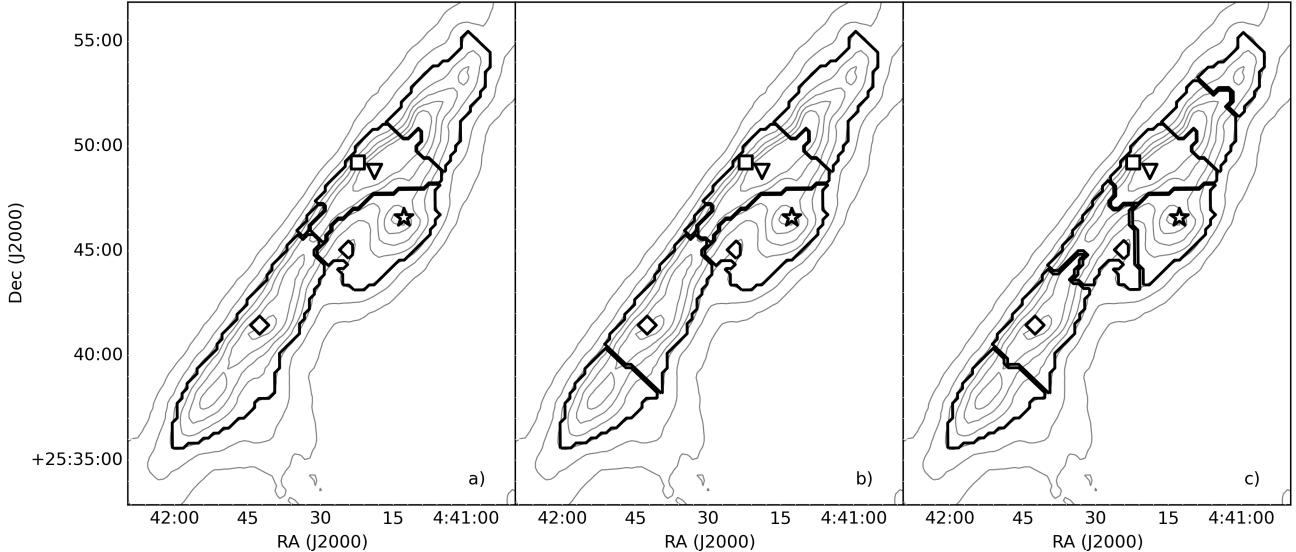


Fig. C.1. Results of the CLUMPFIND decomposition. The algorithm was run with a: 1σ ; b: 0.5σ and c: 0.2σ contouring stepsizes. $N(\text{H}_2)$ from *Herschel* and the positions of TMC-1 CP, IRAS 04381+2540, the SO peak, and our NH_3 maximum are indicated as in Fig. 5a.

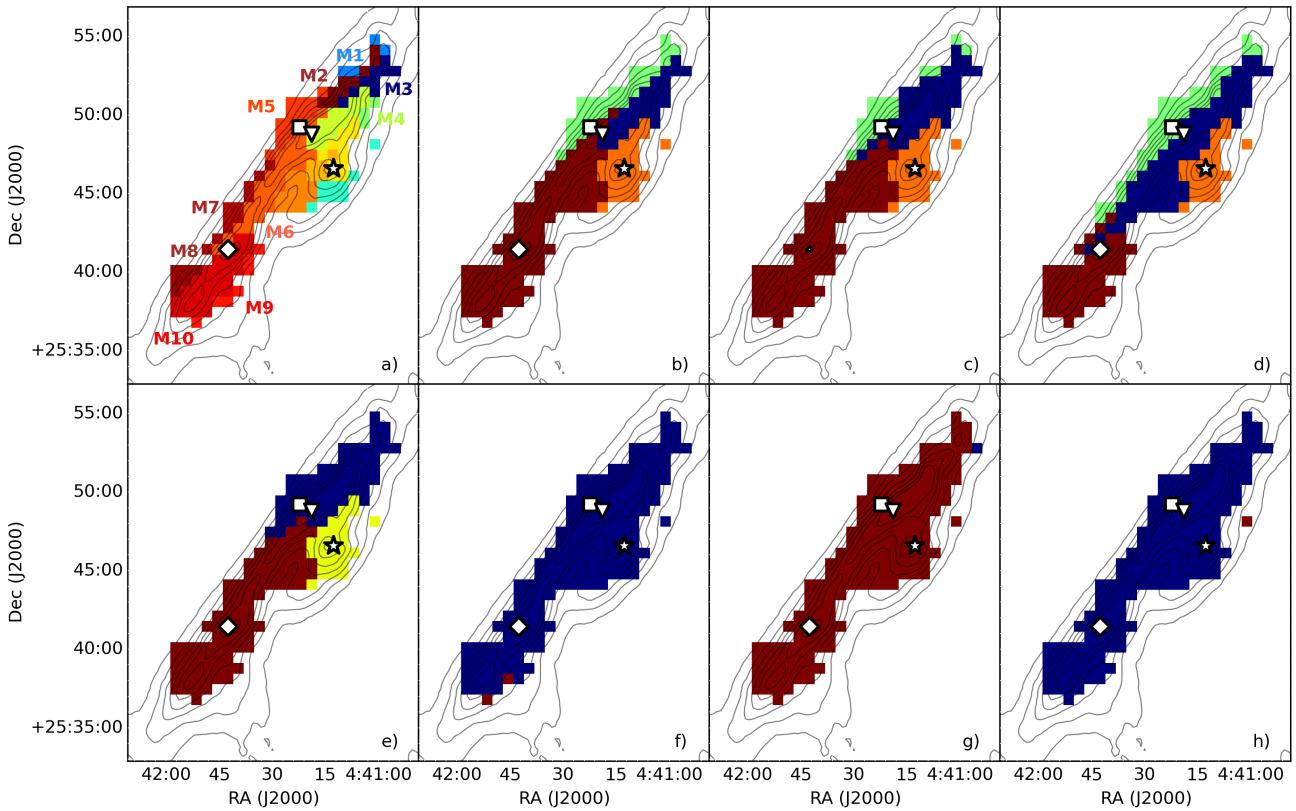


Fig. C.2. Best results of the other *NbClust* runs: HAC methods with a) McQuitty's; b) Ward's; c) Ward's second; d) the complete; e) the average; f) the median; g) the centroid and h) the single criteria. $N(\text{H}_2)$ from *Herschel* and the positions of TMC-1 CP, IRAS 04381+2540, the SO peak, and our NH_3 maximum are indicated as in Fig. 5a. The clusters in the main ridge on panel a) are numbered from M1 to M10 and discussed shortly in Sect. 4.2.3.

centres are updated, and then the process repeats itself until the cluster centres no longer change.

C.3. Hierarchical clustering

The hierarchical agglomerative clustering (HAC) uses distance connectivity and seeks to build a hierarchy, where each observation starts in its own cluster, then pairs of clusters are merged

according to an agglomeration criterion as one moves up the hierarchy. Another approach is hierarchical divisive clustering where all observations start in one cluster, and splits are performed recursively as one moves down the hierarchy.

The HAC methods require an agglomeration (or linkage) criterion. When building the hierarchy of clusters found in a dataset, the two clusters with the shortest distance between each other are merged at each step. The linkage criterion defines

the way this between-cluster distance is measured. It can be defined, for example as the distance of the elements that are the farthest away in two clusters (complete-linkage; [Sorenson 1948](#)), the shortest distance between two elements not yet belonging to the same cluster (single; [Sokal & Michener 1958](#)), the average distance of the elements inside two clusters (average; [Sokal & Michener 1958](#)), and the squared Euclidean distance between the gravity centres of two clusters (centroid; [Sokal & Michener 1958](#)). Linkage criteria also can be determined with the weighted mean of the between-cluster distances ([McQuitty; McQuitty 1966](#)), the median of the between-cluster distances (median; [Gower 1966](#)), or with minimizing within-cluster variance (Ward and its second variant; [Ward 1963](#); [Murtagh & Legendre 2014](#)).

C.4. The NbClust package

The R project for statistical computing provides a wide variety of the previously listed algorithms. We used the *NbClust* package ([Charrad et al. 2014](#)) to select the best clustering method with the most applicable input parameters to partition the TMC-1 joined *Herschel/NH₃* dataset into clusters. The *k*-means and HAC methods with eight different agglomeration criteria are both available in *NbClust*. The package has been developed to not only perform several types of clustering methods, but to evaluate the quality of clusters (the degree with which a clustering scheme fits a specific dataset) and the optimal number of clusters (which is a required input parameter for the algorithms) in a partitioning. This is the so-called cluster validity. For this purpose, a variety of validity measures (indices) were defined in the literature. The main benefit of using this package lies in the possibility of simultaneously evaluating several clustering schemes while varying the number of clusters. This helps to determine the

most appropriate number of clusters for the dataset, then the resulting ideal clustering scheme can be run by the same package.

One first has to define the distance measure to be used by the clustering methods. The distance measure defines how the distance of two vectors (elements or points in the dataset) are calculated during the clustering process. Euclidean distance is the usual square distance between two vectors, Manhattan distance measures the absolute distance and maximum distance is the greatest distance of two vectors. The Canberra, binary, and Minkowski distance measures are also available in *NbClust* ([Seber 2009](#)).

The *NbClust* algorithm can be run on a range of cluster numbers, choosing from the measures and methods described above, while cluster validity is assessed using the validity indices. Cluster validity (evaluating the quality of clusters, optimizing the number of clusters derived) can be described using three different approaches ([Theodoridis & Koutroumbas 2006](#)): external criteria (comparing the results to externally provided classes or labels), internal criteria (the information obtained within the clustering process is used to evaluate how well the results fit the data), and relative criteria (compares the results with other clustering schemes resulting from the same algorithm but with different parameters). Most of the validity measures determined for the third approach ([Milligan & Cooper 1985](#); [Dunn 1974](#); [Rousseeuw 1987](#)) were incorporated in the *NbClust* package. The validity indices combine information about within-cluster compactness, between-cluster isolation and other geometric or statistical properties of the data, and the number of elements and distance measures. See [Charrad et al. \(2014\)](#) and references therein for the detailed description of *NbClust* and the definitions of distance measures, linkage criteria, and cluster validity indices.