Astronomy & Astrophysics manuscript no. pdf˙bg˙04022014
March 13, 2014

Understanding star formation in molecular clouds

I. A universal probability distribution of column densities ?

N. Schneider[1,2], V. Ossenkopf[3], T. Csengeri[4], R. Klessen[5], C. Federrath[5,6], P. Tremblin[7], P. Girichidis[8], S. Bontemps[1,2], and Ph. André[9]

[1] Univ. Bordeaux, LAB, UMR 5804, 33270 Floirac, France
[2] CNRS, LAB, UMR 5804, 33270 Floirac, France
[3] I. Physikalisches Institut, Universität zu Köln, Zülpicher Straße 77, 50937 Köln, Germany
[4] Max-Planck Institut für Radioastronomie, Auf dem Hügel, Bonn, Germany
[5] Universität Heidelberg, Zentrum für Astronomie, Institut für Theoretische Astrophysik, Albert-Ueberle Str. 2, 69120 Heidelberg, Germany
[6] Monash Centre for Astrophysics, School of Mathematical Sciences, Monash University, Victoria 3800, Australia
[7] Astrophysics Group, University of Exeter, EX4 4QL Exeter, UK
[8] Max-Planck Institut für Astrophysik, 85741, Garching, Germany
[9] IRFU/SAp CEA/DSM, Laboratoire AIM CNRS - Université Paris Diderot, 91191 Gif-sur-Yvette, France

March 13, 2014

ABSTRACT

Column density maps of molecular clouds are one of the most important observables in the context of molecular cloud- and star-formation (SF) studies. With *Herschel* it is now possible to reveal rather precisely the column density of dust, which is the best tracer of the bulk of material in molecular clouds. However, line-of-sight (LOS) contamination from fore- or background clouds can lead to an overestimation of the dust emission of molecular clouds, in particular for distant clouds. This implies too high values for column density and mass, and a misleading interpretation of probability distribution functions (PDFs) of the column density. In this paper, we demonstrate by using observations and simulations how LOS contamination affects the PDF. We apply a first-order approximation (removing a constant level) to the molecular clouds of Auriga and Maddalena (low-mass star-forming), and Carina and NGC3603 (both high-mass SF regions). In perfect agreement with the simulations, we find that the PDFs become broader, the peak shifts to lower column densities, and the power-law tail of the PDF for higher column densities flattens after correction. All corrected PDFs have a lognormal part for low column densities with a peak at A_v ∼2 and a deviation point (DP) from the lognormal at A_v(DP)∼4–5 (corresponding to a surface density Σ of ∼45 $M_\odot$ pc^{-2}). For higher column densities, all PDFs have a power-law tail with an average slope that corresponds to an exponent $\alpha = 1.9\pm0.2$ for an equivalent spherical density distribution $\rho \propto r^{-\alpha}$ consistent with a structure dominated by self-gravity (local free-fall of individual cores and global collapse of gas on larger scales, such as filaments). Our PDF study suggests that there is a common, universal column density break at A_v ∼4–5 for all cloud types where the transition between supersonic turbulence and self-gravity takes place.

Key words. interstellar medium: clouds

1. Introduction

The large-scale far-infrared (FIR) photometric observations of *Herschel*[1] (Pilbratt et al. 2010) have an unprecedented sensitivity and high angular resolution (∼6″ to ∼37″) that provide an exceptional database to better understand the composition and structure of the interstellar medium. Several large programs such as the *Herschel* Gould Belt Survey (HGBS, André et al. 2010), *Herschel* imaging survey of OB Young Stellar Objects (HOBYS, Motte et al. 2010), and Hi-GAL (Molinari et al. 2010) surveys, but also open time programs, study in detail the very complex cloud structure in column density and dust temperature, constructed from spectral energy distribution (SED) fits to the *Herschel* wavelengths (70″ to 500″). Some major results are the recognition of the importance of filaments for the star-formation process (e.g., André et al. 2010, Arzoumanian et al. 2011), and the functionality of probability distribution functions (PDFs) of column density to single out the influence of various physical processes (turbulence, gravity, external compression) on the column density structure of clouds (Schneider et al. 2012, 2013; Russeil et al. 2013; Tremblin et al. 2014). Although significant progress was made to understand the link between the column density/spatial structure of molecular clouds and star formation, a number of important questions are still open. 1. What is the *relative importance of turbulence, gravity, magnetic fields, and radiative feedback* on regulating the overall column density structure of molecular clouds? 2. Are there differences in the column density structure for clouds forming *low-mass stars or high-mass stars*? 3. Is there a *universal (column) density threshold* for the formation of self-gravitating prestellar cores? 4. Does the *star formation efficiency* (SFE) and *star-formation rate* (SFR) depend on the column density structure of molecular clouds?

In a series of papers using column density maps obtained with *Herschel* data and with near-IR extinction, and results of numerical simulations, we will address these questions. This paper makes a start with a detailed study of the validity of column

Correspondence to: nicola.schneider@obs.u-bordeaux1.fr

[1] Herschel is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA.

Table 1. Molecular cloud parameters from *Herschel* data. All values in brackets are the ones determined from the original maps before correction for LOS contamination. The last two lines give the average values (and standard deviation) from the corrected and uncorrected maps.

Cloud	D [kpc] (1)	M [10^4 M$_\odot$] (2)	Σ [M$_\odot$ pc^{-2}] (3)	$\langle N(H_2)\rangle$ [10^{21} cm^{-2}] (4)	ΔA_v [mag] (5)	$A_{v,pk}$ [mag] (6)	A_v(DP) [mag] (7)	σ_η (8)	s (9)	α (10)
High-mass SF regions										
NGC3603	7.0	50.4 (97.1)	60 (116)	3.24 (6.24)	3.0±0.5	1.7 (4.8)	4.4 (8.5)	0.52 (0.27)	-1.33 (-1.69)	2.50 (2.18)
Carina	2.3	34.5 (59.5)	50 (89)	2.79 (4.79)	2.0±0.2	2.5 (4.6)	5.0 (8.0)	0.38 (0.20)	-2.67 (-3.10)	1.75 (1.65)
Low-mass SF regions										
Maddalena	2.2	35.2 (68.2)	37 (76)	2.13 (4.13)	2.0±0.25	1.9 (3.9)	5.0 (7.6)	0.32 (0.20)	-3.65 (-5.21)	1.55 (1.38)
Auriga	0.45	2.2 (3.7)	36 (47)	1.51 (2.52)	0.8±0.1	1.4 (2.3)	3.5 (4.2)	0.45 (0.25)	-2.59 (-3.03)	1.77 (1.66)
⟨Corrected⟩			45±6	2.42±0.38		1.88±0.23	4.48±0.35	0.42±0.04	-2.56±0.47	1.89±0.21
⟨Original⟩			82±14	4.42±0.77		3.90±0.57	7.08±0.97	0.23±0.02	-3.25±0.73	1.72±0.17

(1) Distance D of the cloud.
(2) Mass M $\propto N(H_2)D^2$. The H_2 column density determination assumes a mean atomic weight per molecule of 2.8. A visual extinction A_v of 1 corresponds to $N(H_2) = 0.94\times10^{21}$ A_v cm^{-2} mag^{-1} (Bohlin et al. 1978). The total mass was determined above an A_v level of ∼1, which is a typical value for estimating molecular cloud extend (e.g., Lada et al. 2010). For the uncorrected maps, we determine the mass within the same area (above a threshold of $A_v = 1 + \Delta A_v$).
(3) surface density (Σ = M/area).
(4) average column density (above a level of 10^{21} cm^{-2}).
(5) background/foreground level of visual extinction. The error is the root mean square from the pixel statistics used to determine the contamination.
(6) peak of PDF in visual extinction.
(7) visual extinction value of the deviation point where the PDF starts to deviate from the lognormal shape at high column densities.
(8) dispersion of the fitted lognormal PDF.
(9) slope of the high-density tail of the PDF, determined by linear regression (the χ^2 value is given in the panels of the PDF). We excluded the more noisy and less well sampled points at the high column density end of the PDF.
(10) exponent of the spherical density distribution $\rho \propto r^{-\alpha}$, determined from s with $\alpha = -2/s + 1$ (Federrath & Klessen 2013).

density maps and their PDFs obtained from *Herschel*. We show that line-of-sight (LOS) confusion, i.e. emission from unrelated clouds in front or behind the bulk emission of the molecular cloud leads to a significant overestimation in the column density maps. Apart from this observational approach, we quantify how the PDF properties change by using numerical simulations in which we add noise and foreground/background emission to an uncontaminated PDF.

Probability distribution functions of (column) density form the basis for modern theories of star formation (e.g., Federrath & Klessen 2012; Hennebelle & Chabrier 2008, 2009; Padoan et al. 2014), and are frequently used as an analysis tool for simulations and observations. For further theoretical details, we refer to, e.g., Padoan et al. (1997), Federrath et al. (2008, 2010), and how column density PDFs are obtained from *Herschel* data, see Schneider et al. (2013) and Appendix A.

2. Herschel column density maps

For our study, we obtained the raw data from the *Herschel* archive for the following sources: **Auriga-California:** RA(2000)=4^h21^m, Dec(2000)=37°35′, OT1 PI: P. Harvey, see Harvey et al. (2013); **Maddalena:** RA(2000)=6^h48^m, Dec(2000)=-3°39′, OT1 PI: J. Kaufmann; **NGC3603:** RA(2000)=11^h15^m, Dec(2000)=-61°15′, Hi-GAL ("*Herschel* Infrared GALactic plane survey", Molinari et al. (2010); **Carina:** RA(2000)=10^h43^m, Dec(2000)=-59°24′, OT1 PI: T. Preibisch, see Preibisch et al. (2012), Gaczkowski et al. (2012), Roccatagliata et al. (2013). These clouds were selected in order to cover different masses, sizes, and levels of SF. The Maddalena cloud is very massive but shows only a low level of SF, Auriga is less massive but forms low-mass stars, and Carina and NGC3603 are high-mass SF regions with associated OB clusters.

The process of data reduction and determination of column density maps at an angular resolution of ∼37″ is explained in detail in, e.g., André et al. (2010), Könyves et al. (2010), Schneider et al. (2012). We additionally give a summary in Appendix B where we also show the resulting column density maps and individual 250 μm maps. Generally, throughout the paper, we use the definition by, e.g., Bergin & Tafalla (2007) of *low-mass SF regions* as molecular clouds with a mass of 10^3–10^4 M$_\odot$ and a size of up to a few 10 pc (such as Maddalena and Auriga), and *high-mass SF regions* as giant molecular clouds with a mass of 10^5–10^6 M$_\odot$, a size of up to a ∼100 pc and signs of active massive star-formation (Carina and NGC3603).

A notorious problem of continuum maps (extinction maps or column density maps derived from SED fitting using *Herschel*) is line-of-sight confusion. This is particularly true in the Galactic plane and along spiral arms where dust emission not related to the cloud can significantly add, and thus artifically increase, the derived column density for the cloud studied at a single distance. In Marshall et al. (2006) values for the total interstellar extinction at 15′ resolution are given along the Galactic plane (l=±100°, b=±10°). For the Carina and NGC3603 clouds for example, we note that these regions are in a location of relatively low average column density (from a few A_v up to maximum values in the clouds of up to A_v ∼20). However, it is not easy to distinguish between bulk emission of the cloud and background or foreground contribution from unrelated clouds if the distance information is missing. A simple approach we present in this paper is to determine a mean value for the contaminating column density by averaging pixels outside the molecular cloud close to the map borders. We find values between ΔA_v ∼0.8 and

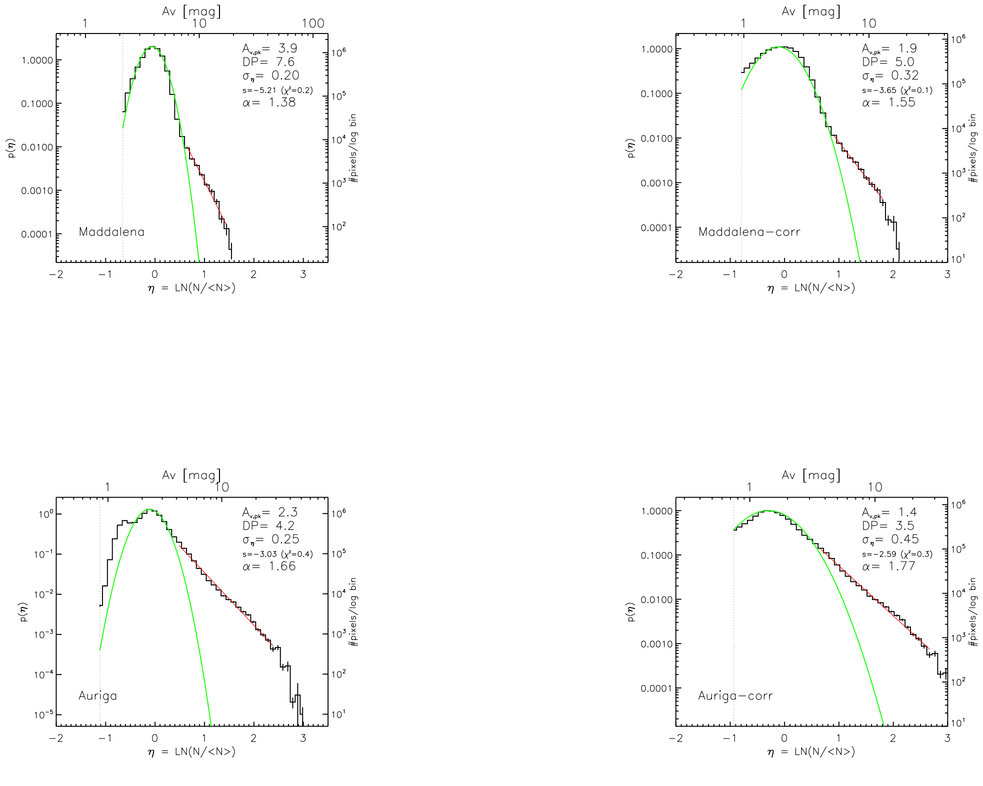


Fig. 1. PDFs derived from *Herschel* column density maps towards the **low-mass star-forming regions** Maddalena and Auriga. The left panel shows the PDF from the original map, the right one from the corrected map. The vertical dashed line indicates the noise level of the map. The left y-axis gives the normalized probability $p(\eta)$, the right y-axis the number of pixels per log bin. The upper x-axis is the visual extinction and the lower x-axis the logarithm of the normalized column density. The green curve indicates the fitted PDF. For Auriga, diffuse LOS-contamination (or a well defined seperate cloud) shows up as an individual PDF at low extinctions. The red line indicates a power-law fit to the high A_v tail. Inside each panel, we give the value where the PDF peaks (A_{pk}), the deviation point from lognormal to power-law tail (DP), the dispersion of the fitted PDF (σ_η), the slope s and the X^2 of the fit (linear regression), and the exponent α of an equivalent spherical density distribution. These values are also summarized in Table 1.

3 (see Table 1), with an uncertainty of 0.1 to 0.5, that are then subtracted as an offset from the original column density map. A similar method was proposed for extinction maps in Lombardi et al. (2008). This approach ignores possible gradients and generally the spatial variation of the emission distribution of clouds that are located along the line-of-sight (more distant background clouds are expected to be smaller than foreground clouds that are assumed to be bigger in angular size), and can lead to unrealistic features in the PDF (see Sec. 4). For simplicity, we here only consider correcting for foreground/background contamination by subtracting a constant value, which already provides significantly more reliable estimates of the true column density and masses.

An alternative method to estimate the contamination is to use the velocity information from atomic hydrogen or molecular lines and to translate the observed intensity into H_2+HI column density. This has been done for *Herschel* studies of NGC6334 (Russeil et al. 2013), and W3 (Rivera-Ingraham et al. 2013). Recently, a special method was developed to obtain an H_2 map from ^{13}CO and C^{18}O observations in W43 that does not suffer from a cut-off at high column densities (Carlhoff et al. 2013). However, these methods are subject to other uncertainties such as variable CO/H_2 conversion factor or the variation in excitation temperature T_{ex} (see Fig. 11 in Carlhoff et al. 2013 how the PDF shifts by changing only a few Kelvin T_{ex}).

3. Column density maps and PDFs

The LOS contamination correction reduces significantly (up to a factor ∼2) the absolute values for average column density $\langle N(H_2)\rangle$, mass, and surface density (Table 1). The values for $\langle N(H_2)\rangle$ range between 1.5 and 3.2 $\times 10^{21}$ cm^{-2} with an average value of $(2.4\pm0.4)\times10^{21}$ cm^{-2}. The surface density Σ varies between 36 for Auriga and 60 for NGC3603 with an average value of Σ = 45 $M_\odot$ pc^{-2} (instead of 82 $M_\odot$ pc^{-2} that was obtained from the original maps).[2] This value corresponds very well to the one found by Heyer et al. (2009) with Σ = 42 $M_\odot$ pc^{-2} from a sample of clouds (>250) from the FCRAO Galactic Ring Survey investigated with ^{13}CO 1→0 emission, and to the value Σ = 41 $M_\odot$ pc^{-2} obtained from extinction maps of five nearby clouds (Lombardi, Alves, Lada 2010). However, we emphasize

[2] We here use the threshold A_v= 1 to 'define' a molecular cloud. Taking a higher value such as A_v= 2, we obtain a variation in Σ between 47 and 90 with an average value of 63 $M_\odot$ pc^{-2}.

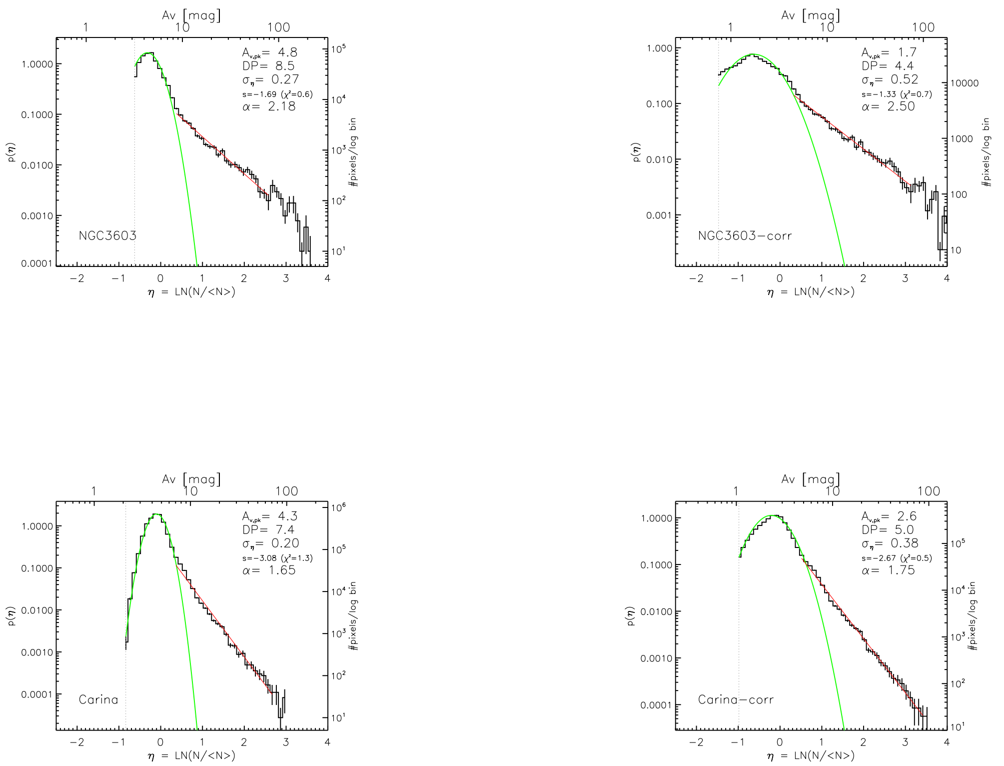


Fig. 2. PDFs derived from *Herschel* column density maps towards the **high-mass star-forming regions** NGC3603 and Carina. The left panel shows the PDF from the original map, the right one from the corrected map. All other parameters as in Fig. 1.

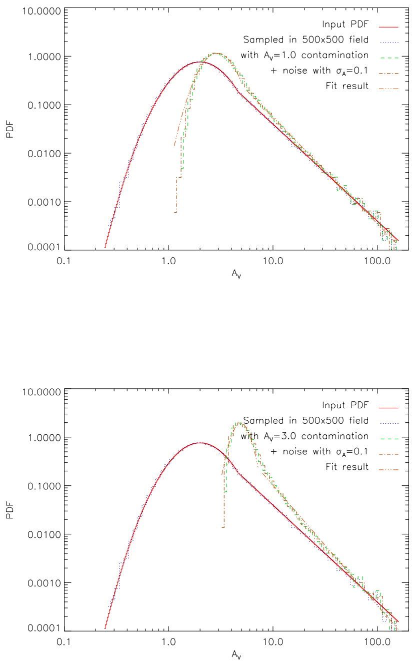
that our simple approach may still under- or overestimate the LOS contamination and the statistic we present here is low (but shows already a trend). In a forthcoming paper (Paper II), we will present a study of more than 20 clouds to obtain a larger statistical significance.

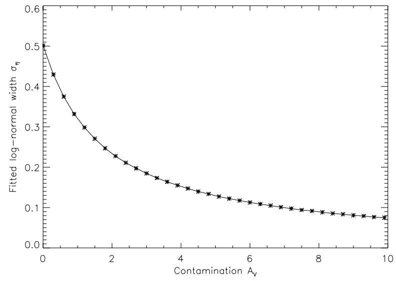
The PDFs obtained from the original and corrected maps are shown in Figs. 1 and 2. The low column density regime of the uncorrected maps is limited by noise and LOS contamination of the map (ΔA_v -value listed in Table 1), and we plot the PDF starting at the respective value. We then fitted with a lognormal function the lower extinction part of the PDF.[3] For the LOS-corrected maps, the low extinction values (left of the PDF peak) do not neccessarily have a perfect lognormal distribution (e.g., Maddalena) because removing a constant offset can lead to negative pixels in the maps which are ignored during the process to make the PDF. As outlined in Sec. 4, noise and 'overcorrection' can lead to a Gaussian pixel distribution that shows up as a linear run in the low column density range of the PDF. We therefore slightly iterated the correction value for the contamination value in order to avoid this effect and optimize the threshold where we start the lognormal fit. The original Auriga PDF (see also Harvey et al. 2013) shows a superposition of two PDFs where the first (low extinction) peak is compatible with the contamination level of the map and nearly disappears when a level of ΔA_v = 0.8 is removed. Note that this 'double-peak' is not the same feature as it was observed for regions with expanding ionization fronts (Schneider et al. 2012, Tremblin et al. 2014) where the second peak appears at high column densities arising from compression and depends on the turbulent state of the cloud. In the case of Auriga, the correction works very well because the contamination is most likely a rather homogeneous layer in front of or behind the Auriga cloud. Such a superposition was already observed in Pipe (Lombardi et al. 2006) and we give more examples in Paper II. The LOS contamination correction has several effects, (i) the PDF is broader (increase of $\langle\sigma_\eta\rangle$ from 0.23±0.02 to 0.42±0.04), (ii) the slope becomes flatter, and (ii) the peak and deviation point of the PDF from lognormal to excess (A_v(DP) from nowon) shift to lower values. In Sec. 4, we quantify these effects by an analytic model. The most dramatic change is observed for the PDFs of high-mass SF regions (Fig. 2). The very narrow distribution[4] ($\sigma_\eta = 0.27$) for NGC3603 becomes much broader (σ_η = 0.52), the PDF peak shifts from A_v, 4.8 to 1.7, and A_v(DP) from 8.5 to 4.4. The same applies for the Carina PDF with a change of σ_η = 0.2 to 0.38 after correction[5]. Other *Herschel* studies of high-mass SF regions such as NGC6334 (Russeil et al. 2013), where a LOS contamination of ΔA_v ∼2–3 was estimated, show the same narrow PDFs. *It is*

[3] Note that we keep the classical approach to fit one lognormal PDF to the low column density range though other functional fits such as a Gaussian can be possible as well (Alves et al. 2014).

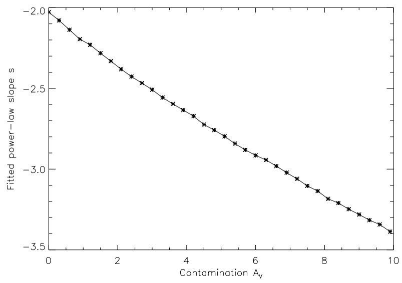
[4] Note that a low angular resolution (such as for extinction maps at a few arcmin resolution) also naturally results in narrower PDFs as the highest density structures are not well resolved. This effect can become important for very distant clouds.

[5] Our PDF is different from the one shown in Preibisch et al. (2012) because their column density map was obtained from a fit using only the *Herschel* wavelengths 70 and 160 μm, which are not a good tracer of *cold* gas.





(a)



(b)

Fig. 3. Simulations of the PDF of an observed cloud with an originally perfectly lognormal core and a power-law tail (solid line). The representation by a finite-size map introduces some uncertainties (dotted line). The addition of an ΔA_V = 1.0 (3) contamination distorts the shape of the PDF and observational noise adds a low-extinction component. The deviation point A_V(DP) shifts by ∼4 magnitudes for a contanimation of ΔA_V = 1.0 (3). The fit of the resulting PDF by a lognormal core and a power-law tail does not recover the initial parameters.

Fig. 4. Dependence of the fitted width of the lognormal core of the PDF *(a)* and the fitted slope of the power-law tail *(b)* as a function of the foreground (and/or background) contamination for the standard parameters of the molecular cloud PDFs.

thus essential to consider the contamination for a correct interpretation of PDFs for high-mass SF regions. In particular the interpretation of the slope of the power-law tail is important for high-mass SF regions. For example, we found a slope variation within different subregions of a cloud (see NGC6334, Russeil et al. 2013), depending on the radiative impact. Tremblin et al. (2014) showed explicitly that the power-law tail becomes steeper going from the interaction zone between HII regions and cloud into the cloud center. This implies that compression of gas takes place, and that self-gravity then takes over in the densest regions to form cores and finally stars. The flat slope we observe for NGC3603 is thus also most likely a consequence of radiative feedback processes.

In Table 1 the PDF properties are listed, and reveal that the correction leads to an *equalization* of the values for all clouds. The PDF peak values have now a range of 1.4 to 2.5 instead of 2.3 to 4.8 and A_V(DP) changes to 3.5 to 5.0 with an average of 4.48±0.35 (original values are 4.2 to 8.5 with an average of 7.08±0.97).

4. Simulations

We have modeled the effect of LOS contamination in a numerical simulation, starting from an "ideal" PDF consisting of a lognormal part and a power-law tail. Inverting the integral over the PDF allows us to create a distribution of random numbers that provide the PDF within statistical uncertainties. For the examples shown here, we chose a field size of 500×500 pixels matching the typical observational map size and grid (see Appendix A) investigated in Sect. 3. Larger and smaller fields were only tested to verify the numerical accuracy of the method.

The generated field was then "contaminated" by adding a constant level to all map values and finally "observed" including Gaussian white noise typical for the *Herschel* observations. The overall model is characterized by six parameters: the width and the center A_V of the lognormal PDF contribution, the exponent of the power-law tail, the deviation point characterizing the transition from the lognormal to the power-law PDF, the additive contamination ΔA_V and the standard deviation of the observational noise. To analyze the impact of the contamination we fixed all other parameters to lie within a range obtained for the observations in Sect. 3, i.e., a center (peak) at A_V= 2.0, σ_η = 0.5, $s = -2.0$, A_V(DP) = 4.3, and a noise rms σ_{A_V} = 0.1.

In Fig. 3 it becomes obvious that the fit of the resulting measurable PDFs provides parameters that clearly deviate from the original input, in particular for the ΔA_V=3.0 case. One observes there that the lognormal part of the PDF is strongly compressed, consistent with what we observe for the NGC3603 cloud (Fig. 2). The statistical sampling of that part becomes very rough. The value for A_V(DP) increases and the slope of the power-law tail becomes steeper than $s = -2.0$ that characterized the original cloud.[6]

[6] Note that the offset correction mathematically results in a modified function without a power-law tail. However, the corrections are small enough, so that fitting a power-law function and infering a slope from this fitting function is still a reasonable procedure.

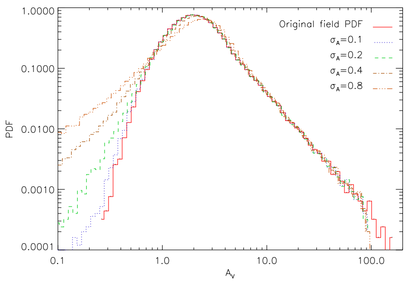


Fig. 5. Simulations of reconstructed PDFs observed with different amplitudes of noise, expressed in visual extinctions $\sigma_{\rm A}$. The reconstruction does not depend on the absolute value of the foreground/background contamination.

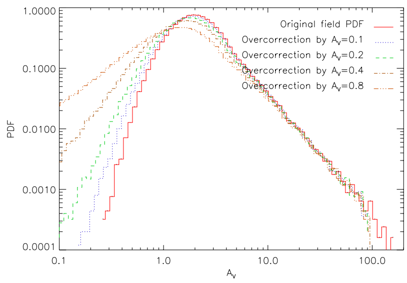


Fig. 6. Simulations of a PDF with negligible noise ($\sigma_{\rm A}$ = 0.01) when applying too high values for the contamination correction ΔA$_{\rm v}$. The reconstruction does not depend on the absolute value of the foreground/background contamination but only on the difference between actual contamination and subtracted contamination.

To quantify this effect systematically, we show in Fig. 4 the change of the width of the PDF and the slope of the power-law tail when measured as a function of LOS contamination, always starting from the standard parameters of the underlying cloud with σ_η = 0.5 and $s = -2.0$. The small irregularities in the curve result from the discrete binning of the randomly sampled PDF. We find a dramatic effect in both parameters. For a contamination of ΔA$_{\rm v}$=2, σ_η is already reduced by more than a factor two, and the slope of the power-law tail has steepened from −2.0 to −2.4. For an ΔA$_{\rm v}\approx$ 10, characterizing distant massive regions or infrared dark clouds, σ_η has decreased by more than a factor of five and the power-law tail has a slope of −3.4.

The addition of observational noise adds a contribution to the structure that is hardly visible in the contaminated PDF, but that may reappear amplified after the correction of the contamination. The noise is assumed to have an approximately Gaussian distribution providing a floor of small fluctuations with zero average. In the logarithmic binning, the core of this Gaussian gives a linear contribution to the PDF for small densities, i.e. below the lognormal part of the PDF of the observed structure. This effect is simulated in Fig. 5 where we varied the observational noise from Fig. 3 and applied the contamination correction to reconstruct the PDF of the original structure.

We find a clear excess at low column densities which turns close to the expected linear behaviour for $\sigma_{A_{\rm V}}$ = 0.8. Starting from $\sigma_{A_{\rm V}}$ = 0.4 we find also a slight, but noticeable, shift of the PDF peak to higher A$_{\rm v}$. For actual observed data, this should be taken into account. The effect is independent of the contamination that is added and subtracted in the PDF transformation.

The observed PDFs in Sect. 3 show the same kind of low-column density excess as these simulations, in fact, the example of the Maddalena cloud exactly matches the simulation for $\sigma_{A_{\rm V}}$ = 0.4. However, the pure observational noise in the column density maps is much lower. This can be explained by additional small scale uncertainties and fluctuations in $\sigma_{\rm A}$. They behave similar to noise, but are no observational noise. They may represent fluctuations in the overall cloud contamination, either variations in the foreground screen or small background contaminating clouds (see also discussion in Alves et al. (2014) for correlated pattern of noise in extinction maps). As long as the fluctuations are relatively small, their impact on the fit of the main lognormal part of the PDF is negligible.

Very low column densities, giving rise to a linear contribution to the PDF for low A$_{\rm v}$, can however also stem from an "overcorrection" of the contamination, shifting part of the real cloud structure to column densities around zero. Therefore, an 'overcorrection' of the map, i.e., using a too high value for the contamination correction by ΔA$_{\rm v}$, leads to a similar effect as increased noise. Figure 6 shows this for an example calculated with negligible noise ($\sigma_{\rm A}$ = 0.01) for different levels of ΔA$_{\rm v}$. We find a similar linear distribution for the low colum density pixels that becomes more important for increasing ΔA$_{\rm v}$. In this case, the peak of the PDF shifts in the other direction, compared to the addition of noise/fluctuations, i.e., to lower column densities.

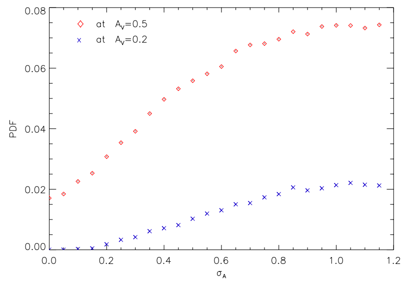


Fig. 7. Amplitude of the small-column density in the PDF as a function of the added level of noise or small-scale fluctuations, corresponding to Fig. 5 for two points in the low-column density wing. The scatter in the points results from the adaptive binning in the PDF computation.

When evaluating the low-column-density excess, we find that an overcorrection always has a stronger impact than noise or fluctuations of the same amplitude. Therefore we can analyse the low-density excess in the observations to get an upper limit of the uncertainty in the contamination correction when comparing the excess with the noise impact. This uncertainty then represents either fluctuations in the contamination or an absolute error in the contamination correction (in which case it would be somewhat overestimated). This is done through the simulation shown in Fig. 7. It gives the amplitude of the PDF at $A_{\rm V}$ = 0.2 and $A_{\rm V}$ = 0.5 as a function of the amplitude of the fluctuations, i.e., this quantifies the low-density excess in Fig. 5 for all pos-

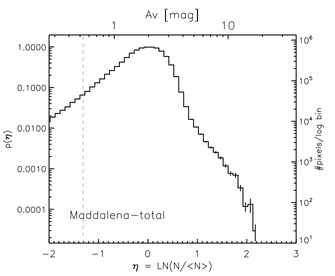


Fig. 8. PDF of Madalena including the very low column density range down to A$_v$~0.2. The dashed line indicates the A$_v$= 0.5 level used to derive the amplitude of the PDF (in this case ≈0.06) to be compared to Fig. 7 in order to determine the uncertainty of the correction (in this case ≈0.5-0.55).

sible noise/fluctuation levels. For all observations, we have used this approach to quantify the uncertainty of the contamination correction, then excluding the low-density part from the fit of the lognormal part of the PDF of the actual cloud structure (see Sect. 3). Figure 8 illustrates an example of this procedure where we chose the 'worst case' scenario for Maddalena. The PDF is plotted over a large column density range and the A$_v$= 0.5 level corresponds to an amplitude of the PDF of ≈0.06. Using Fig. 7 as a look-up table indicates that for this amplitude, the uncertainty of the correction corresponds to approximately 0.5 magnitudes.

In this way, we cannot only estimate the contamination of the cloud extinction by foreground/background material, but also the uncertainty of the contamination. The uncertainty can be due to random fluctuations or a systematic error. For any measured low-density excess, the systematic error is always smaller than the random error, so that we provide an upper limit to the total uncertainty by giving the magnitude of the possible fluctuations only.

5. Discussion and Conclusions

5.1. Is there a difference in the density structure between molecular clouds ?

The 'third Larson law' (Larson 1981) is an empirical relation between the volume density n of a cloud and its size L with $n \propto L^{-1.1}$, and implies that all molecular clouds have approximatly constant mean column densities. However, there is a long-lasting controversy about its validity in theory (e.g., Ballesteros-Paredes & Mac Low 2002) and observations (e.g., Schneider & Brooks 2004, Heyer et al. 2009). Recently, a study of Lombardi, Alves, & Lada (2010) using extinction maps of five molecular clouds concluded that all clouds in their sample have identical average column densities above a given extinction threshold. However, their investigation does not include very massive and dense clouds (the only high-mass SF region in their sample is Orion A with a column density below 10^{23} cm^{-2}, see, e.g. Polychroni et al. 2013).

We showed that a correction of column density maps for the effect of LOS contamination leads to a typical value of 1.5–3 $\times 10^{21}$ cm^{-2} for the average column density for the individual clouds with a total average for all clouds of $\langle N(H_2) \rangle$ = $(2.4 \pm 0.4)\times 10^{21}$ cm^{-2}. This range is rather narrow, though in our sample the high-mass SF clouds - NGC3603 and Carina - have slightly higher average column densities and surface densities compared to the low-mass SF clouds Auriga and Maddalena. Taking the whole dynamic range of column densities of molecular clouds across two magnitudes (a few 10^{21} cm^{-2} up to a few 10^{23}cm^{-2}), we find that there is no considerable difference in the column density structure of molecular clouds.

5.2. What does the common deviation point of the PDF at A$_v$= 4–5 signify ?

The LOS-correction changes the PDFs significantly and leads to a common peak at A$_v$ around 2 and a width of σ_η around 0.4-0.5. Particularly interesting is the common deviation point of the PDF from lognormal to a power-law tail at A$_v$(DP)~4–5 (variation between 3.5 and 5.0). Kainulainen et al. (2011) explained the value they found for PDFs obtained from extinction maps (A$_v$(DP) around 2–4) caused by a phase transition between lower-density interclump gas and pressure-confined clumps. Froebrich & Rowles (2011) interprete the value they found (A$_v$(DP)=6.0±1.5, also extinction maps), due to the effect of gravity. The scatter in their values is large (16 clouds were studied with values of A$_v$(DP) between 3.4 and 8.6) and based on our study, we expect that it is LOS contamination that led to the large scatter.

In our study, we find an average slope of the power-law tail s that corresponds to an exponent $\alpha = 1.9 \pm 0.2$ of an equivalent radial density profile (see the studies of the link between PDF and self-gravity by Klessen (2000), Kritsuk et al. (2011), Federrath & Klessen (2013), Girichidis et al. (2014)). The collapse of an isothermal sphere has been studied since long (Larson 1969, Penston 1969, Shu 1977, Whitworth & Summers 1985), and though all models start with different initial conditions, they arrive to the same $\alpha = 2$ for early stages and $\alpha = 1.5$ after a singularity formed at the center of the sphere. Can the power-law tail then stem only from local core collapse ? In most of the clouds, this is unlikely because cores constitute only a small mass fraction of the total gas mass (e.g. 15% of dense gas in Aquila, André et al. (2014)). In addition, there are clear observational signatures for gravitational collapse on much larger scales, e.g., the radial collapse of the DR21 filament on a few pc (Schneider et al. 2010). Gravitational fragmentation of dense filaments into prestellar cores, possibly supported by accretion via filaments oriented orthogonal to the main filament (called 'striations', see Palmeirim et al. 2013), is proposed as the main process to form solar-type protostars (André et al. 2014). Mass accretion by larger subfilaments is considered further as an important process to build up the large mass reservoir to form massive star(s) (see e.g., Schneider et al. 2010, 2012; Galvan-Madrid et al. 2010; Ngyuen-Luong et al 2011; Hennemann et al. 2012; Peretto et al. 2013 for observations and Girichidis et al. 2012, Smith et al. 2011, 2012 for theory). The power-law tail of the PDF is thus not only due to local core collapse, but also to the aforementioned processes, and the fact that the power-law tail in the PDFs we present here is continuous, points toward self-similarity. This was already noticed in the hydrodynamic turbulence models including self-gravity of Klessen, Heitsch, Mac Low (2000), Heitsch, Mac Low, Klessen (2001), and Ballesteros-Paredes et al. (2011) and Kritsuk et al. (2011). We thus interprete the deviation point of the PDF at A$_v$= 4–5 as the transition of low (column)-density, turbulent gas into a denser phase that is controlled by gravity on all scales (global collapse and accretion, core collapse). Girichidis et al. (2014) state that if *each density parcel collapses on its free-fall timescale then the evolution of the overall PDF is independent of the number of fragements and* *the details of the fragmentation process.* It is thus challenging to single out the different 'apparitions' of gravity in the power-law tail of the PDF. It is already possible to investigate the column density environment of cores, though there is a large variation in these values. Onishi et al. (1998) studied the properties of $C^{18}O$ 1→0 cores in *Taurus* and found that the average column density of star-forming cores is 5.6×10^{21} cm^{-2} for warm sources and 8.9×10^{21} cm^{-2} for cold sources. Heiderman et al. (2010) suggest a 'star-formation threshold' at A_v=8 that was defined as a steep decrease of the star-formation rate surface density Σ_{SFR} over gas surface density Σ_{gas}. Recent *Herschel* studies show also some variations: André et al. (2014) found that 90% of the bound low-mass prestellar cores in Aquila (Könyves et al. 2010) are located above a background extinction of A_v ∼7. Polychroni et al. (2013) detected 84% of the gravitationally bound cores in L1641 (Orion A) in filaments above A_v ∼5 (and most of them exhibit A_v>20). The same value was found by Harvey et al. (2013) in Auriga for the location of YSO.

Similarily, there is no observational support for a clear threshold *volume density n* of for the formation of self-gravitating cores either. Submillimeter continuum studies found values around $n \sim 10^{4-5}$ cm^{-3} (e.g., Lada et al. 1992, Motte et al. 1998, Kirk et al. 2005) but also $n \sim 10^{5-6}$ cm^{-3} (Ward-Thompson et al. 1994). Bergin et al. (2001) found that N_2H^+ (critical density 10^5 cm^{-3}) exists only in dense cores potentially forming stars at an A_v above 4 which is consistent with predictions of chemical theory. From a theoretical point of view, outlined in Padoan et al. (2014), a universal volume density threshold of a few times 10^4 cm^{-3} was determined, based on the star formation rate theories of Krumholz & McKee (2005), Hennebelle & Chabrier (2008), Padoan & Nordlund (2011), and Federrath & Klessen (2012).

Our study attempts to correct for the most important observational complication - line-of-sight confusion - and we derive a 'threshold' value of A_v∼4–5 (column density $N(H_2)$∼3.8-4.7×10^{21} cm^{-2}) defined as the transition from turbulent gas (lognormal part of the PDF) to self-gravitating gas (power-law tail emerging for extinctions/column densities above A_v∼4–5). Taking this scenario face value, and assuming a minimum star-formation *volume density* $n \approx 10^4$ cm^{-3}, we calculate a size scale of 0.12 pc from N/n with N = 3.8×10^{21} cm^{-2}. (Taking other combinations of n and N leads to a range of size scales between 0.012 pc to 0.15 pc.) Values around 0.1 pc correspond to the already observed typical size scale of collapsing cores, but also to the width of gravitationally unstable filaments (André et al. 2010, Arzoumanian et al. 2011) in low-mass SF regions.

The final interpretation from this study of the column density PDFs is thus that the gas mass reservoir above an extinction value A_v∼4–5 is strongly affected by self-gravity effects, and may be globally contracting in most clouds (whether massive or not). A clear seperation in the PDF between global (such as filament) collapse, and local core collapse (and other effects like radiative feedback) awaits further studies that make the link between the core population (pre- and protostellar) and the PDF.

Acknowledgements. N.S., S.B., and P.A. acknowledge support by the ANR-11-BS56-010 project "STARFICH". R.S.K. acknowledges subsidies from the Deutsche Forschungsgemeinschaft, priority program 1573 ("Physics of the Interstellar Medium") and the collaborative research project SFB 881 ("The Milky Way System", subprojects B1, B2, and B5). C. Federrath acknowledges the Australian Research Council for a Discovery Projects Fellowship (grant No. DP110102191). T.Cs. acknowledges financial support for the ERC Advanced Grant GLOSTAR under contract no. 247078. V.O. and N.S. acknowledge supported by the Deutsche Forschungsgemeinschaft, DFG, through project number 0s 177/2-1 and central funds of the DFG-priority program 1573 (ISM-SPP).

References

Alves, J., Lombardi, M., Lada, C., 2014, A&A, in press
André, Ph., Men'shchikov A., Bontemps S., et al., 2010, A&A 518, L102
André, Ph., Di Francesco, J., Ward-Thompson, D., et al., 2014, Protostars and Planets VI, University of Arizona Press (2014), eds. H. Beuther, R. Klessen, Dullemond, Th. Henning, arXiv:1312.6232
Arzoumanian, D., André, Ph., Didelon, P., et al., 2011, A&A, 529, L1
Ballesteros-Paredes, J., Mac Low, M.-M., 2002, ApJ, 570, 734
Ballesteros-Paredes, J., Vázquez-Semadeni, E., Gazol, A., 2011, MNRAS, 416, 1436
Bergin, E. A., Ciardi, D.R., Lada, C.J., et al., 2001, ApJ, 557, 209
Bergin, E. A., Tafalla, M., 2007, ARA, 26, 45, 339
Bohlin, R.C., Savage, B.D., Drake, J.F., 1978, ApJ 224, 132
Carlhoff P., Schilke P., Nguyen-Luong Q., Motte, F., Schneider N., et al., 2013, A&A, 560, 24
Federrath, C., Klessen, R.S., Schmidt, W., 2008, ApJ, 688, L79
Federrath, C., Roman-Duval, J., Klessen, R.J., et al., 2010, A&A, 512, 81
Federrath, C., Klessen, R. S., 2012, ApJ, 761, 156
Federrath, C., Klessen, R. S., 2013, ApJ, 763, 51
Froebrich, D., Rowles, J., 2011, MNRAS, 406, 1350
Gaczkowski, B., Preibisch, T., Ratzka, T., et al., 2012, A&A, 549, 67
Galvan-Madrid, R., Zhang, Q., Keto, E., et al., 2010, ApJ, 725, 17
Girichidis, P., Federrath, C., Banerjee, R., Klessen, R. S., 2012, MNRAS, 420, 613
Girichidis, P., Konstandin, L., Whitworth, A.P., Klessen, R., et al., 2014, ApJ, 781, 91
Griffin, M., Abergel, A., Abreau, A., et al., 2010, A&A, 518, L3
Harvey, P., Fallscheer, C., Ginsburg, A., et al., 2013, ApJ, 764, 133
Heiderman, A., Evans, N.J. II, Allen, L.E., et al., 2010, ApJ, 723, 1019
Heitsch, F., Mac Low, M.-M., Klessen, R.S., 2001, ApJ, 547, 280
Hennebelle, P., Chabrier, G., 2008, ApJ, 684, 395
Hennebelle, P., Chabrier, G., 2009, ApJ, 702, 1428
Hennemann, M., Motte, F., Schneider, N., et al., 2012, A&A, 543, L3
Heyer, M.H., Krawczyk, C., Duval, J., 2009, ApJ, 699, 1092
Johnstone, D., Di Francesco, J., Kirk, H., 2004, ApJ, L45
Kainulainen, J., Beuther, H., Henning, T., & Plume, R., 2009, A&A, 508, L35
Kainulainen, J., Beuther, H., Banerjee, R., et al., 2011, A&A, 530, 64
Kirk, J.M., Ward-Thompson, D., André, Ph., 2005, MNRAS, 360, 1506
Klessen, R. S., 2000, ApJ, 535, 869
Klessen, R. S., Heitsch, F., Mac Low, M.-M., 2000, ApJ, 535, 887
Könyves, V., André Ph., Men'shchikov A., et al., 2010, A&A, 518, L106
Kritsuk, A.G., Norman, M.L., Wagner, R., 2011, ApJ, 727, L20
Krumholz, M., & McKee, C.F., 2005, ApJ, 630, 250
Lada, C.J., Lombardi, M., Alves, J., 2010, ApJ, 724, 687
Lada, E., A., 1992, ApJ, 393, 25
Larson, R.B., 1969, 145, 271
Larson, R.B., 1981, MNRAS, 194, 809
Lombardi,M., Alves,J., Lada, C., 2006, A&A, 454, 781
Lombardi,M., Lada, C., Alves, J., 2008, A&A, 489, 143
Lombardi,M., Alves, J., Lada, 2010, A&A, 519, L7
Marshall, D.J., Robin, A.C., Reylé, C., et al., 2006, A&A, 453, 635
Molinari, S., Swinyard, B., Bally, J. et al. 2010, A&A, 518, L100
Motte, F., André, Ph., Neri, R., 1998, A&A, 336, 150
Motte, F., Zavagno A., Bontemps S., et al., 2010, A&A 518, L77
Ngyuen-Luong Q., Motte F., Hennemann M., et al., 2011, A&A, 535, 76
Onishi, T., Mizuno, A., Kawamura, A., et al., 1998, ApJ, 502, 296
Padoan, P., Jones, J.T., Nordlund, A.A., 1997, ApJ, 474, 730
Padoan, P., Nordlund, A.A., 2011, ApJ, 730, 40
Padoan, P., Federrath, C., Chabrier, G., et al., 2014, Protostars and Planets VI, University of Arizona Press (2014), eds. H. Beuther, R. S. Klessen, C. P. Dullemond, Th. Henning, arXiv:1312.5365
Palmeirim, P., André, Ph., Kirk, J., et al., 2013, A&A, 550, 38
Penston, M.V., 1969, MNRAS, 144, 425
Peretto, N., Fuller, G.A., Duarte-Cabral, A., et al., 2013, A&A, 555, 112
Pilbratt, G., Riedinger, J., Passvogel, T., et al., 2010, A&A 518, L1
Poglitsch, A., Waelkens, C., Geis, N., et al., 2010, A&A 518, L2
Polychroni, D., Schisano, E., Elia, D., et al., 2013, ApJ, 777, L33
Preibisch, T., Roccatagliata, V., Gaczkowski, B., Ratzka, T., 2012, A&A, 541, 132
Rivera-Ingraham, A., Martin, P.G., Polychroni, D., et al., 2013, ApJ, 766, 85
Roccatagliata, V., Preibisch, T., Ratzka, T., Gaczkowski, B., 2013, A&A, 554, 6
Roussel, H., 2013, PASP, 125,1126
Russeil, D., Schneider, N., Anderson, L., et al., 2013, A&A, 554, 42
Schneider, N., Brooks, K., 2004, PASA, 21, 290
Schneider, N., Csengeri, T., Bontemps S., et al., 2010, A&A, 520, 49
Schneider, N., Csengeri, T., Hennemann, M., et al., 2012, A&A, 540, L11

Schneider, N., André, Ph., Könyves, V., Bontemps, S., et al., 2013, ApJ, 766, L17
Shu, F.H., 1977, ApJ, 214, 488
Smith, R.J., Glover, S., Bonnell, I., Clark, P., Klessen, R.S., 2011, MNRAS, 411, 1354
Smith, R.J., Shetty, R., Stutz, A., Klessen, R.S., 2012, ApJ, 750, 64
Tremblin, P., Schneider, N., Minier, V., et al., 2014, A&A, in press, astro-ph:1401.7333
Ward-Thompson, D., Scott, P.F., Hills, R.E., André, Ph., 1994, MNRAS, 268, 276
Whitworth, A., & Summers, D., 1985, MNRAS, 214, 1

Appendix A: Probability distribution functions of column density

We define a column density PDF as the probability to find gas within a range [N,N+dN]. It is given by the surface-weighted PDF with $\int_{N}^{N+dN} p_N(N')\, dN'$, where $p_N(N)$ corresponds to the probability distribution of the column density. (Note that the column density N can be replaced by the visual extinction A_v). We define $\eta \equiv \ln(N/\langle N\rangle)$ as the natural logarithm of the column density N, divided by the mean column density $\langle N\rangle$, and the quantity $p_\eta(\eta)$ then corresponds to the probability distribution function of η with the normalization $\int_{-\infty}^{+\infty} p_\eta d\eta = \int_0^{+\infty} p_N dN = 1$.

The column density maps from which we determine the PDF are typically on a half-beam (18″) or Nyquist sampled grid. The maps are on average large (>1°) and have a size of a few hundred^2 px^2 up to a few thousand^2 px^2. The binsize that provides the best compromise between resolving small features in the PDF and averaging out too strong variations due to small pixel statistics is 0.1. However, testing binsizes of 0.05 up to 0.2 showed that the PDF properties (width, peak, A_v(DP), slope of power-law tail) do not change. In order to derive the characteristic properties of the PDF (width σ, peak, deviations from the lognormal shape), we fit the lognormal function

$$p_\eta\, d\eta = \frac{1}{\sqrt{2\pi\sigma_\eta^2}} \exp\left[-\frac{(\eta-\mu)^2}{2\sigma_\eta^2} \right] d\eta \tag{A.1}$$

where σ_η is the dispersion and μ is the mean logarithmic column density. We do this systematically by performing several fits on a grid of parameters for η and μ and then calculate the positive and negative residuals. Because the power-law tail is expected to lie above the lognormal form, we select fits with the least negative residuals. We then determine the range of lognormality, when the difference between the model and p_η is less than three times the statistical noise in p_η.

As shown by various authors (e.g. Kainulainen et al. 2009, Schneider et al. 2012, 2013), a power-law tail for high column densities emerges for star-forming regions. We perform a linear regression fit in order to determine the slope s. The values taken into account start at the deviation point (DP) where the lognormal PDF turns into a power-law distribution and stop where the power-law is no longer well defined (at high column densities) due to a smaller pixels statistic caused by resolution effects. In case the tail is only due to gravity, and if we assume spherical symmetry, the PDF slope s of the power-law tail is related to the exponent α of the radial density profile $\rho \propto r^{-\alpha}$ with $\alpha = -2/s{+}1$ (Federrath & Klessen 2013).

Appendix B: Herschel maps of molecular clouds

For the data reduction of the five *Herschel* wavelengths, we used the HIPE10 pipeline including the destriper task for SPIRE (250, 350, 500 μm), and HIPE10 and scanamorphos v12 (Roussel et al. 2013) for PACS (70″ and 160″, Poglitsch et al. 2010). The SPIRE (Griffin et al. 2010) maps include the turnaround-data and were calibrated for extended emission. The column density maps were determined from a pixel-to-pixel greybody fit to the red wavelength of PACS (160 μm) and SPIRE (250, 350, 500 μm). The maps were first convolved to have the same angualr reslution of 36″ and then regridded on a 14″ grid. All maps have an absolute flux calibration, using the zeroPointCorrection task in HIPE10 for SPIRE and IRAS maps for PACS. We emphasize that such a correction is indispensible for an accurate determination of column density maps. For the SED fit, we fixed the specific dust opacity per unit mass (dust+gas) approximated by the power law $\kappa_\nu = 0.1\,(\nu/1000\mathrm{GHz})^\beta$ cm^2/g and β=2, and left the dust temperature and column density as free parameters (see André et al. 2010; Könyves et al. 2010 for details). For the transformation H_2 column density into visual extinction we use the conversion formula N(H_2)/A_v= 0.94×10^{21} cm^{-2} mag^{-1} (Bohlin et al. 1978). The angular resolution of the column density maps is ∼37″ and they are shown in Figs. B.1 to B.4, together with SPIRE 250 μm images.

We estimate the final uncertainties in the *Herschel* column density maps to be around ∼30–40%, mainly arising from the uncertainty in the assumed form of the opacity law, and possible temperature gradients along the line-of-sight (see Sec. 4.1 in Russeil et al. 2013 for a quantitative discussion on the various error sources).

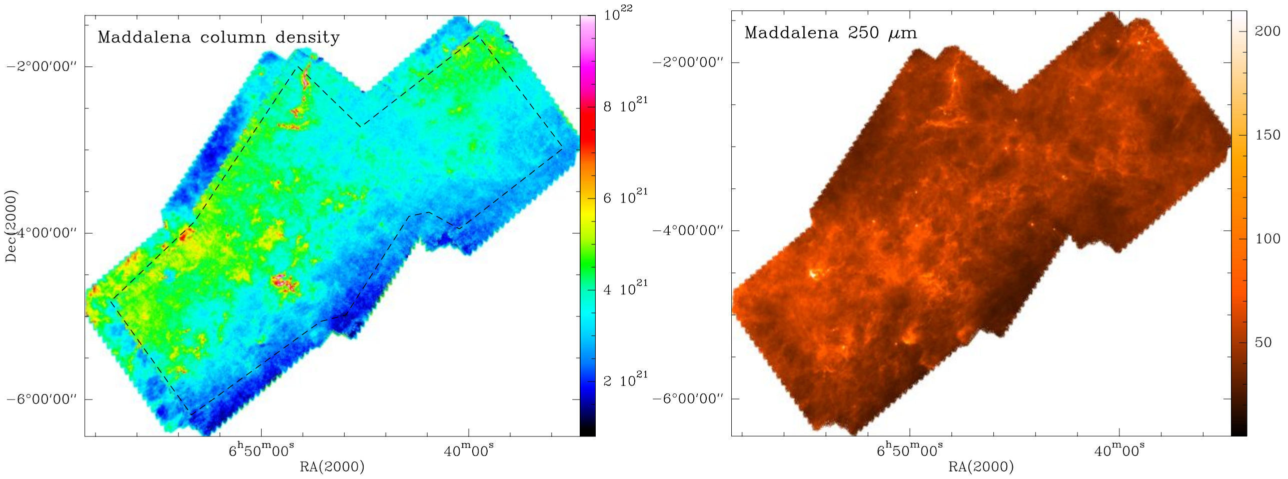


Fig. B.1. Left: *Herschel* column density map of the Maddalena cloud in [cm^{-2}]. The PDF was determined using the pixels inside the dashed area (this is the common overlap region of SPIRE and PACS). **Right:** SPIRE 250 μm map in units [MJy/sr].

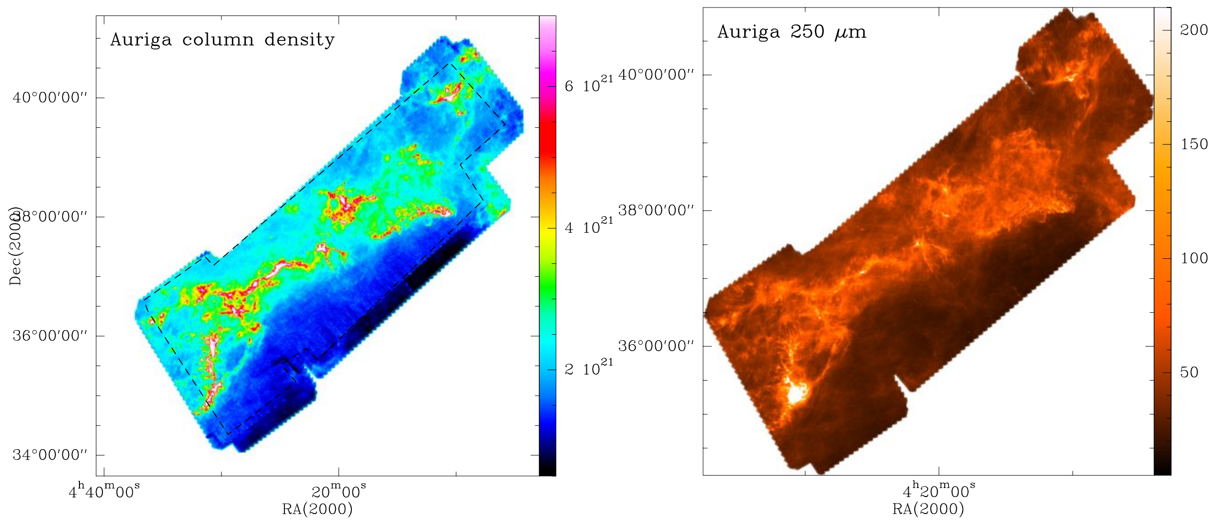


Fig. B.2. Left: *Herschel* column density map of the Auriga cloud in [cm^{-2}]. The PDF was determined using the pixels inside the dashed area (this is the common overlap region of SPIRE and PACS). **Right:** SPIRE 250 μm map in units [MJy/sr].

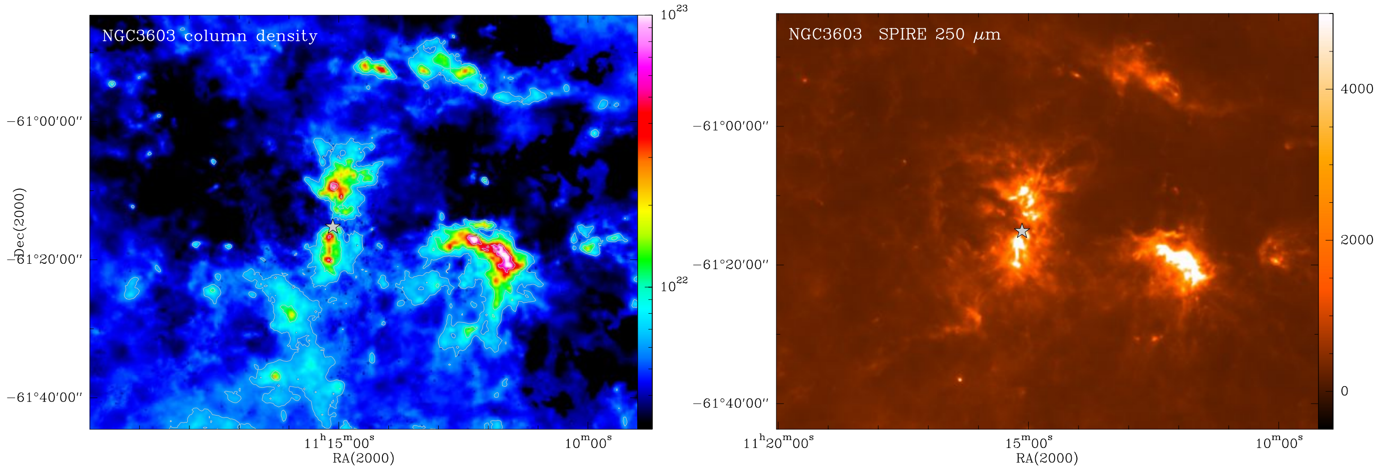


Fig. B.3. Left: *Herschel* column density map of NGC3603 in [cm^{-2}]. **Right:** SPIRE 250 μm map in units [MJy/sr]. The grey star indicates the location of the central cluster.

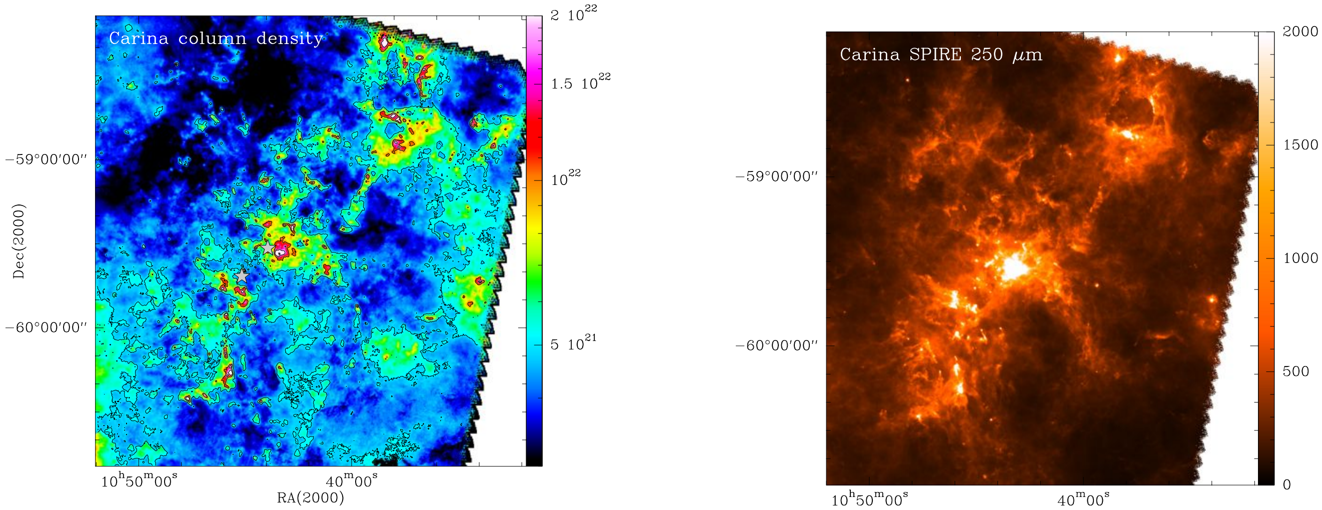


Fig. B.4. Left: *Herschel* column density map of Carina in [cm^{-2}]. **Right:** SPIRE 250 μm map in units [MJy/sr]. The grey stars indicate the location of the OB clusters Tr14 and Tr16.