

# UV-driven chemistry in simulations of the interstellar medium

## I. Post-processed chemistry with the Meudon PDR code

F. Levrier<sup>1</sup>, F. Le Petit<sup>2</sup>, P. Hennebelle<sup>1</sup>, P. Lesaffre<sup>1</sup>, M. Gerin<sup>1</sup>, and E. Falgarone<sup>1</sup>

<sup>1</sup> LERMA/LRA – ENS Paris – UMR 8112 du CNRS, 24 rue Lhomond, 75231 Paris Cedex 05, France  
e-mail: francois.levrier@ens.fr

<sup>2</sup> LUTH, Observatoire de Paris, CNRS, Université Paris Diderot, 5 place Jules Janssen, 92190 Meudon, France

Received 23 January 2012 / Accepted 30 May 2012

### ABSTRACT

**Context.** Observations have long demonstrated the molecular diversity of the diffuse interstellar medium (ISM). Only now, with the advent of high-performance computing, does it become possible for numerical simulations of astrophysical fluids to include a treatment of chemistry, to faithfully reproduce the abundances of the many observed species, and especially that of CO, which is used as a proxy for molecular hydrogen. When applying photon-dominated region (PDR) codes to describe the UV-driven chemistry of uniform density cloud models, it is found that the observed abundances of CO are not well reproduced.

**Aims.** Our main purpose is to estimate the effect of assuming uniform density on the line-of-sight in PDR chemistry models, compared to a more realistic distribution for which total gas densities may well vary by several orders of magnitude. A secondary goal of this paper is to estimate the amount of molecular hydrogen that is not properly traced by the CO ( $J = 1 \rightarrow 0$ ) line, the so-called “dark molecular gas”.

**Methods.** We used results from a magnetohydrodynamical (MHD) simulation as a model for the density structures found in a turbulent diffuse ISM with no star-formation activity. The Meudon PDR code was then applied to several lines of sight through this model, to derive their chemical structures.

**Results.** We found that compared to the uniform density assumption, maximal chemical abundances for H<sub>2</sub>, CO, CH and CN are increased by a factor  $\sim 2$ – $4$  when taking into account density fluctuations on the line of sight. The correlations between column densities of CO, CH and CN with respect to those of H<sub>2</sub> are also found to be in better overall agreement with observations. For instance, at  $N(\text{H}_2) \gtrsim 2 \times 10^{20} \text{ cm}^{-2}$ , while observations suggest that  $d[\log N(\text{CO})]/d[\log N(\text{H}_2)] \simeq 3.07 \pm 0.73$ , we find  $d[\log N(\text{CO})]/d[\log N(\text{H}_2)] \simeq 14$  when assuming uniform density, and  $d[\log N(\text{CO})]/d[\log N(\text{H}_2)] \simeq 5.2$  when including density fluctuations.

**Key words.** ISM: structure – ISM: clouds – photon-dominated region (PDR)

## 1. Introduction

The interstellar medium (ISM) is a complex system: its structure and dynamics are governed by the interaction of many processes that cover a wide range of scales, involving both micro- and macrophysics. To understand how the ISM works is also essential for understanding star and planetary system formation. This field has seen much progress in recent years, both on the observational side, with the results from the Herschel Space Observatory (Pilbratt et al. 2010; de Graauw et al. 2010), and on the theoretical side, with ever-improving numerical simulations (see e.g. Banerjee et al. 2009; Vázquez-Semadeni et al. 2010), which now consistently treat self-gravity, thermodynamics and magnetohydrodynamics (MHD).

The new challenge for numerical simulations of the ISM is now to incorporate some treatment of chemistry and grain physics, to be able to compare the results with observations of atomic and molecular lines. One possibility, followed for instance by Glover et al. (2010), is to perform multi-fluid simulations, each fluid representing a chemical species and being coupled to the others via a network of reactions. This approach has the advantage that it naturally solves for the time-dependent distribution and dynamics of the various species in three-dimensional space, but its computational cost requires making use of simplifying assumptions, most notably a small number of species and reactions. For instance,

Glover et al. (2010) treated a network of 218 reactions for 32 species, 13 of which were assumed to be in instantaneous chemical equilibrium, so that only 19 were fully treated as time-dependent quantities.

In this paper, we follow a different approach: we post-processed a single-fluid MHD simulation with the Meudon PDR code. This has the advantage of providing a full chemical network (99 species, 1362 reactions), but on the other hand, it implies a one-dimensional, steady-state treatment. With this approach, our main goal is to discuss the effects of realistic density fluctuations on photon-dominated region (PDR) chemistry, since previous studies have focused on uniform density models, e.g. Le Petit et al. (2006), or on simplified clumpy models, such as those by Wolfire et al. (2010). A second goal of this study is to estimate the amount of “dark molecular gas” that can be expected from observations of the diffuse ISM, i.e. gas where hydrogen is mostly in the form of H<sub>2</sub>, but where CO is too scarce to be seen in the usual tracer that is the ( $J = 1 \rightarrow 0$ ) line, given current sensitivities. We wish to compare this estimate with available observations (Grenier et al. 2005; Leroy et al. 2007; Abdo et al. 2010; Velusamy et al. 2010) and models (Wolfire et al. 2010).

The paper is organised as follows: Sect. 2 describes the MHD simulation used and presents the PDR code in its current state, while Sect. 3 gives an overview of the post-processing

method used to combine the two. Section 4 discusses the lines of sight selected for running our analysis. Results and comparison to observations are presented in Sect. 5. Section 6 offers a more detailed discussion and a summary. Details on using the Meudon PDR code with density profiles and our strategy for post-processing results can be found in the appendices at the end of the paper.

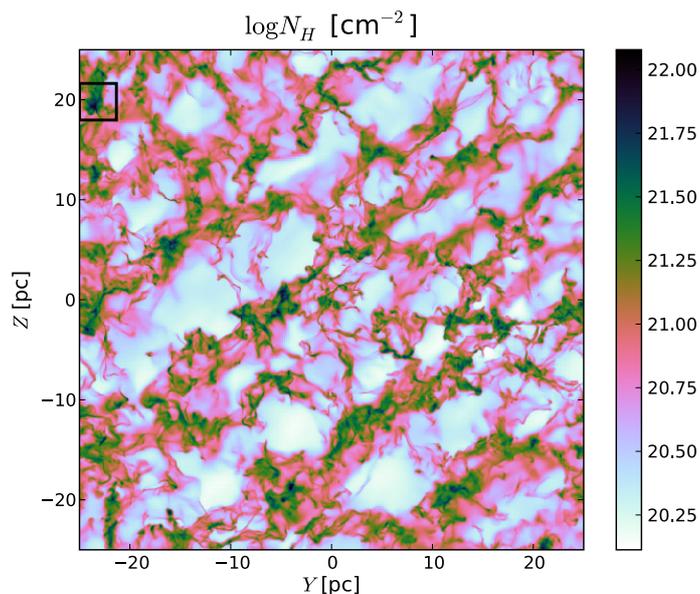
## 2. Tools

### 2.1. The MHD simulation

As a model of the turbulent diffuse ISM, we used data from a MHD simulation performed by Hennebelle et al. (2008) using the RAMSES code (Teyssier 2002; Fromang et al. 2006). A notable advantage of this code is its adaptive mesh refinement (AMR) capabilities, which enable to locally reach extremely high spatial resolutions. The initial setup of the simulation is a cube of homogeneous warm neutral gas (WNM) of atomic gas (a 10:1 mixture of H and He) with density  $n_{\text{H}} = 1 \text{ cm}^{-3}$  and temperature  $T = 8000 \text{ K}$ . The cube is  $L = 50 \text{ pc}$  on each side, and two converging flows of WNM are injected from opposing faces along the  $X$  axis with relative velocity  $\Delta V_X \sim 40 \text{ km s}^{-1}$ , which means that the Mach number of each flow with respect to the ambient WNM is  $\mathcal{M} \sim 2$ . Transverse and longitudinal velocity modulations are imposed, with amplitudes roughly equal to the mean flow velocity ( $20 \text{ km s}^{-1}$ ). Periodic boundary conditions are applied on the remaining four faces. The simulation starts on a  $256^3$  grid, with two extra levels of refinement based on density thresholds, so that the effective resolution of the simulation is  $\sim 0.05 \text{ pc}$ . For our purposes, we regridded the data regularly on a  $1024^3$  cube, therefore using the maximum resolution over the whole domain. A magnetic field is present in the simulation, and is initially parallel to the  $X$  axis with an intensity of about  $5 \mu\text{G}$ , consistent with observational values at these densities (Crutcher et al. 2010). The converging flows collide near the midplane  $X = 0$  about 1 Myr into the simulation, and as the total mass grows from  $\sim 3000 M_{\odot}$  initially to about 10 times this value at the end of the simulation ( $t = 13.11 \text{ Myr}$ ), gravity eventually takes over and, combined with the effects of thermal instability (Field 1965; Hennebelle & Péroult 2000; Koyama & Inutsuka 2002; Heitsch et al. 2005; Hennebelle & Audit 2007; Banerjee et al. 2009), leads to the formation of cold dense clumps ( $n_{\text{H}} > 100 \text{ cm}^{-3}$ , while  $T \sim 10\text{--}50 \text{ K}$ ) within a much more diffuse and warm interclump medium ( $n_{\text{H}} \sim 1\text{--}10 \text{ cm}^{-3}$  and  $T \sim 10^3\text{--}10^4 \text{ K}$ ). This occurs at  $t \sim 12 \text{ Myr}$ . See Hennebelle et al. (2008) for a more detailed description. Figure 1 shows the column density of the gas viewed along the  $X$  axis, which is that of the incoming flows.

### 2.2. The Meudon PDR code

The Meudon PDR code (<http://pdr.obspm.fr/>) is a publicly available set of routines (Le Bourlot et al. 1993; Le Petit et al. 2006; Gonzalez Garcia et al. 2008) whose purpose is to describe the UV-driven chemistry of interstellar clouds, and in particular of PDR. It is a steady-state one-dimensional code in which a plane-parallel slab of gas and dust is illuminated on either or both sides by the light from a star or by the standard interstellar radiation field (ISRF), which is defined using expressions from Mathis et al. (1983) and Black (1994). Usually run with homogeneous gas densities, the code can also accept input of density profiles via a text file. At each point along the line of sight, radiative transfer in the UV is treated



**Fig. 1.** Total gas column density along the  $X$  axis of the MHD simulation snapshot used here. The box in the upper left side marks the position of the clump selected for running our analysis.

to solve for the H/H<sub>2</sub> transition, using either the approximations by Federman et al. (1979), or an exact method based on a spherical harmonics expansion of the specific intensity (Goicoechea & Le Bourlot 2007). A number of heating (photoelectric effect on grains, cosmic rays) and cooling (infrared and millimeter emission lines) processes contribute to the computation of thermal balance. Outputs of the code include gas properties such as temperature and ionisation fraction, radiation energy density, chemical abundances and column densities, level populations and line intensities. The code is iterative and therefore requires the user to check whether convergence has been achieved.

## 3. Method overview

Naively, one would like to run the PDR code on all lines of sight through the simulated cube to derive a three-dimensional chemical structure, as well as line-of-sight integrated observables (an emission map for the CII [158  $\mu\text{m}$ ] line, for instance). However, this is neither feasible nor desirable.

Two computational reasons preclude this brute force approach. Firstly, the cost is prohibitive: just one global iteration of the PDR code for a single run typically completes in a couple hours on a GNU/Linux machine equipped with four dual-core 64-bit x86 processors and 64GB of memory. Because a run usually converges in some ten iterations, it takes about a day to process a single line of sight. The PDR code, however, has been ported on the EGEE grid<sup>1</sup>, which allows us to perform  $\sim 100$  runs simultaneously in the same timeframe. Still, it was not reasonable to consider treating more than  $\sim 10^3$  lines of sight in this work.

The second computational problem to consider is that the PDR code has trouble converging in low-density regions, of which there are many in the MHD simulation. These are the vicinities of  $X = \pm 25 \text{ pc}$ , where WNM gas ( $n_{\text{H}} = 1 \text{ cm}^{-3}$ ) is entering the box, but low-density regions are found everywhere throughout the cube (the volume filling factor of regions where  $n_{\text{H}} \leq 20 \text{ cm}^{-3}$  is  $f_v = 0.96$ ). This convergence problem might

<sup>1</sup> <http://www.eu-egee.org/>

be related to the fact that Ly- $\alpha$  emission, which is a main cooling process for  $n_{\text{H}} \lesssim 5 \text{ cm}^{-3}$ , is not included in the code. This renders the computation of thermal balance by the PDR code in these regions unreliable.

Besides these computational hurdles, it should be noted that the code is one-dimensional, and treats a density profile as that of a plane-parallel slab of gas. Imagine then that a line of sight intercepts a dense clump of matter: the gas lying directly behind that clump will be shielded from incoming radiation, since the most energetic UV photons will have been absorbed by the clump. This leads to shadowing artefacts in regions that in reality may well be illuminated from other directions, because the ISM has a complex, fractal-like, and evolving structure.

In this paper, we deal with these artefacts in the following way: the physical conditions and chemical composition at every grid point  $(X, Y, Z)$  considered in the analysis may be obtained by running the PDR code in  $p$  directions going through that grid point, for instance along the main coordinate axes. Thus, each quantity  $F(X, Y, Z)$  output by the code has  $p$  possible values  $f_1, f_2, \dots, f_p$ . This is the case, in particular, of the radiation energy density  $E$ , which has possible values  $e_1, e_2, \dots, e_p$ . To combine the  $p$  runs at this grid point, we selected a direction  $p_0$  for which the radiation energy density there is maximum,  $e_{p_0} = \max\{e_i\}_{1 \leq i \leq p}$ . This choice is discussed in Sect. 6. Of course  $p_0$  is a function of  $(X, Y, Z)$ . We then chose  $F(X, Y, Z) = f_{p_0}$  for every quantity  $F$  output by the code. This procedure takes better account of the porosity of the simulated structures to the ISRF, while ensuring element conservation at each grid point. Ideally, the more directions  $p$ , the better, but this obviously comes with an increased computational cost. In this paper, we chose  $p = 2$  as a compromise, running the PDR code in two orthogonal directions.

The analysis in this paper focuses on a small subset of a single simulation snapshot, which we discuss in the next section. After applying a density threshold  $n_0 = 20 \text{ cm}^{-3}$ , which is the lowest value considered in the grid of models run by [Le Petit et al. \(2006\)](#) and therefore deemed sufficient to ensure convergence, we extracted a number of one-dimensional density profiles from this thresholded subset. We ran the PDR code on these profiles and combined results to derive the chemical structure of the clump.

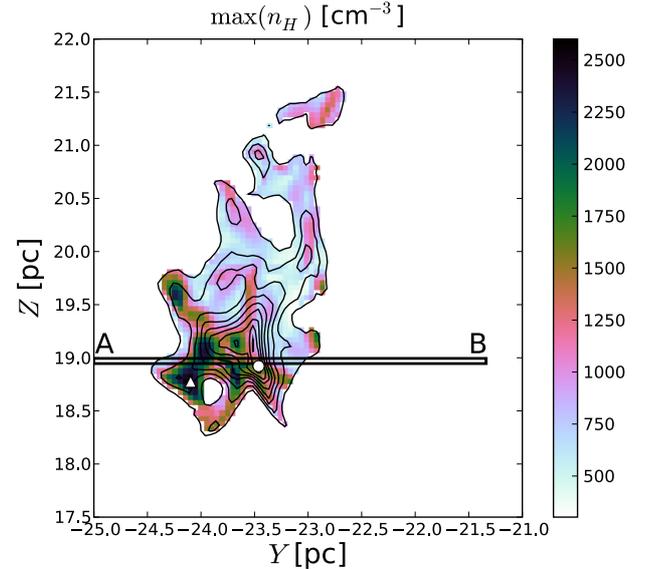
## 4. Lines of sight selected for PDR computations

### 4.1. Simulation snapshot

The snapshot chosen to run our analysis on was timed at 7.35 Myr, when the densest parts of the cloud reach  $n_{\text{max}} \sim 9 \times 10^3 \text{ cm}^{-3}$ . Some of the structures present in the simulation at this time are self-gravitating, but we are confident that they are still diffuse enough that the simulation snapshot is representative of a non-starforming region of the ISM. We therefore ran our analysis without any illuminating star, with the ISRF being the only source of primary UV photons.

### 4.2. Selected clump

To select a representative subset, we note that observational PDRs such as the Horsehead Nebula (see e.g. [Pety et al. 2007](#)) are found at the edge of dense and cold clouds of gas and dust, illuminated by ambient FUV light and possibly nearby young stars. It thus makes sense to focus on a ‘‘clump’’, defined observationally as a connected structure with a significantly higher column density than its surroundings. We identified clumps via a



**Fig. 2.** Maximum gas density  $n_{\text{H}}$  along the lines of sight parallel to the  $X$  axis within the selected clump. Contours show the total gas column density  $N_{\text{H}}$  from  $3 \times 10^{21} \text{ cm}^{-2}$  to  $1.1 \times 10^{22} \text{ cm}^{-2}$  in steps of  $10^{21} \text{ cm}^{-2}$ . The ‘‘clump’’ displayed here has a size roughly  $1.5 \text{ pc} \times 2.5 \text{ pc}$ . The two-dimensional slab of gas under study is seen projected as the single-pixel-wide **AB** strip. The white triangle marks the position of the maximum value of  $\max(n_{\text{H}})$  in this region, and the white circle that of the maximum of  $N_{\text{H}}$ . They are separated by 0.65 pc.

**Table 1.** Properties of the selected observational clump.

Average density	$\langle n_{\text{H}} \rangle$	$30 \text{ cm}^{-3}$
Average temperature	$\bar{T}$	270 K
Line-of-sight centroid velocity	$\bar{V}_X$	$-0.12 \text{ km s}^{-1}$
Mass	$M$	$124 M_{\odot}$
Turbulent dispersion	$\sigma_{3\text{D}}$	$1.8 \text{ km s}^{-1}$

**Notes.**  $\langle F \rangle$  refers to the direct average and  $\bar{F}$  to the density-weighted average of any quantity  $F$  in this table. The mass is computed as  $M = \mathcal{V} \mu m_{\text{H}} \langle n_{\text{H}} \rangle$ , where  $\mathcal{V}$  is the three-dimensional volume corresponding to the two-dimensional clump,  $m_{\text{H}} = 1.66 \times 10^{-24} \text{ g}$  is the mass of the hydrogen atom and  $\mu = 1.4$  corresponds to a 1:10 number ratio for He with respect to H. The turbulent velocity dispersion includes all three velocity components,  $\sigma_{3\text{D}}^2 = \sigma_X^2 + \sigma_Y^2 + \sigma_Z^2$ .

friend-of-friend algorithm on the column density map along the  $X$  axis (Fig. 1), using a threshold  $N_0 = 3 \times 10^{21} \text{ cm}^{-2}$ , which corresponds to a mean density  $\langle n_{\text{H}} \rangle = n_0 = 20 \text{ cm}^{-3}$  over 50 pc.

The selected clump, which lies at the top left corner of the simulation’s field, harbours an interesting feature: Fig. 2 shows in colour scale the map of the maximum gas density  $\max(n_{\text{H}})$  encountered along the  $X$  axis for every line of sight within the clump. It so happens that the peak of  $N_{\text{H}}$  does not match that of  $\max(n_{\text{H}})$ , or even a local maximum of the latter. This is important to note for species that may be sensitive to the local gas density rather than to the total column density. Properties of that selected clump are listed in Table 1.

### 4.3. Selected lines of sight

The observational clump shown in Fig. 2 has a size  $\Delta Y \times \Delta Z \simeq 1.5 \text{ pc} \times 2.5 \text{ pc}$ . In the  $X$  direction, most of its gas is located in the central  $\Delta X \simeq 15 \text{ pc}$  around  $X = 0$ . Given the pixel size

$\delta \simeq 0.05$  pc, applying the PDR code on a structure of that size along the three coordinate axes requires some 25 000 runs, which is beyond the scope of this work. Consequently, we restricted our study to a two-dimensional slab of gas across the observational clump. It is projected in Fig. 2 as the one-pixel-wide strip **AB**, which is therefore  $\sim 3.7$  pc long,  $\sim 0.05$  pc wide and contains 76 lines of sight along the  $X$  axis. These sample a wide range of column densities, from  $N_{\text{H}} = 6.11 \times 10^{20} \text{ cm}^{-2}$  (near the **A** end) to  $N_{\text{H}} = 1.11 \times 10^{22} \text{ cm}^{-2}$ , corresponding to visual extinctions  $A_{\text{V}} = 0.33$  to  $A_{\text{V}} = 5.93$ , using the conversion from total hydrogen column density  $N_{\text{H}} = N(\text{H}) + 2N(\text{H}_2)$

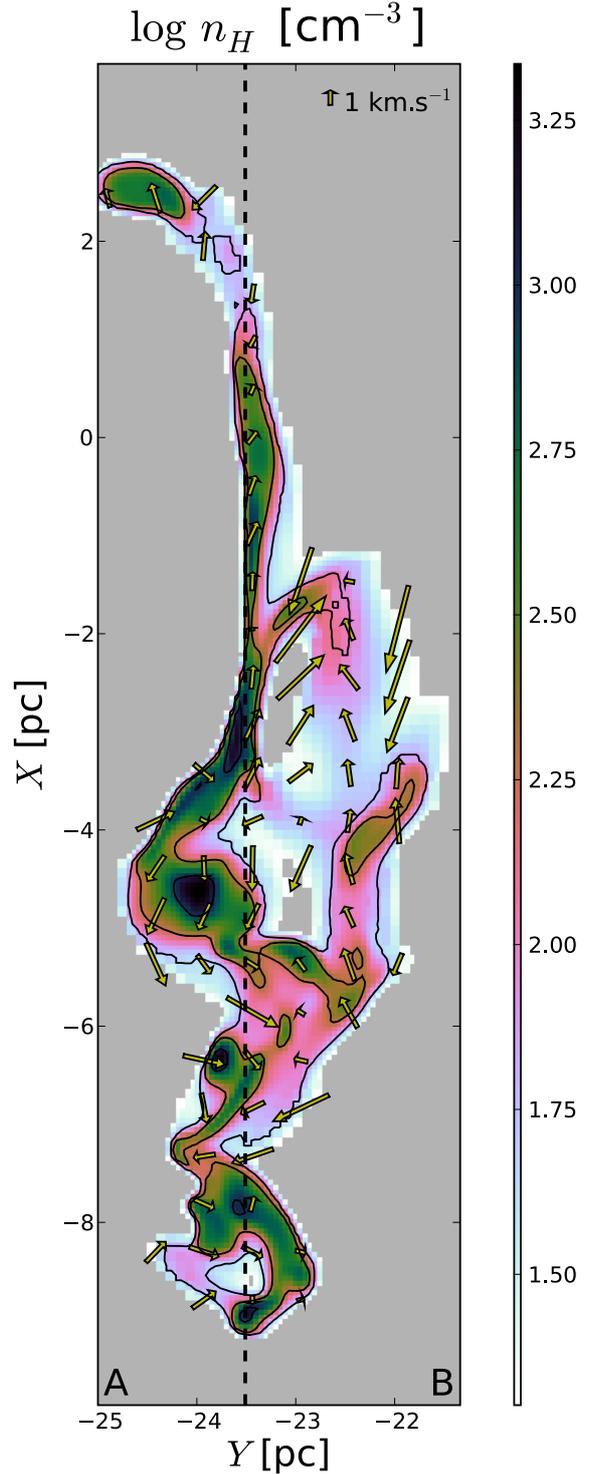
$$A_{\text{V}} = \frac{R_{\text{V}}}{C_{\text{D}}} \left( \frac{N_{\text{H}}}{1 \text{ cm}^{-2}} \right) \quad (1)$$

with  $R_{\text{V}} = 3.1$  and  $C_{\text{D}} = 5.8 \times 10^{21} \text{ cm}^{-2} \text{ mag}^{-1}$  (see Table A.1 and Le Petit et al. 2006). The structure of the gas on these lines of sight (Fig. 3) is complex, with many small, dense and cold regions ( $n_{\text{H}} \simeq 10^3 \text{ cm}^{-3}$ ,  $T \simeq 20$  K) interconnected via a filamentary structure and embedded within a much more diffuse and warm medium ( $n_{\text{H}} \simeq 1 \text{ cm}^{-3}$ ,  $T \simeq 10^4$  K). The overdense regions ( $n_{\text{H}} \geq n_0$ ) are located near the midplane of the simulation ( $-9 \text{ pc} \lesssim X \lesssim 2 \text{ pc}$ ), where the flows collide and the gas condenses into cold structures.

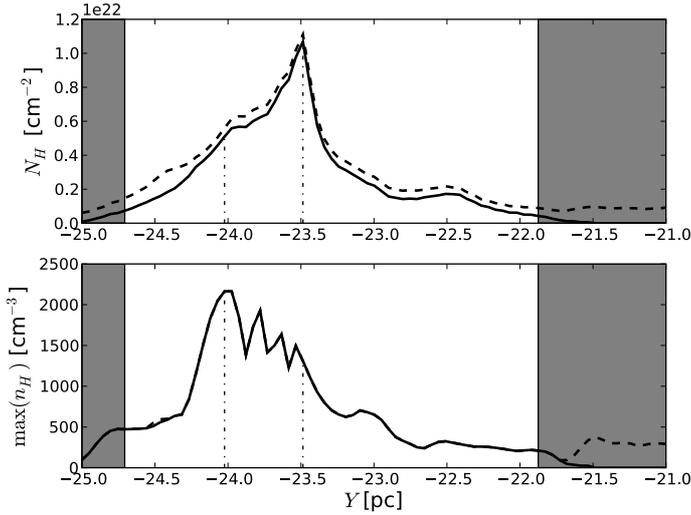
It is this subset of the simulated cube, shown in Fig. 3, which is the focus of our study, and from which we extracted one-dimensional density profiles. As Fig. 4 shows, this subset indeed contains most of the gas on the lines of sight within the **AB** strip: Except in the outermost regions in which  $N_{\text{H}} \leq 1.5 \times 10^{21} \text{ cm}^{-2}$ , column densities along  $X$  over that region represent more than half the total column densities over the full 50 pc lines of sight. Figure 4 also shows that within the **AB** strip the peaks of  $N_{\text{H}}$  and  $\max(n_{\text{H}})$  are still separated by about 0.55 pc.

As explained in Sect. 3, we considered two ( $p = 2$ ) possible directions for the one-dimensional density profiles extracted from the two-dimensional subset, namely those parallel to the  $X$  or  $Y$  axis. The dashed line in Fig. 3 marks the location of such an extracted profile, which is shown in Fig. 5. Note that we considered profiles to be connectedly overdense, which means that along each line of sight, we may extract several profiles separated by underdense regions  $n_{\text{H}} < n_0$ . For instance, this is the case of the line of sight parallel to the  $X$  axis located at  $Y = -23$  pc.

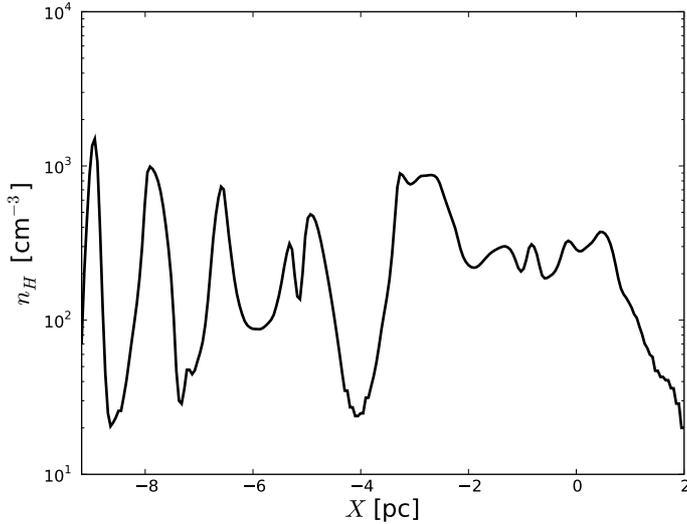
We thus extracted 447 density profiles, of which 156 were parallel to the  $X$  axis and 291 were parallel to the  $Y$  axis. Their statistical properties are summarised in Tables 2 and 3, emphasising the wide dynamic range they sample in column densities ( $\sim 10^4$ ) and mean densities ( $\sim 30$ ). The high values in density-weighted temperatures correspond to density profiles that never much deviate from  $n_0 = 20 \text{ cm}^{-3}$ . For each of these 447 profiles, we ran the PDR code assuming identical illumination on both sides. Since one of the objectives of this paper is to assess the effect of realistic density distributions along the line of sight on the chemical composition of interstellar clouds, we also applied the PDR code to a homogeneous reference model for each extracted profile. We specify this model, called `uniform`, as having the same mean density  $\langle n_{\text{H}} \rangle$  and total visual extinction  $A_{\text{V}}$  as the inhomogeneous model, which we dub `los` from now on. Unless otherwise specified, results presented in the next section refer to these `los` models. The setup for all runs is detailed in Appendix A and their post-processing is described in Appendix B.



**Fig. 3.** Structure of the gas in the two-dimensional slab under study. This is a close-up view on the region  $X \simeq 0$  in which the incoming flows of WNM collide to form cold structures, and only regions in which  $n_{\text{H}} \geq n_0 = 20 \text{ cm}^{-3}$  are shown. The colour image shows the gas density  $n_{\text{H}}$  in logarithmic scale, while contours show the gas temperature  $T_{\text{MHD}}$  at 20 K (higher densities), 30 K and 50 K (lower densities). The two-dimensional projection of the velocity field ( $V_x, V_y$ ) is also shown as yellow arrows, the length of which indicate the velocity modulus at that location (at the centre of each arrow). The **A** and **B** extremities of the observational one-dimensional strip are indicated for reference, and the dashed line marks an example location for the extracted density profiles on which the PDR code is run (Fig. 5). The grey areas are outside of the domain used for PDR computations.



**Fig. 4.** Total gas column densities along the  $X$  axis within the **AB** strip under study (*top plot*) and maximum gas density  $\max(n_{\text{H}})$  on the same lines of sight (*bottom plot*). Shown are the column densities over the full 50 pc lines of sight along the  $X$  axis (dashed line) and the column densities for the overdense regions  $n_{\text{H}} \geq n_0$  shown in Fig. 3 (solid line). Grey areas mark lines of sight for which less than 50% of the mass is in the overdense region. The dash-dotted lines mark the positions of the maxima of  $N_{\text{H}}$  and  $\max(n_{\text{H}})$ , which are separated by 0.55 pc.



**Fig. 5.** Example of a density profile used in the PDR code. This is the profile extracted at the location of the dashed line in Fig. 3.

## 5. Results

### 5.1. Temperature comparison

The PDR code and the MHD simulation both treat thermal balance, so that we have two estimates of the gas temperature, which we can compare: Fig. 6 shows the ratio  $r = T_{\text{PDR}}/T_{\text{MHD}}$  of the gas temperature  $T_{\text{PDR}}$  output by the PDR code to the gas temperature  $T_{\text{MHD}}$  computed in the MHD simulation at every point in the subset under study versus the total gas density  $n_{\text{H}}$  at that point. Average ratios  $\langle r \rangle$  in selected density bins are also shown. It appears quite clearly that  $\langle r \rangle$  is close to 1, with  $0.3 \lesssim \langle r \rangle \lesssim 2.0$  over the whole range of densities.

The fact that  $T_{\text{PDR}} \sim T_{\text{MHD}}$  comes as a pleasant surprise, considering the differences between the PDR and MHD computations: while the former is one-dimensional, steady-state, and

**Table 2.** Properties of the 156 one-dimensional profiles extracted parallel to the  $X$  axis.

Parameter ( $F$ )	$\min(F)$	$\max(F)$	$\langle F \rangle$
Size $L$ [pc]	0.15	11.2	2.18
Density $\langle n_{\text{H}} \rangle$ [ $\text{cm}^{-3}$ ]	20	571	155
Column density $N_{\text{H}}$ [ $10^{20} \text{ cm}^{-2}$ ]	0.117	107	11.9
Visual extinction $A_{\text{V}}$	$6.3 \times 10^{-3}$	5.7	0.64
Temperature $\bar{T}$ [K]	22	924	88
Velocity dispersion $\sigma_{\text{X}}$ [ $\text{km s}^{-1}$ ]	–	1.9	0.50

**Notes.** Listed are the minimum, maximum and ensemble average values for the size, average density, column density, visual extinction (corresponding to  $N_{\text{H}}$  via Eq. (1)), density-weighted average temperature and line-of-sight velocity dispersion. For that last quantity, the minimum value is not shown, because it is too low to be meaningful.

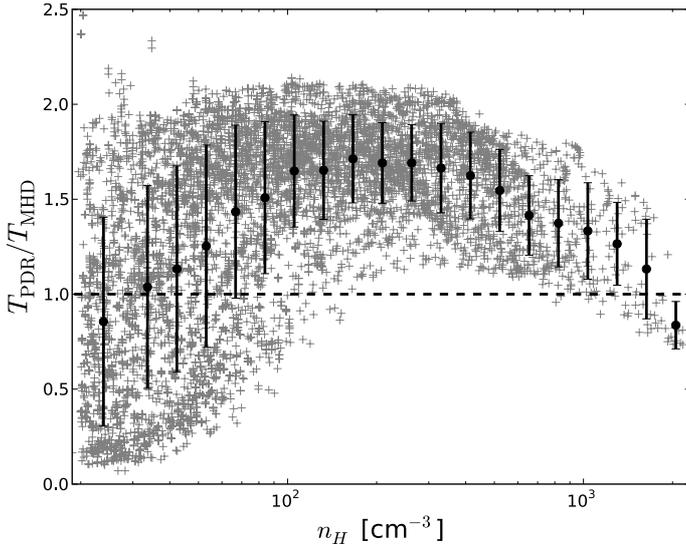
**Table 3.** Same as Table 2 but for the 291 one-dimensional profiles extracted parallel to the  $Y$  axis.

Parameter ( $F$ )	$\min(F)$	$\max(F)$	$\langle F \rangle$
Size $L$ [pc]	0.24	2.78	1.17
Density $\langle n_{\text{H}} \rangle$ [ $\text{cm}^{-3}$ ]	22	655	188
Column density $N_{\text{H}}$ [ $10^{20} \text{ cm}^{-2}$ ]	0.165	31.6	6.39
Visual extinction $A_{\text{V}}$	$8.8 \times 10^{-3}$	1.7	0.34
Temperature $\bar{T}$ [K]	21	464	56
Velocity dispersion $\sigma_{\text{Y}}$ [ $\text{km s}^{-1}$ ]	–	1.51	0.37

includes cooling via the infrared and submillimeter lines from atomic and molecular species, especially  $\text{H}_2$  transitions (Le Petit et al. 2006), the latter is three-dimensional, dynamical, and only includes cooling via the fine structure lines of  $\text{CII}$  [ $158 \mu\text{m}$ ] and  $\text{OI}$  [ $63 \mu\text{m}$ ], as well as the recombination of electrons with ionised PAHs. This suggests that unless a very precise knowledge of the temperature is needed, it is probably not necessary, at least as a first approximation and in the range of densities and temperatures probed here, to refine the details of cooling processes in our MHD simulations, because the simple cooling function currently used already yields gas temperatures close to those found using the more detailed processes of the PDR code. Similar conclusions were reached by Glover et al. (2010).

### 5.2. Chemical structure and comparison to observations

The spatial distributions of  $\text{H}$ ,  $\text{H}_2$ ,  $\text{C}^+$ ,  $\text{C}$ ,  $\text{CO}$ ,  $\text{CS}$ ,  $\text{CH}$ , and  $\text{CN}$  in the simulation subset are shown in Figs. 7 and 8. In these figures, points were clipped where the density  $n(\alpha)$  of species  $\alpha$  was below  $n_{0,\alpha} = X_{\alpha}n_0$ , with  $X_{\alpha}$  a typical threshold abundance for the detection of that species (see Table 4). Although shadowing effects remain (for instance in the atomic hydrogen map), these figures show how some species (e.g.  $\text{CO}$ ,  $\text{CN}$ ) trace denser gas than others ( $\text{C}^+$ ,  $\text{CH}$ ). This appears more clearly when plotting these abundances, averaged over density bins, versus total gas density  $n_{\text{H}}$  (Fig. 9):  $\text{C}^+$  traces gas uniformly up to  $n_{\text{H}} \geq 10^3 \text{ cm}^{-3}$ , while  $\text{CO}$  starts rising up at  $n_{\text{H}} \geq 250 \text{ cm}^{-3}$ , right about where  $\text{CH}$  flattens out. This break in the slope for  $\text{CO}$  occurs after the molecular transition  $\langle f_{\text{H}_2} \rangle = 2 \langle n(\text{H}_2) \rangle / n_{\text{H}} = 1/2$ , which is at  $n_{\text{H}} \approx 100 \text{ cm}^{-3}$ . Considering the abundances of  $\text{C}$  and  $\text{C}^+$ , this means that a significant fraction of the molecular gas (i.e. where hydrogen is mostly in the form of  $\text{H}_2$ ) is better traced by  $\text{C}$  and  $\text{C}^+$  than by  $\text{CO}$ . This “dark molecular gas” fraction is the subject of Sect. 5.4.  $\text{CH}$ , on the other hand, nicely follows  $\text{H}_2$  (Sheffer et al. 2008). To complete the picture,  $\text{C}$  and  $\text{CN}$



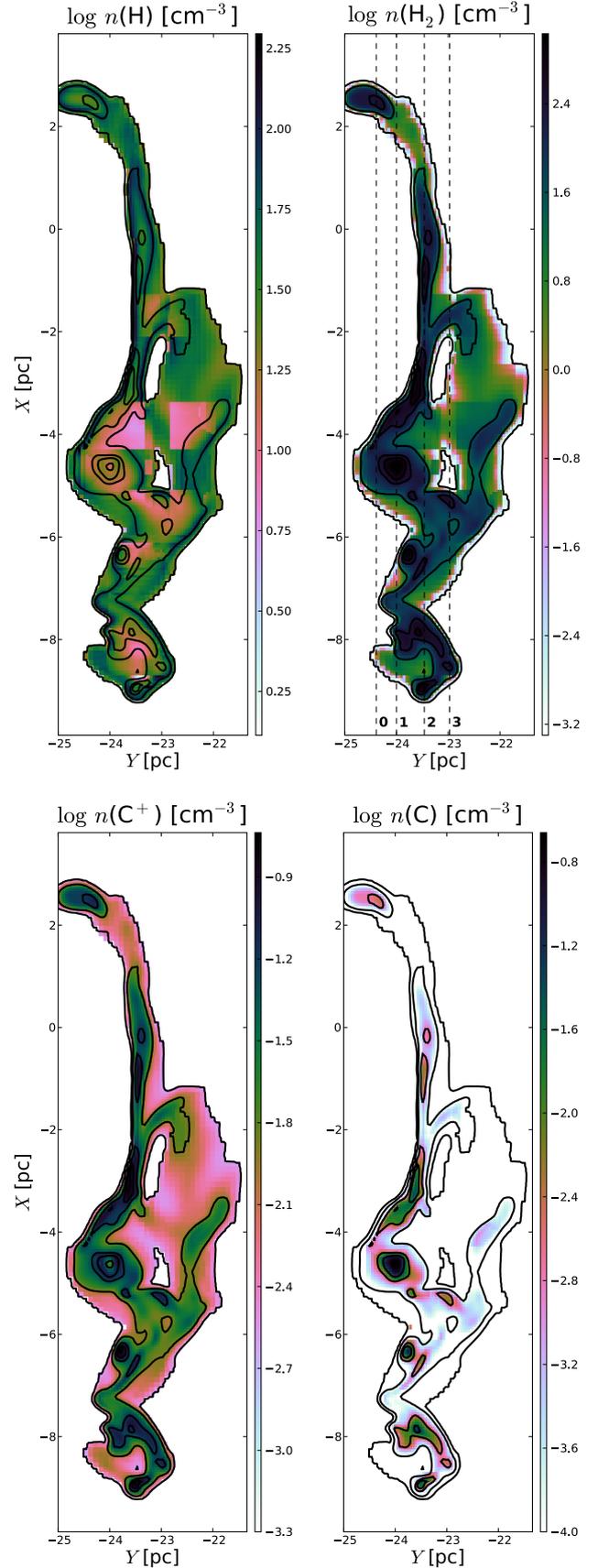
**Fig. 6.** Ratio of the gas temperatures  $T_{\text{PDR}}/T_{\text{MHD}}$  versus total gas density  $n_{\text{H}}$ . Grey crosses show all points, while black circles are average values over density bins in logarithmic scale, with error bars standing for  $\pm 1\sigma$ . The dashed line corresponds to  $T_{\text{PDR}}/T_{\text{MHD}} = 1$ .

have a similar slope throughout the density range, although CN seems to break away slightly at  $n_{\text{H}} \gtrsim 10^3 \text{ cm}^{-3}$ , to follow CO.

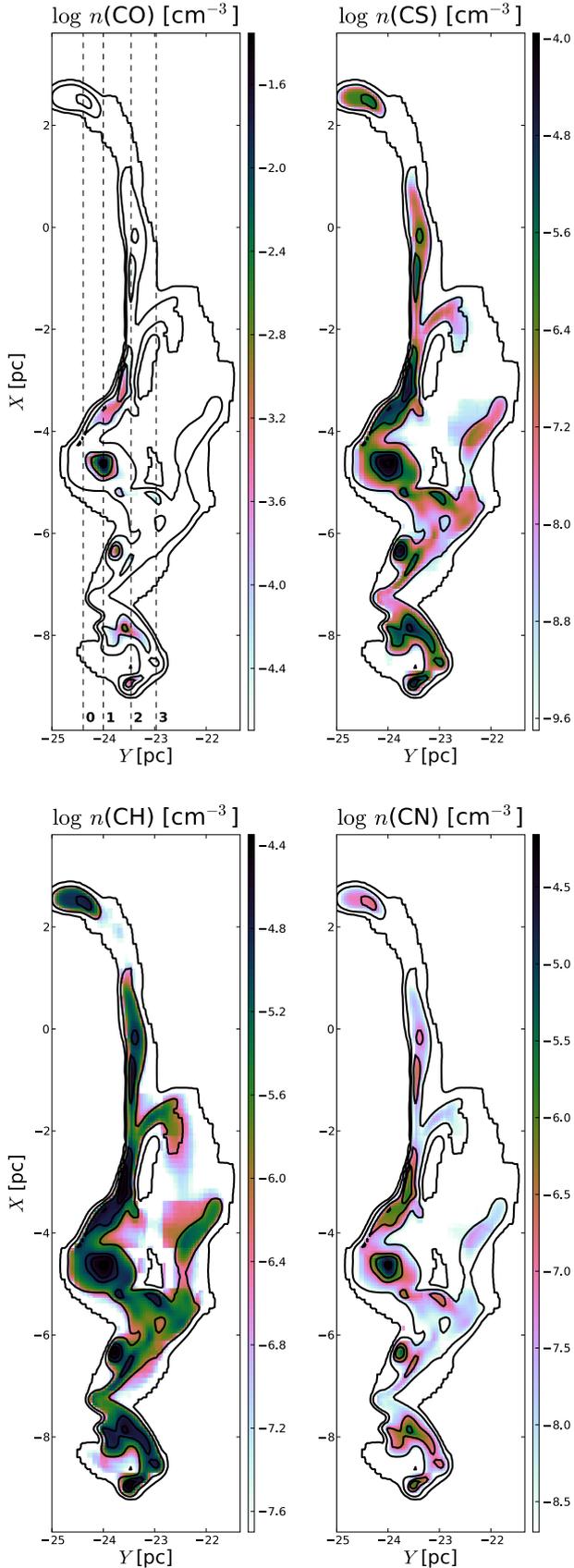
To be more precise on these apparent correlations (CH vs.  $\text{H}_2$ , CS vs. C and CN vs. CO), we show in Fig. 10 abundance ratios for these similarly distributed species in the region where they are all significantly present, that is the clouddlet at ( $X \simeq -4.7 \text{ pc}, Y \simeq -24 \text{ pc}$ ) (see Figs. 7 and 8). We can see that  $n(\text{CH})/n(\text{H}_2)$  and  $n(\text{CN})/n(\text{CO})$  have a similar “ring-like” behaviour, rising to maximum values  $n(\text{CH})/n(\text{H}_2) \sim 5.5 \times 10^{-8}$  and  $n(\text{CN})/n(\text{CO}) \sim 10^{-2}$  at total gas densities  $n_{\text{H}} \sim 400\text{--}500 \text{ cm}^{-3}$ , then falling at higher densities, to  $n(\text{CH})/n(\text{H}_2) \sim 4 \times 10^{-8}$  and  $n(\text{CN})/n(\text{CO}) \sim 2 \times 10^{-3}$ , respectively. For  $n(\text{CS})/n(\text{C})$ , there is a slight loss of azimuthal symmetry, but the overall trend is the same, although maximum values of  $\sim 10^{-3}$  are reached at higher densities  $n_{\text{H}} \sim 1000 \text{ cm}^{-3}$  before falling to  $\sim 3 \times 10^{-4}$  at the peak.

The comparison to observational data requires us to work not with densities but with column densities, which are what observers have access to. To his end, we computed column densities for CO, CH and CN on each line of sight parallel to the X or Y axis and plot them versus those of  $\text{H}_2$  in Fig. 11 to match the observational data plots in Sheffer et al. (2008). We plot data points that correspond to the long lines of sight parallel to X (squares) and to the shorter lines of sight parallel to Y (circles) using a colour scheme to specify the mean gas density  $\langle n_{\text{H}} \rangle$  on the line of sight. Fits to observational data derived by Sheffer et al. (2008) are shown as dashed lines, and in the CO panel we also plot actual data points from that same paper.

Our data shows a significant deficit around  $N(\text{H}_2) \sim 10^{20} \text{ cm}^{-2}$  for CO, which may be a general problem with PDR chemistry computations (Sonnentrucker et al. 2007). However, the agreement with the observational fit improves, both in values and in slope, for  $N(\text{H}_2) \gtrsim 2 \times 10^{20} \text{ cm}^{-2}$ . The slope found is  $d[\log N(\text{CO})]/d[\log N(\text{H}_2)] \simeq 5.2$ , while the observational fit yields  $3.07 \pm 0.73$ . On the longer lines of sight, we can see a “loop” structure that needs to be understood. It is easily identified with lines of sight between  $Y = -24.5 \text{ pc}$  and  $Y = -23 \text{ pc}$  (positions 0 to 3 on the  $\text{H}_2$  and CO maps from Figs. 7 and 8). To be more accurate, following the loop clockwise corresponds



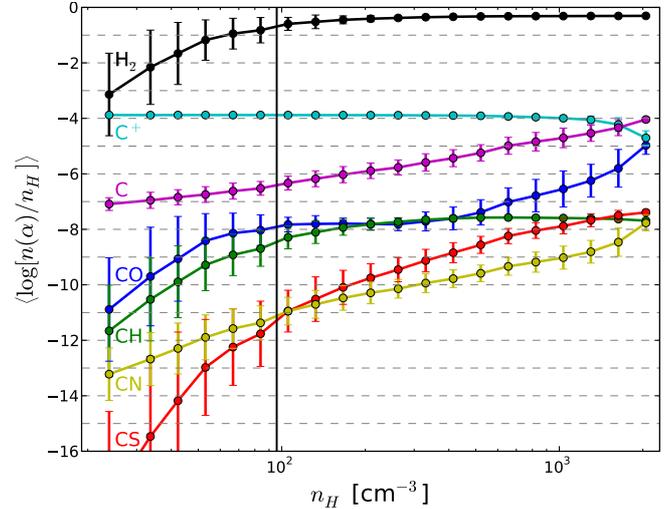
**Fig. 7.** H (top left),  $\text{H}_2$  (top right),  $\text{C}^+$  (bottom left) and C (bottom right) abundances for the los models. Contours mark total gas densities 20, 100, 500, 1000, and 2000  $\text{cm}^{-3}$ .  $\text{C}^+$  and C abundance maps are clipped below  $2 \times 10^{-4} \text{ cm}^{-3}$  and  $10^{-4} \text{ cm}^{-3}$ , respectively. Lines of sight 0, 1, 2, and 3 in the  $n(\text{H}_2)$  map refer to the discussion in the text.



**Fig. 8.** Same as Fig. 7 but for CO (top left), CS (top right), CH (bottom left), and CN (bottom right). Abundance maps are clipped below  $2 \times 10^{-5} \text{ cm}^{-3}$  for CO,  $2 \times 10^{-10} \text{ cm}^{-3}$  for CS,  $2 \times 10^{-8} \text{ cm}^{-3}$  for CH and  $2 \times 10^{-9} \text{ cm}^{-3}$  for CN.

**Table 4.** Typical threshold abundances  $X_\alpha$  used in Figs. 7 and 8.

$\text{C}^+ : 10^{-5}$	$\text{CO} : 10^{-6}$	$\text{CH} : 10^{-9}$
$\text{C} : 5 \times 10^{-6}$	$\text{CS} : 10^{-11}$	$\text{CN} : 10^{-10}$



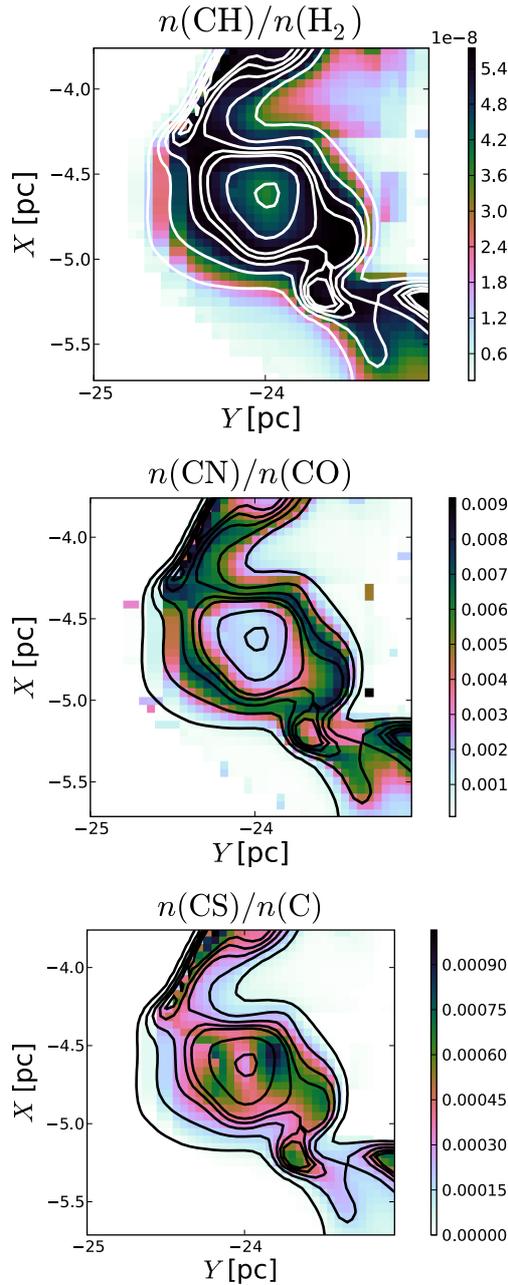
**Fig. 9.** Abundances of  $\text{H}_2$ ,  $\text{C}^+$ ,  $\text{C}$ ,  $\text{CO}$ ,  $\text{CH}$ ,  $\text{CS}$  and  $\text{CN}$  versus total gas density  $n_{\text{H}}$ . Data points are averaged in the same  $n_{\text{H}}$  bins as in Fig. 6. The vertical line marks the position of the average molecular transition where  $\langle f_{\text{H}_2} \rangle = 2 \langle n(\text{H}_2) \rangle / n_{\text{H}} = 1/2$ .

to scanning lines of sight parallel to  $X$  from **0** to **3**, with turnovers at positions **1** and **2**. Considering the  $\text{H}_2$  and  $\text{CO}$  maps alongside Fig. 11 helps to understand this “loop”. Firstly, lines of sight from **0** to **1** basically intercept just one dense structure with both  $\text{H}_2$  and  $\text{CO}$ , so that we have a very similar behaviour to that of the short lines of sight parallel to  $Y$ . Secondly, lines of sight from **1** to **2** pass through many less dense  $\text{CO}$  structures, with a lot of molecular hydrogen in between, leading to the sharp drop in the  $N(\text{CO})$  vs.  $N(\text{H}_2)$  relation. At a given  $N(\text{CO})$ , the difference  $\Delta N(\text{H}_2)$  between  $\text{H}_2$  column densities in the branches **0–1** and **1–2** thus represents the “dark gas” (see Sect. 5.4). Finally, lines of sight from **2** to **3** barely have any  $\text{CO}$  and contain less and less  $\text{H}_2$  as we approach position **3**, where we return to a situation similar to that at position **0**. The split between the two branches is definitely related to the mean gas density on the line of sight, the upper branch having  $\langle n_{\text{H}} \rangle \sim 250 \text{ cm}^{-3}$ , the lower one having  $\langle n_{\text{H}} \rangle \sim 150 \text{ cm}^{-3}$ . The overall deficit in  $\text{CO}$  suggests that we may not shield it sufficiently from the ambient UV field, a hypothesis that we discuss in Sect. 6.

The behaviour of  $N(\text{CH})$  with respect to  $N(\text{H}_2)$  (Fig. 11 – middle panel) agrees remarkably well with the observational data fitted by Sheffer et al. (2008). There is a discrepancy for  $N(\text{H}_2) \lesssim 10^{20} \text{ cm}^{-2}$ , but it is irrelevant, because there are no detections in that range, only upper limits.

We typically have a factor 10 deficit for  $\text{CN}$  (Fig. 11 – bottom panel) compared to observations. It appears that the reaction rate coefficient for the  $\text{CN} + \text{N} \rightarrow \text{C} + \text{N}_2$  reaction in the KIDA database<sup>2</sup> may be too high (E. Roueff, priv. comm.), which could partly explain that deficit.

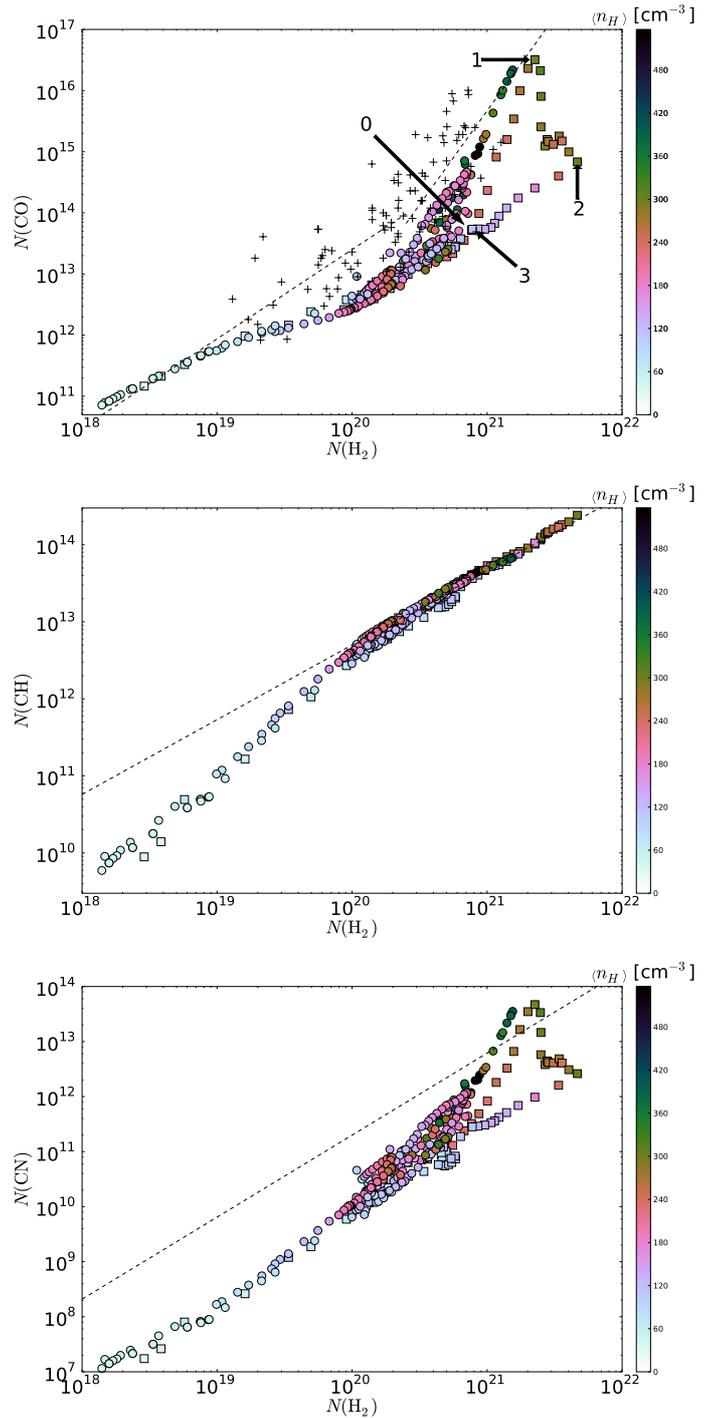
<sup>2</sup> <http://kida.obs.u-bordeaux1.fr/>



**Fig. 10.** Abundance ratios  $n(\text{CH})/n(\text{H}_2)$  (top),  $n(\text{CN})/n(\text{CO})$  (middle), and  $n(\text{CS})/n(\text{C})$  (bottom) in the vicinity of the total gas density peak, located at  $(X \approx -4.7 \text{ pc}, Y \approx -24 \text{ pc})$ . Contours mark total gas densities  $n_{\text{H}}$  of 100, 200, 300, 400, 500, 1000, and 2000  $\text{cm}^{-3}$ .

### 5.3. Comparison of the los and uniform models

To assess the effects of taking into account density fluctuations, as opposed to the assumption of a homogeneous medium, which is usually made when modelling PDRs, we computed the column densities of CO and  $\text{H}_2$  derived from the uniform models, and plot them in Fig. 12. The behaviour at low  $N(\text{H}_2)$  is very similar to that in the los models (Fig. 11), and the data suffer from the same deficit in CO around  $N(\text{H}_2) \sim 10^{20} \text{ cm}^{-2}$ . However, there is a definite difference at higher  $\text{H}_2$  column densities: the slope of the relation between both column densities is much steeper,  $d[\log N(\text{CO})]/d[\log N(\text{H}_2)] \approx 14$ , than in the los models and in observational data fits. This break occurs later, at  $N(\text{H}_2) \sim 5 \times 10^{20} \text{ cm}^{-2}$ , and applies to the short lines of sight



**Fig. 11.** Column densities of CO (top), CH (middle), and CN (bottom) versus column densities of  $\text{H}_2$ , in the los models. Circles correspond to lines of sight parallel to  $Y$  and squares to lines of sight parallel to  $X$ . Their colours reflect the mean gas density  $\langle n_{\text{H}} \rangle$  on these lines of sight. Plus signs on the top (CO) plot stand for observational data points (Sheffer et al. 2008; Crenny & Federman 2004; Pan et al. 2005; Lacour et al. 2005; Rachford et al. 2002, 2009; Snow et al. 2008). The dashed lines are power-law fits from Sheffer et al. (2008). The lines of sight parallel to  $X$  marked 0, 1, 2, and 3 on the top panel are the same as in Figs. 7 and 8.

(parallel to  $Y$  and parallel to  $X$  between positions 0 and 1) that show essentially one structure in both  $\text{H}_2$  and CO. However, the maximum column densities reached are slightly lower than in the los models by a factor  $\sim 3$ , for both CO and  $\text{H}_2$ . This shows

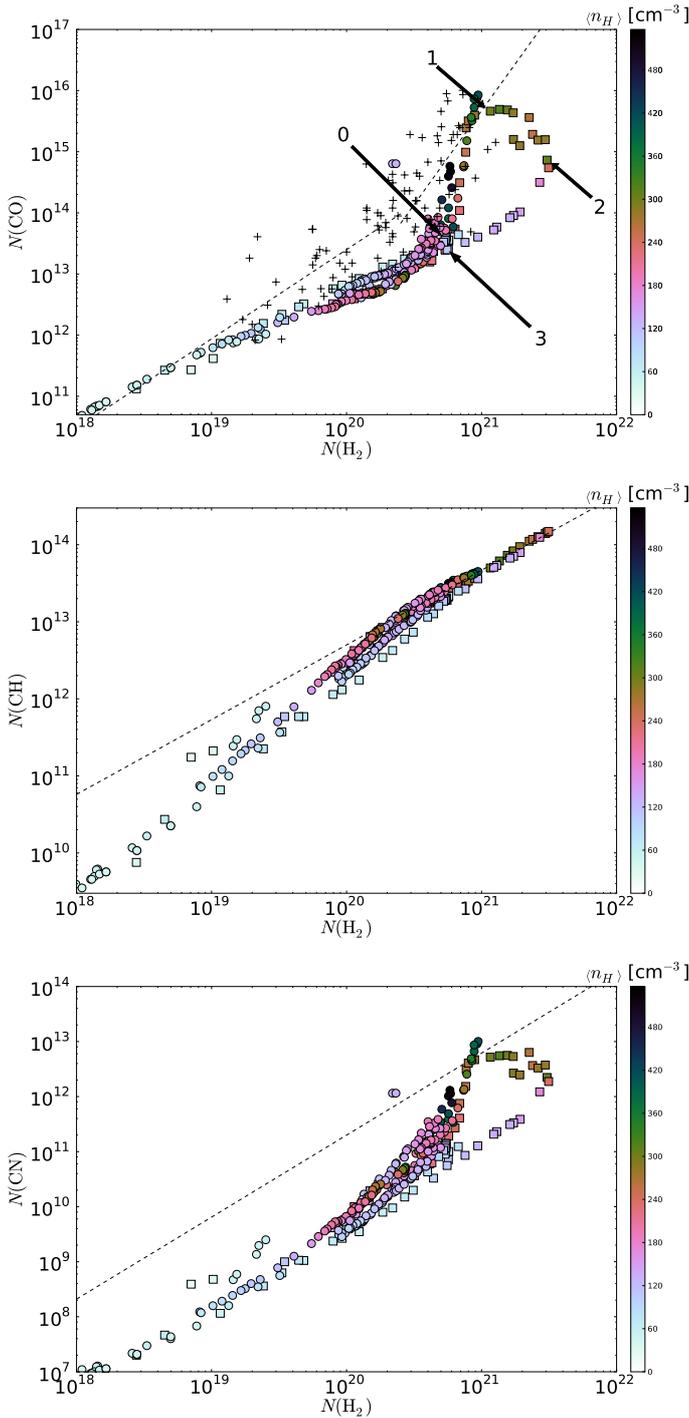


Fig. 12. Same as Fig. 11 but for the uniform models.

the importance of taking into account density fluctuations along the line of sight when modelling PDRs.

For CH, the behaviour is very similar in the uniform and los models, although here also the maximum column densities reached are slightly lower in the uniform models, by a factor  $\sim 2$ . The observational fit is recovered at somewhat higher  $H_2$  column densities ( $2 \times 10^{20} \text{ cm}^{-2}$  instead of  $10^{20} \text{ cm}^{-2}$ ), and the scatter of data points is slightly larger.

Finally, regarding CN, if we ignore the deficit already seen in the los models, we reach the same conclusions: In the uniform models, a break in the slope occurs at higher  $H_2$  column densities  $N(H_2) \sim 6 \times 10^{20} \text{ cm}^{-2}$  (instead of  $N(H_2) \sim 4 \times 10^{20} \text{ cm}^{-2}$  in the

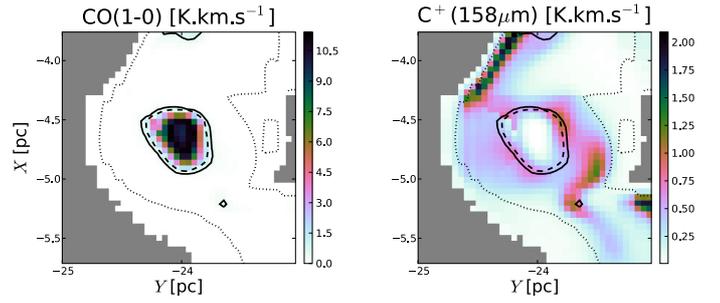


Fig. 13. Synthetic emission maps in  $CO(J = 1 \rightarrow 0)$  (left) and  $[CII]$  at  $158 \mu\text{m}$  (right). The solid contour marks the assumed  $0.4 \text{ K km s}^{-1}$  detection threshold for CO, the dashed contour marks the position of the line centre optical depth  $\tau_{CO} = 1$ , and the dotted contour marks the position of the molecular transition  $f_{H_2} = 1/2$ . Grey areas are outside of the computational domain.

los models), the slope is definitely steeper in that high-column-density regime, and the maximum column densities reached are lower, here by a factor  $\sim 4$ .

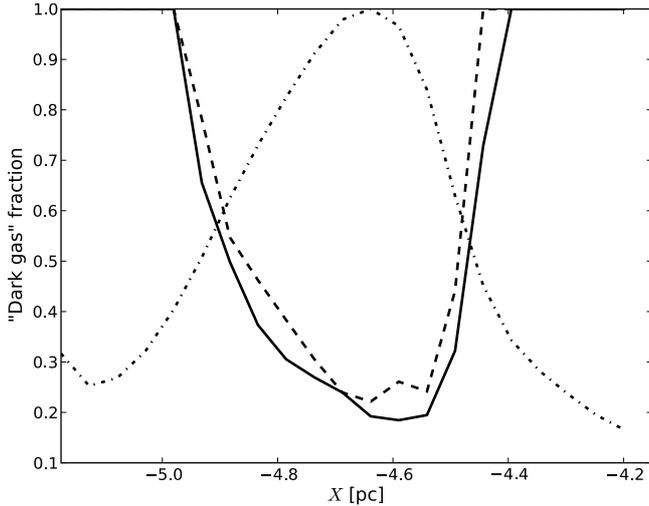
#### 5.4. Dark molecular gas fraction

To estimate the amount of molecular gas not seen in CO in our simulation – the so-called “dark gas” (Grenier et al. 2005; Planck Collaboration 2011) – we need to compute the CO line emission and compare its spatial distribution with that of  $H_2$ . To that effect, we focused on the cloulet at  $(X \simeq -4.7 \text{ pc}, Y \simeq -24 \text{ pc})$ , where the gas density peak is found, and assumed a very crude cylindrical cloud model by replicating the density maps along the  $Z$  axis over a line-of-sight  $L = 1 \text{ pc}$ , which roughly corresponds to the cloulet’s extent in the  $(X, Y)$  plane. This effectively yields column density maps that we can use with the RADEX radiative transfer code (Van der Tak et al. 2007) to obtain emission maps in the  $CO(J = 1 \rightarrow 0)$  rotational transition line at  $115.271 \text{ GHz}$  and in the  $[CII]$  fine structure transition line at  $158 \mu\text{m}$ . To be precise, at each position  $(X, Y)$ , we treated radiative transfer along  $Z$  in a plane-parallel slab geometry. RADEX works with the escape probability formalism (Sobolev 1960), which requires specification of the line width  $\Delta V$ . We estimate it to be  $\sigma_{3D}/\sqrt{3} = 1 \text{ km s}^{-1}$ , where  $\sigma_{3D} = 1.8 \text{ km s}^{-1}$  is the total gas velocity dispersion listed in Table 1. Indeed, for any species  $\alpha$  with molecular weight  $\mu_\alpha$  ( $\mu_{CO} = 28$  and  $\mu_{C^+} = 12$ ), the ratio of thermal to one-dimensional turbulent velocity dispersions reads

$$3 \frac{\sigma_{\text{th}}^2(\alpha)}{\sigma_{3D}^2} \simeq 2.1 \times \left( \frac{T}{270 \text{ K}} \right) \left( \frac{1}{\mu_\alpha} \right).$$

In the cloulet under study,  $T \lesssim 100 \text{ K}$ , so the above ratio is typically  $\lesssim 0.03$  for CO and  $\lesssim 0.06$  for  $C^+$ . It is thus reasonable to take  $\Delta V = \sigma_{3D}/\sqrt{3}$  for all RADEX runs. The code also requires specification of the gas kinetic temperature and the densities of collisional partners ( $H_2$  for CO;  $H_2$ , H and electrons for  $[CII]$ ), which we obtained from the PDR code outputs. RADEX is thus run on every line of sight parallel to the  $Z$  axis, and results are combined into a  $CO(J = 1 \rightarrow 0)$  emission map and a  $[CII]$  emission map, both shown in Fig. 13.

Between the solid and dotted contours is the “dark molecular gas” region where hydrogen is predominantly in its molecular form but CO emission fails to detect it. We assumed a detection threshold  $W_{CO} = 0.4 \text{ K km s}^{-1}$  consistent with the noise level in e.g. the CO survey of Taurus by Goldsmith et al. (2008).



**Fig. 14.** Fraction of “dark gas” (solid line) along one-dimensional cuts parallel to the  $Y$  axis going through the cloulet at ( $X \simeq -4.7$  pc,  $Y \simeq -24$  pc). Also shown are the fraction of “dark gas” computed using the definition by [Wolfire et al. \(2010\)](#) (dashed line), and a normalised profile of the total gas column densities  $N_{\text{H}}$  along the same cuts (dash-dotted line).

On the other hand, that same gas can definitely be traced in the [CII] line, because it has a typical integrated emission of  $\sim 0.4\text{--}0.8$  K km s $^{-1}$ , while the sensitivity quoted by [Velusamy et al. \(2010\)](#) for the GOTC+ key program is  $\sim 0.1\text{--}0.2$  K km s $^{-1}$ .

The “dark molecular gas” fraction associated with this cloulet can be estimated by taking one-dimensional cuts parallel to the  $Y$  axis going through the CO emission region. Along such a cut, which is parameterised by  $X$ , we defined  $Y_0(X)$  and  $Y_1(X)$  as the boundaries of the computational domain<sup>3</sup> (sharp transition from grey to white in the panels of [Fig. 13](#)), and we denote with  $Y_{\text{CO}}(X)$  the region where the integrated emission of CO ( $J = 1 \rightarrow 0$ ) exceeds the detection threshold  $W_{\text{CO}}$  (region enclosed by the solid contour in [Fig. 13](#)). We then define the “dark gas” fraction as

$$f_{\text{DG}}(X) = 1 - \frac{\int_{Y_{\text{CO}}(X)} n(\text{H}_2) dY}{\int_{Y_0(X)} n(\text{H}_2) dY}.$$

[Figure 14](#) shows this fraction as a function of the position  $X$  of the one-dimensional cut. Obviously,  $f_{\text{DG}} = 1$  when the cut does not pass through the CO emission region, and  $f_{\text{DG}} < 1$  when some of the  $\text{H}_2$  is traced by CO. We found that in this cloulet, at least 20% of  $\text{H}_2$  is not traced by CO, even at the peak of the gas density. To obtain a mean fraction of dark gas in this cloulet, we averaged  $f_{\text{DG}}$  over the range of  $X$  coordinates at which CO is seen (i.e.  $f_{\text{DG}}(X) < 1$ ), weighted by the total gas column density  $N_{\text{H}}$ . This yields  $\overline{f_{\text{DG}}} = 0.32$ , which is somewhat higher than the findings of [Velusamy et al. \(2010\)](#), who identified 53 “transition clouds” with H I and  $^{12}\text{CO}$  emission but no  $^{13}\text{CO}$ , and found that  $\sim 25\%$  of  $\text{H}_2$  in these clouds belong to an  $\text{H}_2/\text{C}^+$  layer not seen in CO. However, the scatter in observational values is large ([Grenier et al. 2005](#); [Abdo et al. 2010](#)), so the small discrepancy is no cause for alarm. Another possible

<sup>3</sup> The dependence of  $f_{\text{DG}}$  on the boundaries of the computational domain is necessarily small, because there is little mass at low densities.

comparison is with [Wolfire et al. \(2010\)](#), who constructed spherical models of molecular clouds to study the dark gas fraction, which they defined in a similar way, except that the boundary of their CO region is specified by the condition of unit optical thickness at the line centre,  $\tau_{\text{CO}} = 1$ . In our clump, that isocontour is very close to our own condition  $I_{\text{CO}} = W_{\text{CO}}$ , as can be seen in [Fig. 13](#). Computing the average dark gas fraction with [Wolfire et al. \(2010\)](#)’s condition yields  $\overline{f_{\text{DG}}} = 0.36$ , which is quite close to their results  $f_{\text{DG}} \sim 0.25\text{--}0.33$ . There is a notable difference between their models and ours, however, since our cloulet has a mass  $\sim 9.5 M_{\odot}$  inside the CO region, while [Wolfire et al. \(2010\)](#) studied GMCs with masses in the range  $10^5 M_{\odot}$  to  $3 \times 10^6 M_{\odot}$ . Their impinging UV field was also notably higher ( $\chi = 3\text{--}30$ ).

## 6. Discussion and summary

### 6.1. Illumination effects

We bypassed the one-dimensionality of the PDR code by combining runs in two orthogonal directions, taking at each grid point the chemical composition corresponding to maximum radiation energy density  $E$ . However, this choice is questionable: if a position is shielded from radiation in almost every direction but for one small hole, the illumination at this location resulting from our procedure is too high. As this means forming less molecules, we estimated if it might account for some of the CO deficit seen in [Fig. 11](#). To do so in a simple way, we computed the chemical composition under the opposite assumption, i.e. based on the criterion of minimum local illumination. The result is plotted in [Fig. 15](#), and shows how this indeed helps recovering observed CO column densities for  $N(\text{H}_2) \gtrsim 10^{20}$  cm $^{-2}$ , with a consistent scatter. Below  $N(\text{H}_2) \simeq 10^{20}$  cm $^{-2}$ , a significant CO deficit remains, however.

Obviously, this choice of minimum local illumination is also unphysical, and the reality must lie somewhere in between. A physically better, but more computationally intensive method is being pursued and will be presented in a future paper.

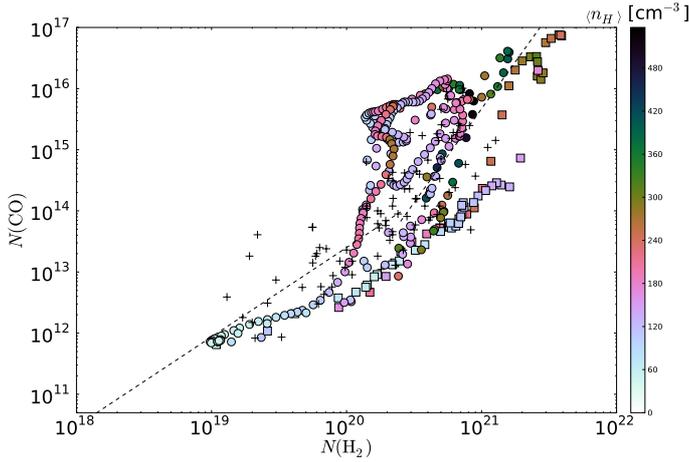
### 6.2. FGK approximation

Our study used the [Federman et al. \(1979, FGK\)](#) approximation to compute self-shielding. This may underestimate the shielding of CO by molecular hydrogen lines, therefore we performed a few runs of the PDR code using exact radiative transfer ([Goicoechea & Le Bourlot 2007](#)). We did this on some lines of sight for which  $N(\text{H}_2) \sim 10^{20}$  cm $^{-2}$ , to see whether this helps fill the CO deficit in that region. It turns out that the CO column densities so obtained are indeed higher than those found in the FGK approximation, but by a factor  $\lesssim 2$ , which is not sufficient to explain our CO deficit. Because this increased computational time by a factor  $\sim 5\text{--}6$ , we feel that this approach is not to be recommended.

### 6.3. Steady-state assumption

We took the simulation cube to be a static background, under the assumption that timescales for chemistry and photoprocesses are much shorter than those of the MHD simulation.

From the analysis performed by [Le Petit et al. \(2006\)](#) on uniform density PDRs, it appears that timescales for  $\text{H}_2$  photodissociation at the edges of a cloud are  $\sim 1000/\chi$  yr, where  $\chi$  is the FUV radiation strength in units of the [Draine \(1978\)](#) field. In our analysis, we chose  $\chi = 1$  so that the corresponding timescale is about 1000 yr.



**Fig. 15.** CO column densities versus H<sub>2</sub> column densities in the los models when combining data based on a criterion of minimum local illumination. Symbols are the same as in Fig. 11, top panel.

Estimating timescales for the MHD simulation is more difficult, because what we are interested in is the time over which structures remain coherent, and we do not have access to this information due to the Eulerian nature of the simulation, which makes it impossible to confidently identify structures. For a rough estimate, we may consider the overall crossing time  $\tau_{\text{cross}} = L/V_X \simeq 2.4$  Myr, but this does not correspond to the time over which gas is mixed by turbulence at a given scale. For this, we may use the velocity dispersion  $\sigma_{3D}$  listed in Table 1 within the observational clump, whose size is about 2 pc. This yields a dynamical timescale  $\tau_{\text{dyn}} \simeq 1.1$  Myr, which is very similar to the values quoted by Wolfire et al. (2010) to validate the steady-state assumption in their models.

The chemical timescale is that of the formation of molecular hydrogen. As shown by Glover & Mac Low (2007) using numerical simulations of decaying ISM turbulence that include a simplified chemical network, the formation timescale for H<sub>2</sub> in turbulent magnetised molecular clouds is  $\tau_{\text{chem}} \sim 1\text{--}2$  Myr. It is therefore of the order of the estimated dynamical timescales in our simulation, so that our steady-state assumption seems only marginally valid. However, H<sub>2</sub> is formed in dense regions and transported throughout the entire volume via turbulent motions, so we may safely assume steady state, provided we consider a sufficiently late snapshot. Indeed, if H<sub>2</sub> starts forming when the converging WNM flows collide near the midplane ( $\tau_{\text{coll}} \simeq 1$  Myr), and if it is fully formed and transported throughout the entire volume after  $\tau_{\text{chem}} + \tau_{\text{dyn}} + \tau_{\text{cross}}$ , this requires taking a snapshot timed at no earlier than 5.4 Myr, which is the case here ( $t = 7.35$  Myr). We conclude that our steady-state assumption is a legitimate one.

#### 6.4. Warm chemistry

We note that chemistry is here driven by UV radiation only, but that there is an important pathway for the formation of many molecular species, which is warm chemistry in turbulence dissipation regions (TDR), studied by Joulain et al. (1998) and Godard et al. (2009). In particular, the CO abundances found in the models by Godard et al. (2009) are higher than in corresponding PDR models, sometimes by almost an order of magnitude. More generally, Godard et al. (2009) argued that observed chemical abundances are on the whole well reproduced if dissipation is caused by ion-neutral friction in sheared structures

$\sim 100$  AU thick. A TDR post-processing of our MHD simulation is in preparation, to compare both types of chemistry.

#### 6.5. Summary

We have presented a first analysis of UV-driven chemistry in a simulation of the diffuse ISM, by post-processing it with the Meudon PDR code. Our results show that assuming a uniform density medium when modelling PDRs leads to significant errors: for instance, the maximum CO column densities found with this simplistic assumption are a factor  $\sim 3$  lower than those found using actual density fluctuations. The slope of the H<sub>2</sub>-CO correlation at  $N(\text{H}_2) \gtrsim 5 \times 10^{20} \text{ cm}^{-2}$  is also a factor  $\sim 3$  higher than in the more realistic case, and therefore agrees less well with observations. A second result of our study is that in the densest parts of the simulation ( $n_{\text{H}} \gtrsim 10^3 \text{ cm}^{-3}$ ), some 35% to 40% of the molecular gas is “dark”, in the sense that it is not traced by the CO( $J = 1 \rightarrow 0$ ) line, given current sensitivities. It is detectable via the [CII] fine structure transition line at  $158 \mu\text{m}$ , however. As a side result, we found that the simplified cooling used in the MHD simulation by Hennebelle et al. (2008) yields gas temperatures that agree reasonably well with those found using the more detailed processes included in the PDR code.

*Acknowledgements.* The authors acknowledge support from ANR through the SCHISM project (ANR-09-BLAN-231), and computing resources and services from France Grilles and the EGI e-infrastructure. Some kinetic data have been downloaded from the online KIDA (Kinetic Database for Astrochemistry, <http://kida.obs.u-bordeaux1.fr>) database. Colour figures in this paper use the cubehelix colour map by Green (2011).

#### Appendix A: Using the Meudon PDR code with density profiles

This Appendix is meant as an introduction to using the Meudon PDR code<sup>4</sup> with fluctuating density profiles. For more detailed presentations of the code, the reader is referred to Le Bourlot et al. (1999), Le Petit et al. (2006), and Gonzalez Garcia et al. (2008).

The code requires two input files: a .pfl file listing visual extinction  $A_V$ , temperature  $T$  (in K) and total gas density  $n_{\text{H}}$  (in  $\text{cm}^{-3}$ ) along the line of sight, and a .in file supplying the parameters of the run to perform. These are listed in Table A.1, and some of them require a short comment:

- `modele` is the basename chosen for output files. For the los models, we used a name of the generic form `los_zZ_yY_xX_m-X_p` or `los_zZ_xX_yY_m-Y_p`, reflecting the position of the extracted profile in the cube. Note that  $Z$  is a constant throughout this paper, as is evident from Fig. 2. For the uniform models, we used names of the generic form `uniform_zZ_yY_xX_m-X_p` or `uniform_zZ_xX_yY_m-Y_p`.
- `ifafm` is the number of global iterations to use. As Le Petit et al. (2006) pointed out, for diffuse clouds ( $A_V < 0.5$ ) proper convergence may require up to 20 iterations, therefore we selected `ifafm=20` for all our models. For information, 415 of the 447 profiles have total  $A_V < 0.5$ .
- `Avmax` is that same total visual extinction through the cloud, which is simply the last  $A_V$  value in the .pfl file.
- We set the density `densh`, temperature `tgaz`, and pressure `presse=densh*tgaz` parameters to the average values<sup>5</sup> for each profile. They are not used by the code with a

<sup>4</sup> We used version 1.4.1 of the PDR code, with a fixed H<sub>2</sub> formation rate  $R_f = 3 \times 10^{-17} \sqrt{T/100 \text{ K}} \text{ cm}^3 \text{ s}^{-1}$ .

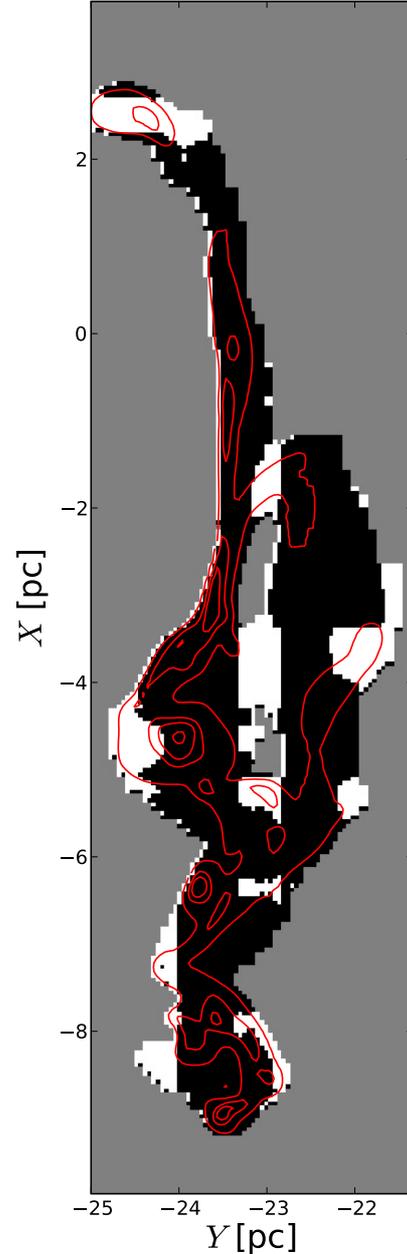
<sup>5</sup> I.e. density-weighted averages for temperature and pressure.

**Table A.1.** Parameters used in the PDR code as input in the .in files.

Parameter	Description	Value
modele	Basename for the output files	see Appendix A
chimie	Chemistry file	chimie08 <sup>a</sup>
ifafm	Number of global iterations	20
Avmax	Integration limit in $A_V$	see Appendix A
densh	Initial density ( $\text{cm}^{-3}$ )	see Appendix A
F_ISRF	ISRF expression flag	1 <sup>b</sup>
radm	ISRF scaling factor $\chi_m$	see Appendix A
radp	ISRF scaling factor $\chi_p$	see Appendix A
srcpp	Additional radiation field source	none.txt
d_sour	Star distance (pc)	0 <sup>c</sup>
fmr	Cosmic rays ionisation rate ( $10^{-17} \text{ s}^{-1}$ )	5
ieqth	Thermal balance computation flag	1 <sup>d</sup>
tgaz	Initial temperature (K)	see Appendix A
ifisob	State equation flag	see Appendix A
fprofil	Density-Temperature profile file	see Appendix A
presse	Initial pressure ( $\text{K cm}^{-3}$ )	see Appendix A
vturb	Turbulent velocity ( $\text{km s}^{-1}$ )	see Appendix A
itrfer	UV transfer method flag	0 <sup>e</sup>
jfgkh2	Minimum $J$ level for FGK approximation	0
ichh2	H + H <sub>2</sub> collision rate model flag	2 <sup>f</sup>
los_ext	Line of sight extinction curve	Galaxy <sup>g</sup>
rrr	Reddening coefficient $R_V = A_V/E_{B-V}$	3.1 <sup>h</sup>
cdunit	Gas-to-dust ratio $C_D = N_H/E_{B-V}$ ( $\text{cm}^{-2}$ )	$5.8 \times 10^{21}$ <sup>g</sup>
alb	Dust albedo	0.42 <sup>g</sup>
gg	Diffusion anisotropy factor $\langle \cos \theta \rangle$	0.6 <sup>g</sup>
gratio	Mass ratio of grains/gas	0.01 <sup>g</sup>
rhogr	Grains mass density ( $\text{g cm}^{-3}$ )	2.59 <sup>i</sup>
alpgr	Grains distribution index	3.5 <sup>g</sup>
rgrmin	Grains minimum radius (cm)	$3 \times 10^{-7}$ <sup>g</sup>
rgrmax	Grains maximum radius (cm)	$3 \times 10^{-5}$ <sup>g</sup>
F_DUSTEM	DUSTEM activation flag	0 <sup>j</sup>
iforh2	H <sub>2</sub> formation on grains model flag	0 <sup>k</sup>
istic	H sticking on grain model flag	4 <sup>l</sup>

**Notes.** <sup>(a)</sup> This chemistry file does not include deuterated species. <sup>(b)</sup> Uses expressions based on Mathis et al. (1983) and Black (1994) rather than Draine (1978). <sup>(c)</sup> This means that no additional star is present. <sup>(d)</sup> Thermal balance is computed for each point in the cloud, using the temperatures given by the .pfl file as initial guesses. <sup>(e)</sup> Use the Federman et al. (1979, FGK) approximation for the H<sub>2</sub> lines in the UV. <sup>(f)</sup> Use the values compiled by Le Bourlot et al. (1999) with reactive collisions from Schofield (1967). <sup>(g)</sup> See Table 4 of Le Petit et al. (2006), which quotes values from Fitzpatrick & Massa (1990) for the extinction curve, Bohlin (1978); Rachford et al. (2002) for  $C_D$ , Mathis (1996) for the dust albedo, Weingartner & Draine (2001) for  $\langle \cos \theta \rangle$ , Mathis et al. (1977) for the grain size distribution parameters. <sup>(h)</sup> See for instance Cardelli et al. (1989). <sup>(i)</sup> Taken from Gonzalez Garcia et al. (2008). <sup>(j)</sup> Do not couple to the DUSTEM code (Compiègne et al. 2011). <sup>(k)</sup> Energy released by H<sub>2</sub> formation on dust grains is equally split between grain excitation, H<sub>2</sub> kinetic energy and internal energy. See Sect. 6.1.2 in Le Petit et al. (2006) for details. <sup>(l)</sup> See Appendix E5 in Le Petit et al. (2006).

## Illumination mask



**Fig. A.1.** Illumination mask computed by comparing at each point ( $X, Y$ ) the local radiative energy densities  $E_X$  and  $E_Y$  output by the PDR code along the  $X$  and  $Y$  directions, respectively. Regions where  $E_Y \geq E_X$  are marked in black and regions where  $E_X > E_Y$  are marked in white. Contour lines of equal total gas density are overlaid at 20, 100, 500, 1000, and 2000  $\text{cm}^{-3}$ . Grey areas are outside of the computational domain.

density-temperature profile, but they are used, in the reference uniform models, as initial guesses for thermal balance computation.

- radm and radp specify the strengths  $\chi_m$  and  $\chi_p$  of the incident radiation field in units of the ISRF on the left and right sides of the profile, respectively. For the runs described in this paper, we used  $\chi_m = \chi_p = 1$ .
- fprofil specifies the .pfl density-temperature profile file. For consistency, we used the same naming scheme as for modele.

- vturb is the “turbulent velocity dispersion”. It does not include thermal dispersion, therefore we took it to be the standard deviation of the line-of-sight velocity within each structure, noted  $\sigma_X$  and  $\sigma_Y$  in Tables 2 and 3, respectively.
- ifisob is a flag specifying whether to use a density profile. For the los models, we therefore set ifisob=1, to enforce the use of a density-temperature .pfl file. However, since thermal balance is solved (ieqth=1), the temperature values in the file are only used as initial guesses. For the uniform models, we set ifisob=0 to use a constant density (specified by densh).

## Appendix B: Post-processing of raw outputs

### B.1. Resampling

Outputs of the PDR code are FITS data files and XML description files, and we used dedicated scripts to extract specific quantities into plain text files for the subsequent analysis. Among the quantities retrieved are the distance  $d$  from the surface of the structure, visual extinction  $A_V$ , proton column density  $N_H$ , temperature  $T_{\text{PDR}}$ , proton density  $n_H$ , pressure  $p$ , ionization fraction  $x_e$ , and abundances  $n(\alpha)$  of 99 chemical species.

Because the PDR code produces its own mesh refinement to better solve for the H/H<sub>2</sub> transition, these quantities are sampled irregularly. Consequently, we resampled outputs on the same regular grid as the MHD simulation, using a simple linear interpolation method. This allowed us to build raw maps for all output quantities from PDR code runs along the  $X$  and  $Y$  directions.

### B.2. Missing data

For reasons that are unclear, a few<sup>6</sup> of the 894 PDR runs did not complete successfully on the EGEE grid. To supplement the missing data, we interpolated along the perpendicular direction. Consider the ensemble of runs along the  $X$  direction: completed runs yield quantities  $F_X(X, Y)$ , and if a run is missing at coordinate  $Y = Y_0$ , we supplied  $F_X(X, Y_0)$  by linearly interpolating  $G(Y) = F_X(X_0, Y)$  at constant  $X = X_0$ . This yields a satisfactory completion of the raw data.

### B.3. Combination of $X$ and $Y$ runs

We then proceed to the combination of data from runs along the  $X$  and  $Y$  directions, as described in 3. Figure A.1 shows the “illumination mask” computed by comparing the local radiation energy densities  $E_X$  and  $E_Y$  output by the PDR code in 1 $\sigma$  models parallel to the  $X$  and  $Y$  directions, respectively. This mask is then used to build a single data array for each quantity  $F$  output by the PDR code at each grid point  $(X, Y)$ , according to the rule:

$$F = \begin{cases} F_X & \text{if } E_X > E_Y \\ F_Y & \text{if } E_X \leq E_Y. \end{cases}$$

This helps to reduce the shadowing artefacts due to the one-dimensionality of the PDR code, while ensuring element conservation in each grid cell, and yields the final maps that are analysed and discussed in the main body of the paper. In the discussion (Sect. 6), we also used the inverse choice:

$$F = \begin{cases} F_Y & \text{if } E_X > E_Y \\ F_X & \text{if } E_X \leq E_Y. \end{cases}$$

## References

Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, *ApJ*, 710, 133  
Banerjee, R., Vázquez-Semadeni, E., Hennebelle, P., & Klessen, R. S. 2009, *MNRAS*, 398, 1082

Black, J. H. 1994, in *The First Symposium on the Infrared Cirrus and Diffuse Interstellar Clouds*, ASP Conf. Ser., 58, 355  
Bohlin, R. C., Savage, B. D., & Drake, J. F. 1978, *ApJ*, 224, 132  
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, 345, 245  
Compiègne, M., Verstraete, L., Jones, A., et al. 2011, *A&A*, 525, A103  
Crenny, T., & Federman, S. R. 2004, *ApJ*, 605, 278  
Crutcher, R. M., Wandelt, B., Heiles, C., Falgarone, E., & Troland, T. H. 2010, *ApJ*, 725, 466  
de Graauw, T., Helmich, F. P., Phillips, T. G., et al. 2010, *A&A*, 518, L6  
Draine, B. 1978, *ApJS*, 36, 595  
Sternberg, A., & Dalgarno, A. 1995, *ApJS*, 99, 565  
Federman, S. R., Glassgold, A. E., & Kwan, J. 1979, *ApJ*, 227, 466  
Field, G. B. 1965, *ApJ*, 142, 531  
Fitzpatrick, E. L., & Massa, D. 1990, *ApJS*, 72, 163  
Fromang, S., Hennebelle, P., & Teyssier, R. 2006, *A&A*, 457, 371  
Glover, S. C. O., & Mac-Low, M.-M. 2007, *ApJ*, 659, 1337  
Glover, S. C. O., Federrath, C., Mac-Low, M.-M., & Klessen, R. 2010, *MNRAS*, 404, 2  
Godard, B., Falgarone, E., & Pineau des Forêts, G. 2009, *A&A*, 495, 847  
Goicoechea, J. R., & Le Bourlot, J. 2007, *A&A*, 467, 1  
Goldsmith, P. F., Heyer, M., Narayanan, G., et al. 2008, *ApJ*, 680, 428  
Gonzalez Garcia, M., Le Bourlot, J., Le Petit, F., & Roueff, E. 2008, *A&A*, 485, 127  
Green, D. A. 2011, *BASI*, 39, 289  
Grenier, I. A., Casandjian, J. M., & Terrier, R. 2005, *Science*, 307, 1292  
Heitsch, F., Burkert, A., Hartmann, L. W., Slyz, A. D., & Devriendt, J. E. G. 2005, *ApJ*, 633, L113  
Hennebelle, P., & Audit, E. 2007, *A&A*, 465, 431  
Hennebelle, P., & Pérault, M. 1999, *A&A*, 351, 309  
Hennebelle, P., & Pérault, M. 2000, *A&A*, 359, 1024  
Hennebelle, P., Banerjee, R., Vázquez-Semadeni, E., Klessen, R.S., & Audit, E. 2008, *A&A*, 486, L43  
Hollenbach, D. J., & Tielens, A. G. G. M. 1999, *Rev. Mod. Phys.*, 71, 173  
Joulain, K., Falgarone, E., Pineau des Forêts, G., & Flower, D. 1998, *A&A*, 340, 241  
Koyama, H., & Inutsuka, S.-I. 2002, *ApJ*, 564, L97  
Lacour, S., Ziskin, V., Hébrard, G., et al. 2005, *ApJ*, 627, 251  
Le Bourlot, J., Pineau des Forêts, G., Roueff, E., & Flower, D. R. 1993, *A&A*, 267, 233  
Le Bourlot, J., Pineau des Forêts, G., & Flower, D. R. 1999, *MNRAS*, 305, 802  
Le Petit, F., Nehmé, C., Le Bourlot, J., & Roueff, E. 2006, *ApJS*, 164, 506  
Leroy, A., Bolatto, A., Stanimirovic, S., et al. 2007, *ApJ*, 658, 1027  
Mathis, J. S. 1996, *A&A*, 472, 643  
Mathis, J. S., Rumpl, W., & Nordsieck, K. H. 1977, *ApJ*, 217, 425  
Mathis, J. S., Mezger, P. G., & Panagia, N. 1983, *A&A*, 128, 212  
Pan, K., Federman, S. R., Sheffer, Y., & Andersson, B.-G. 2005, *ApJ*, 633, 986  
Pety, J., Goicoechea, J. R., Gerin, M., et al. 2007, in *Proc. of the Molecules in Space and Laboratory Conf.*, eds. J.-L. Lemaire, & F. Combes (Publ.: S. Diana), id.13  
Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, *A&A*, 518, L1  
Planck Collaboration 2011, *A&A*, 536, A19  
Rachford, B. L., Snow, T. P., Tumlinson, J., et al. 2002, *ApJ*, 577, 221  
Rachford, B. L., Snow, T. P., Destree, J. D., et al. 2009, *ApJS*, 180, 125  
Schofield, K. 1967, *Planet. Space Sci.*, 15, 643  
Sheffer, Y., Rogers, M., Federman, S. R., et al. 2008, *ApJ*, 687, 1075  
Snow, T. P., Ross, T. L., Destree, J. D., et al. 2008, *ApJ*, 688, 1124  
Sobolev, V. V. 1960, *Moving Envelopes of Stars* (Harvard University Press)  
Sonnentrucker, P., Welty, D. E., Thorburn, J. A., & York, D. G. 2007, *ApJS*, 168, 58  
Teyssier, R. 2002, *A&A*, 385, 337  
Van der Tak, F. F. S., Black, J. H., Schier, F. L., Jansen, D. J., & van Dishoeck, E. F. 2007, *A&A*, 468, 627  
Vázquez-Semadeni, E., Gómez, G. C., Jappsen, A. K., et al. 2010, *ApJ*, 657, 870  
Velusamy, T., Langer, W. D., Pineda, J. L., et al. 2010, *A&A*, 521, L18  
Weingartner, J. C., & Draine, B. T. 2001, *ApJ*, 548, 296  
Wolfire, M. G., Hollenbach, D., & McKee, C. F. 2010, *ApJ*, 716, 1191

<sup>6</sup> Namely, for the 1 $\sigma$  models, 16 out of 291 along  $Y$  and 3 out of 156 along  $X$ ; for the uniform models, 8 out of 291 along  $Y$  and 6 out of 156 along  $X$ .