

LETTER TO THE EDITOR

# On the low ortho-to-para H<sub>2</sub> ratio in star-forming filaments

Alessandro Lupi<sup>1,2</sup> , Stefano Bovino<sup>3</sup>, and Tommaso Grassi<sup>4</sup>

<sup>1</sup> Dipartimento di Fisica “G. Occhialini”, Università degli Studi di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy  
e-mail: [alessandro.lupi@unimib.it](mailto:alessandro.lupi@unimib.it)

<sup>2</sup> INFN – Sezione di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy

<sup>3</sup> Departamento de Astronomía, Facultad Ciencias Físicas y Matemáticas, Universidad de Concepción,  
Av. Esteban Iturra s/n Barrio Universitario, Casilla 160, Concepción, Chile

<sup>4</sup> Center for Astrochemical Studies, Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstrasse 1, 85748 Garching,  
Germany

Received 3 September 2021 / Accepted 1 October 2021

## ABSTRACT

The formation of stars and planetary systems is a complex phenomenon that relies on the interplay of multiple physical processes. Nonetheless, it represents a crucial stage for our understanding of the Universe, and in particular of the conditions leading to the formation of key molecules (e.g. water) on comets and planets. *Herschel* observations demonstrated that stars form in gaseous filamentary structures in which the main constituent is molecular hydrogen (H<sub>2</sub>). Depending on its nuclear spin H<sub>2</sub> can be found in two forms: ‘ortho’ with parallel spins and ‘para’ where the spins are anti-parallel. The relative ratio among these isomers, the ortho-to-para ratio (OPR), plays a crucial role in a variety of processes related to the thermodynamics of star-forming gas and to the fundamental chemistry affecting the deuteration of water in molecular clouds, commonly used to determine the origin of water in Solar System bodies. Here, for the first time, we assess the evolution of the OPR starting from the warm neutral medium by means of state-of-the-art 3D magnetohydrodynamic simulations of turbulent molecular clouds. Our results show that star-forming clouds exhibit a low OPR ( $\ll 0.1$ ) already at moderate densities ( $\sim 1000 \text{ cm}^{-3}$ ). We also constrain the cosmic-ray ionisation rate, finding that  $10^{-16} \text{ s}^{-1}$  is the lower limit required to explain the observations of diffuse clouds. Our results represent a step forward in the understanding of the star and planet formation processes providing a robust determination of the chemical initial conditions for both theoretical and observational studies.

**Key words.** ISM: molecules – stars: formation – astrochemistry – magnetohydrodynamics (MHD) – methods: numerical

## 1. Introduction

Star formation is one of the greatest open problems in astrophysics (Bergin & Tafalla 2007; McKee & Ostriker 2007), and despite the huge progress made over the last decades, both observationally and theoretically, some fundamental questions remain open. Star formation occurs within molecular clouds (MCs), dense and cold regions within galaxies mainly composed of molecular hydrogen (H<sub>2</sub>). Within these clouds, stars form out of gaseous filamentary structures (André et al. 2010; Molinari et al. 2010) which show quasi-universal average properties (Arzoumanian et al. 2011). Due to the huge dynamic range involved and the variety of complex physical processes that couple together on different scales, studying star formation from ab initio conditions is still unfeasible. For this reason the problem has been tackled from different sides: on MC scales, with the aim of describing the formation of filaments and cores (e.g. Federrath et al. 2010a, 2016; Padoan & Nordlund 2011; Padoan et al. 2016), or on the small scales typical of these substructures, neglecting the large-scale environment and focussing on the last stages of the gravitational collapse (e.g. Bovino et al. 2019, 2020). Unfortunately, while small-scale simulations are extremely useful for studying the final stages of the gravitational collapse and comparing the results with observations of protostellar cores, the detailed chemical conditions

at the onset of gravitational collapse are still very uncertain, and can be constrained only via self-consistent studies on MC scales.

From a chemical point of view, the evolution of species like CO and H<sub>2</sub> up to the formation of filaments is crucial to assess the formation of key molecules (e.g. water) at later stages as on comets and planets (Hogerheijde et al. 2011; Bergin & van Dishoeck 2012; Ceccarelli et al. 2014; Jørgensen et al. 2020). This process strongly depends on the conditions of the main constituent of star-forming filaments, molecular hydrogen (H<sub>2</sub>), which can be found in two forms, the ortho and para states, where the spins are parallel or anti-parallel, respectively. The relative abundance of these isomers, called the ortho-to-para ratio (OPR), has profound implications for both thermodynamic and chemical processes that affect the formation and deuteration of water (Furuya et al. 2015; Jensen et al. 2021) and its importance in the Solar System (Altwegg et al. 2015).

In this work we assess for the first time the evolution of the OPR starting from large-scale conditions (i.e. the warm neutral medium) by means of state-of-the-art three-dimensional (3D) magnetohydrodynamic (MHD) simulations of turbulent molecular clouds. The Letter is organised as follows: in Sect. 2 we introduce the set-up of our simulations, in Sect. 3 we discuss our results, and in Sect. 4 we draw our conclusions.

## 2. Numerical set-up

The simulations presented in this work were performed with the publicly available MHD code GIZMO (Hopkins 2015, 2016; Hopkins & Raives 2016), a descendant of GADGET2 (Springel 2005), which included gas self-gravity<sup>1</sup>.

### 2.1. Microphysics

For the purpose of this study, we equipped the code with an on-the-fly non-equilibrium chemistry network, implemented via the public chemistry library KROME (Grassi et al. 2014). The chemical network we employed is based on Grassi et al. (2017), which is an updated version of that in Glover et al. (2010). We included isomer-dependent chemistry, by employing the most up-to-date reaction rates (Sipilä et al. 2015; Bovino et al. 2019). The final network includes 40 species: H, H<sup>+</sup>, He, He<sup>+</sup>, He<sup>++</sup>, ortho-H<sub>2</sub>, para-H<sub>2</sub>, ortho-H<sub>2</sub><sup>+</sup>, para-H<sub>2</sub><sup>+</sup>, H<sup>-</sup>, C<sup>+</sup>, C, O<sup>+</sup>, O, OH, HOC<sup>+</sup>, HCO<sup>+</sup>, CO, CH, CH<sub>2</sub>, C<sub>2</sub>, HCO, H<sub>2</sub>O, O<sub>2</sub>, ortho-H<sub>3</sub><sup>+</sup>, para-H<sub>3</sub><sup>+</sup>, CH<sup>+</sup>, CH<sub>2</sub><sup>+</sup>, CO<sup>+</sup>, CH<sub>3</sub><sup>+</sup>, OH<sup>+</sup>, H<sub>2</sub>O<sup>+</sup>, H<sub>3</sub>O<sup>+</sup>, O<sub>2</sub><sup>+</sup>, C<sup>-</sup>, O<sup>-</sup>, electrons, plus GRAIN0, GRAIN-, and GRAIN+, which represent dust grains. A total of 397 reactions connects all these species, including ortho-to-para conversion by proton collisions (H<sup>+</sup> and H<sub>3</sub><sup>+</sup>), adsorption and desorption of CO and water on the surface of grains (Cazaux et al. 2010; Hocuk et al. 2014), ionisation and dissociation induced by impact with cosmic rays, and dissociation of molecules induced by a standard interstellar radiation field (Draine flux, Draine 1978), which includes self-shielding of H<sub>2</sub> (Glover et al. 2010) and CO (Visser et al. 2009). Electron attachment and recombination of positive ions on grains were also included in the chemical network (Walmsley et al. 2004), with the Coulomb factor consistently calculated for the Draine flux (Draine & Sutin 1987), as well as H<sub>2</sub> formation on dust (assuming an initial OPR of 3 at formation; see Watanabe et al. 2010; Gavilan et al. 2012; Hama & Watanabe 2013; Wakelam et al. 2017) and dust cooling (Grassi et al. 2014), which were determined via pre-computed dust tables, in this case density-, temperature-, and A<sub>v</sub>-dependent. The extinction parameter is defined as A<sub>v</sub> = (10<sup>-3</sup>n<sub>H<sub>2</sub></sub>)<sup>α</sup>, with α = 2/3 (see Grassi et al. 2014), and the H<sub>2</sub> column density entering the H<sub>2</sub> self-shielding factor as N<sub>H<sub>2</sub></sub> = 1.87 × 10<sup>21</sup> A<sub>v</sub> cm<sup>-2</sup> (Grassi et al. 2014).

To consistently follow the thermodynamics of the gas, we further included metal line cooling from CI, CII, and OI; CO rotational cooling; chemical cooling induced by endothermic reactions; H<sub>2</sub> roto-vibrational cooling; Compton cooling; continuum cooling, plus chemical heating induced by exothermic reactions; photoheating; and cosmic-ray induced heating. Photoelectric heating was also included (Bakes & Tielens 1994), following recent modifications (Wolfire et al. 2003). A floor of 10 K was imposed. In order to model cosmic-ray attenuation through the cloud, in our fiducial model we employ a variable cosmic-ray flux which depends on local column density, as in Padovani et al. (2018) (see Appendix A for details).

### 2.2. Initial conditions

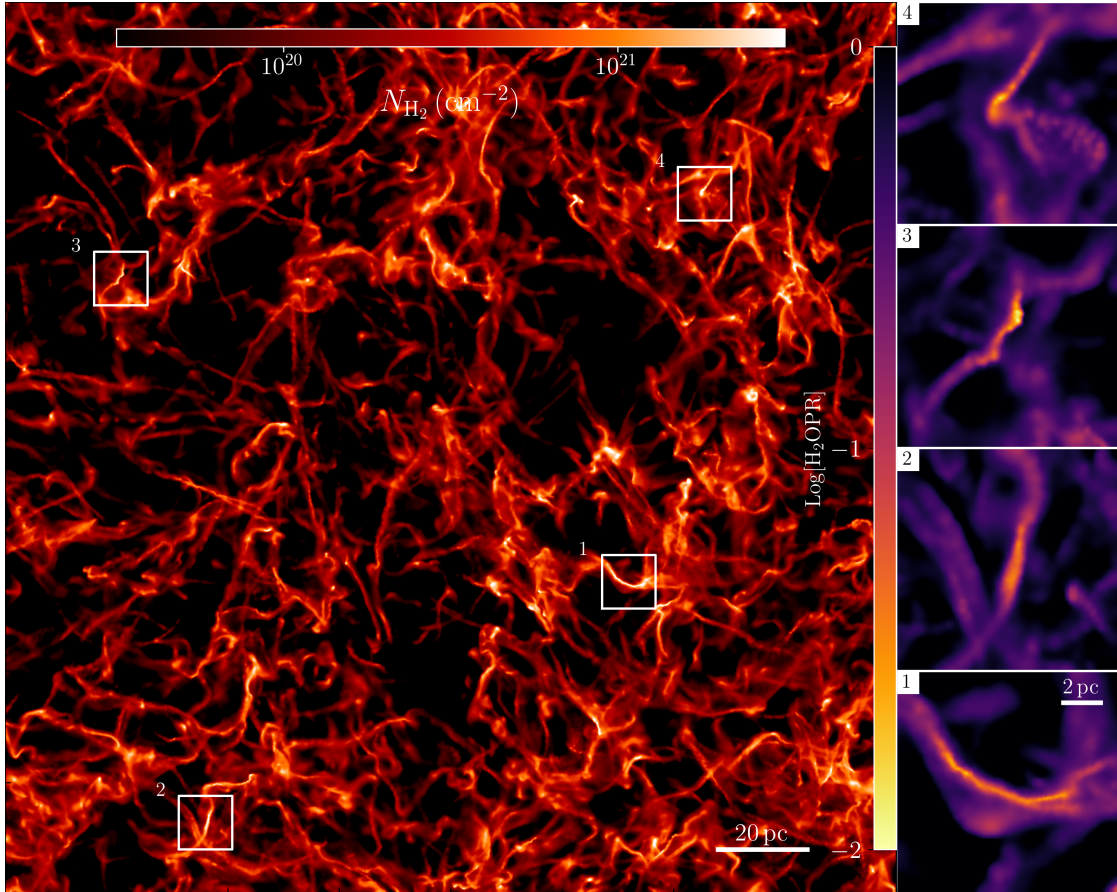
The region we simulated is a cubic box of 200 pc filled with homogeneous atomic gas and dust at a hydrogen nuclei density n<sub>H,tot</sub> = 5 cm<sup>-3</sup>, that corresponds to a total mass of

1.25 × 10<sup>6</sup> M<sub>⊙</sub>, in other words a typical giant molecular cloud. The mass and spatial resolution are 0.2 M<sub>⊙</sub> and ~60 AU, respectively, which allow us to properly resolve the formation of observed filaments and clumps with typical masses of 100–1000 M<sub>⊙</sub> and sizes of a few parsecs. Our simulations evolve the gas according to the ideal MHD equations, starting from an initially constant magnetic field of 3 μG aligned with the x direction. After an initial relaxation phase aimed at reaching a steady-state turbulence, we added self-gravity and on-the-fly non-equilibrium chemistry, the latter including ortho- and para- forms of H<sub>2</sub>, gas-grain interactions, photochemistry, and cosmic-ray induced reactions (see Appendix A for details).

### 2.3. Filament identification and analysis

During the evolution, filaments continuously form and disperse, up to the point at which gravity overcomes the thermal, turbulent, and magnetic support, resulting in the decoupling of the structure from the entire cloud, and the beginning of the collapse phase. In order to infer their properties over time, we identify them from two-dimensional (2D) H<sub>2</sub> column density maps using ASTRODENDRO, imposing a minimum background density N<sub>H<sub>2</sub></sub> = 10<sup>21</sup> cm<sup>-2</sup> and an rms error of σ<sub>N</sub> = 3 × 10<sup>20</sup> cm<sup>-2</sup>. In addition, we require a minimum area of at least 10 pixels. Among the found structures, we assume as filaments only the main branches of the dendrogram, excluding sub-branches and leaves in the hierarchy that likely represent clumps and cores within the filaments. The average properties of the identified structures are then computed by averaging the corresponding 2D maps pixel by pixel. For instance, for H<sub>3</sub><sup>+</sup> we employ column density maps integrated along the line of sight (the z-axis), whereas for magnetic field, temperature, and velocity dispersion we employ the H<sub>2</sub> density-weighted line-of-sight average maps, one per direction depending on the property considered. We found that computing the OPR using column densities or number densities can result in large differences that are due to the dilution effect caused by averaging over different regions along the line of sight. The filament mass is derived from the H<sub>2</sub> column density as M = ∑<sub>i</sub> N<sub>H<sub>2</sub></sub><sup>i</sup> × dS with dS is the pixel area and the sum is over the pixels associated with the filament. The Mach number M ≡ c<sub>s</sub>/σ<sub>v</sub> is determined from the isothermal sound speed c<sub>s</sub> = √(k<sub>B</sub>T<sub>gas</sub>/(μm<sub>H</sub>)), with m<sub>H</sub> the proton mass, k<sub>B</sub> the Boltzmann constant, and μ = 2.4 the molecular weight, and the average 3D velocity dispersion σ<sub>v</sub> = √(∑<sub>k=x,y,z</sub> σ<sub>v,k</sub><sup>2</sup>), where σ<sub>v,k</sub> is the average over the filament of the kth direction average velocity dispersion. Since we are not interested in an extremely accurate measure of the filament lengths and widths, and given the complex geometry of our simulated filaments, far from a perfect cylinder, we opt for a simple and approximate determination of the filament length and width. In detail, we proceed as follows: we first align the structure along its major axis (using the position angle of the ellipse associated to the branch of the dendrogram) and then define the length L as the maximum horizontal distance among the pixels belonging to the filament. The width W is then retrieved as W = A<sub>px</sub>/L, which guarantees that the area is preserved exactly, and the error in the estimate of W is as good as that used for L. Finally, a crucial parameter used to determine the fate of (potentially) star-forming filaments is the mass-to-length ratio M/L, which is typically compared to a critical value [M/L]<sub>c</sub> = 2c<sub>s</sub>/G, with G the gravitational constant.

<sup>1</sup> In this study we employ a cubic spline kernel with an effective number of neighbours of 32.



**Fig. 1.** *Left:* column density map of  $\text{H}_2$  at 4 Myr. *Right:*  $\text{OPR}^N$  in four filaments identified in the snapshot (corresponding to the four white squares in the left panel). The cloud already exhibits clear filamentary structures, and also some clumps forming within the most massive and largest ones. In these regions, the OPR is already low, with values typically below 0.1.

### 3. Results

We evolved the cloud for a few Myr, necessary for the first filaments to form out of the low-density material. During this stage, the temperature evolves self-consistently (see Appendix B), leading to an average Mach number in the cloud of  $\sim 5$ – $6$  after 4 Myr, consistent with typically observed values (Mac Low & Klessen 2004). The gas distribution in our fiducial simulated cloud at 4 Myr (just before the formation of the first sink particle, see Appendix A) is shown in Fig. 1, with the four panels on the right reporting the line-of-sight integrated  $\text{OPR}^N \equiv N_{\text{o-H}_2}/N_{\text{p-H}_2}$  of four massive filamentary structures out of hundreds identified. Qualitatively, Fig. 1 indicates that the OPR is around 0.1 already at these early stages, with peaks of  $\sim 0.01$  in the densest regions (clumps).

More in detail, in Table 1 we report the average properties of the four selected filaments (among the most massive and extended ones, i.e.  $M > 100 M_\odot$ ) at  $t = 4$  Myr. In particular, from left to right, we report the  $\text{H}_2$  column density  $N_{\text{H}_2}$ , the cosmic-ray ionisation rate  $\zeta_{\text{H}_2}$ , the gas temperature  $T_{\text{fil}}$ , the Mach number  $\mathcal{M}$ , the magnetic field  $B$  magnitude, and two values for the OPR, i.e.  $\text{OPR}^N$  and the density-weighted line-of-sight average ratio  $\text{OPR}^n \equiv \langle n_{\text{o-H}_2} \rangle / \langle n_{\text{p-H}_2} \rangle$ . Finally, in the last two columns, we report the mass per unit length  $M/L$  and the filament width  $W \equiv A_{\text{px}}/L$ , with  $A_{\text{px}}$  the effective pixel area of the dendrogram and  $L$  the major axis length of the filament. In general, the filaments identified in our simulations show typical properties consistent with observations (Arzoumanian et al. 2011): lengths  $L$

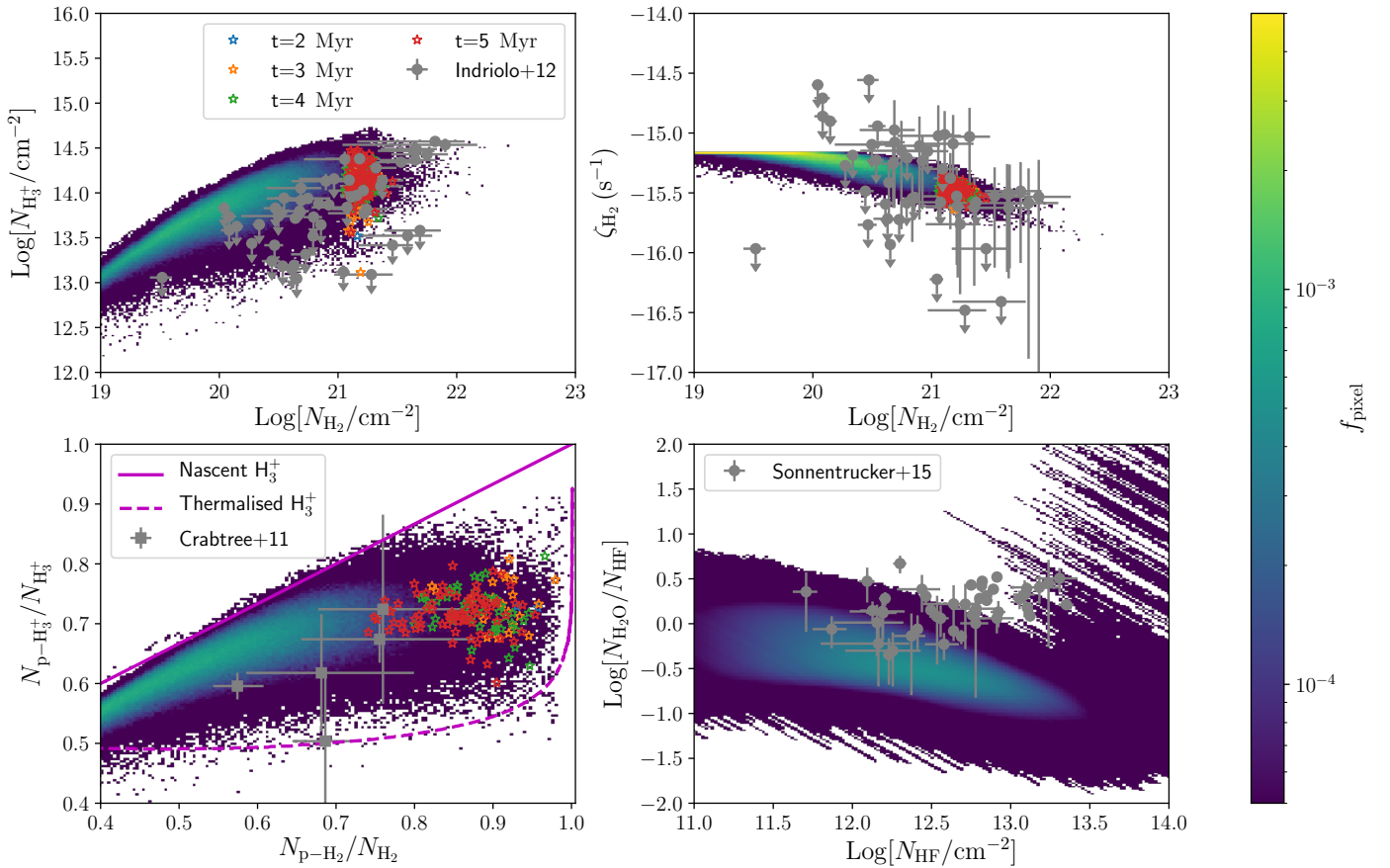
between 1 and 10 pc, axis ratios between 1:2 and 1:20, masses from a few tens up to a thousand solar masses, and densities not exceeding  $10^4 \text{ cm}^{-3}$  (being still in an initial collapse stage) with average temperatures around 30 K. The estimated mass per unit length ( $M/L$ ) is compared with the critical value for collapse, obtaining a full spectrum of values ranging from  $\sim 0.2$  (sub-critical) up to  $\sim 3$  (super-critical), with the four reported in Table 1 lying around 1.0–1.5.

To further confirm the accuracy of our modelling which, being developed with ab initio physics, has not been calibrated to reproduce real clouds, in Fig. 2 we compare our simulated cloud with observations of ‘diffuse’ clouds. In particular, we focus on properties connected to  $\text{H}_2$  and its OPR, like the  $\text{H}_3^+$  abundance (top left panel) and the cosmic-ray ionisation rate  $\zeta_{\text{H}_2}$  estimates (top right panel) (Indriolo 2012), the para-to-total ratio of  $\text{H}_3^+$  and  $\text{H}_2$  (Crabtree et al. 2011) (bottom left panel), and the water-to-HF column density ratio (bottom right panel) (Sonnentrucker et al. 2015). The remarkable agreement between our simulation and observational results suggests that our theoretical framework naturally produces reliable initial conditions for the collapsing filaments within molecular clouds (see Appendix C for a detailed analysis of the other simulations of our suite). In particular, in the bottom left panel, both our simulation and observations lie between the two theoretical curves, with our results more closely following the nascent distribution, which reflects the strong impact of cosmic rays. The abundance of water is slightly underestimated, particularly at high density, likely because of the missing formation channels of water on

**Table 1.** Main properties of four filaments out of hundreds in our fiducial simulation (see Fig. 1).

ID	$\log N_{\text{H}_2}$ ( $\text{cm}^{-2}$ )	$n_{\text{H}_2}$ ( $\text{cm}^{-3}$ )	$\log \zeta_{\text{H}_2}$ ( $\text{s}^{-1}$ )	$T_{\text{fil}}$ (K)	$\mathcal{M}$ –	$B$ ( $\mu\text{G}$ )	OPR <sup>N/n</sup> –	$[M/L]$ ( $M_{\odot} \text{pc}^{-1}$ )	$W$ (pc)
1	21.28	768.24	–15.53	31.4	10.60	7.60	0.10/0.06	77.56	0.10
2	21.28	1237.57	–15.54	28.0	3.55	10.13	0.08/0.03	54.09	0.12
3	21.22	1274.69	–15.58	23.9	7.78	9.92	0.06/0.01	41.33	0.16
4	21.38	2022.84	–15.51	28.3	6.14	9.30	0.06/0.03	52.70	0.08

**Notes.** From left to right, we report  $\text{H}_2$  column density, average number density  $n_{\text{H}_2}$ , average cosmic-ray ionisation rate  $\zeta_{\text{H}_2}$ , gas temperature  $T_{\text{fil}}$ , Mach number  $\mathcal{M}$ , magnetic field magnitude  $B$ , OPR (column density/average number density), mass per unit length  $M/L$ , and filament width  $W$ . All values are consistent with the expected ones derived from observations.

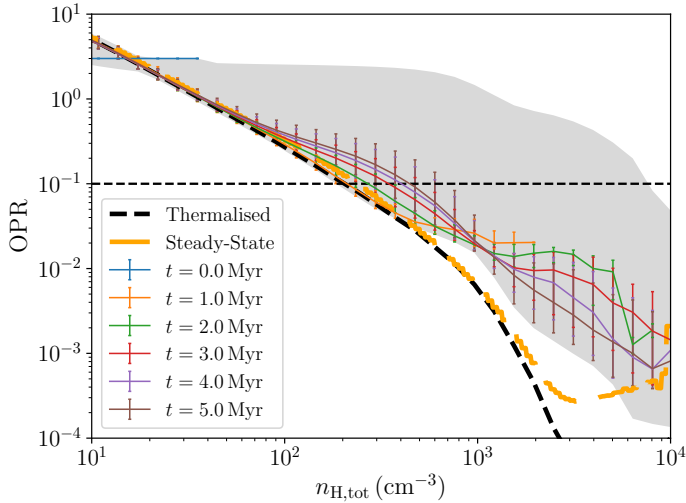


**Fig. 2.** Comparison of our simulation with existing observations of diffuse clouds, shown as grey dots (Crabtree et al. 2011; Indriolo 2012; Sonnentrucker et al. 2015). *Top left panel:* abundance of  $\text{H}_3^+$  relative to  $\text{H}_2$ , *top right panel:* corresponding  $\zeta_{\text{H}_2}$ , *bottom left panel:* para-to-total ratio for both  $\text{H}_3^+$  and  $\text{H}_2$ , and *bottom right panel:*  $\text{H}_2\text{O}$  relative abundance with respect to HF, where the HF column density in our simulation is extracted assuming a rigid scaling relative to  $\text{H}_2$  (Indriolo 2012) with scaling factors in the range  $0.5 \times 10^{-8} - 3.5 \times 10^{-8}$ . The coloured 2D histogram represents the full distribution of pixels (0.25 pc wide) in our simulation, and the blue, orange, green, and red stars the average abundances from the identified filaments at different times. Finally, the magenta lines correspond to the theoretical abundance of  $\text{H}_3^+$  assuming a nascent distribution (solid line), in which the formation of  $\text{H}_3^+$  from cosmic-ray-induced ionisation of  $\text{H}_2$  fully determines the relative abundance of the nuclear spin values, and a thermalised distribution (dashed line), in which the relative abundance is dominated by collisional exchange between  $\text{H}_3^+$  and  $\text{H}_2$  (Crabtree et al. 2011).

dust grains in our network, which are potentially relevant in cold gas (Cazaux et al. 2010; Sonnentrucker et al. 2015).

In order to disentangle whether time or density play the major role in producing these results, we report in Fig. 3 the evolution of the OPR distribution for every simulation element, as a function of the total hydrogen nuclei density. For this analysis, we directly use the local properties of the gas in the simulation, which allows us to avoid the dilution effects resulting from integrating along the line of sight (Ferrada-Chamorro et al. 2021).

We note that the distribution extends to very low values ( $\leq 0.1$ ) already at moderate densities ( $n_{\text{H,tot}} \sim 10^3 - 10^4 \text{ cm}^{-3}$ ). We also see that the distribution does not significantly change with time, suggesting that the dynamical evolution only has a second-order effect. This result suggests that, as soon as filaments form, the OPR is already well below 0.1, and that higher values found for OPR<sup>N</sup> and from observations are hugely affected by dilution effects (by one order of magnitude or more). Our results are robust even against different physical assumptions



**Fig. 3.** Three-dimensional distribution of the OPR as a function of  $n_{\text{H,tot}}$ . Each solid curve corresponds to a different time, with the error bars showing the 20th and 80th percentiles. The black dashed line is the thermalised OPR (i.e. the balance between the ortho-to-para and para-to-ortho conversion), valid for temperatures below  $\sim 100$  K, which correspond to densities higher than  $30\text{--}40\text{ cm}^{-3}$  in our specific case, and the orange long-dashed line to the steady-state value (i.e. the ratio at chemical equilibrium self-consistently computed using our network). For completeness, we also show the total uncertainty resulting from our entire suite of simulations (see Appendix C) as a grey shaded area.

(see Appendix C), with the upper limits (grey shaded area) still exhibiting very low OPR values when  $n_{\text{H,tot}} \gtrsim 10^4\text{ cm}^{-3}$ . At low density, the ratio tightly follows the thermalised value, which is also consistent with the equilibrium value. As soon as the gas becomes fully molecular, instead, the distribution starts to differ, but still decreases towards very small values (0.001 or less). At the highest densities probed, the OPR in the simulation settles on the equilibrium value, which showed a moderate increase above  $n_{\text{H,tot}} \sim 10^3\text{ cm}^{-3}$ , hence departing from the thermalised value.

#### 4. Discussion and conclusions

In this work, we have performed 3D MHD simulations of a molecular cloud with state-of-the-art on-the-fly non-equilibrium chemistry, finding that filaments are characterised by already low OPR values. Our simulations show a remarkable agreement with available observations, without any a priori tuning of the model. It is important to note that a proper comparison with other available works is difficult, as direct measurements or estimates of the OPR in diffuse and molecular clouds are rare. In particular, the few data in diffuse clouds (Crabtree et al. 2011) suggest values around 0.3–0.7, slightly higher than ours, but in relatively good agreement when considering the observational uncertainties (e.g. optical thickness of these lines and the possible dilution effects along the line of sight). Some indirect estimates in the range 0.001–0.2 have also been provided for pre-stellar cores (Troscompt et al. 2009; Brünken et al. 2014; Pagani et al. 2013), in regions with densities higher than those explored by our simulations and representing an advanced stage of the star formation process. However, the strong assumptions made to infer this fundamental quantity from different chemical proxies do not allow a reliable estimation from observations. In this context our study represents a relevant step forward in improving the fundamental knowledge of the star formation process.

The absence of sulfur chemistry in our work might represent a limitation, since it could reduce the ortho-to-para conversion efficiency by removing  $\text{H}^+$  via the reaction path  $\text{S} + \text{H}^+ \rightarrow \text{S}^+ + \text{H}$  (Furuya et al. 2015). However, Furuya et al. (2015) show that (i) sulfur is quickly adsorbed on dust grains as soon as  $\text{S}^+$  recombines and, as a consequence, (ii) this effect is relevant only for extremely high metal abundances, which are not compatible with the measured values of sulfur from observations. Hence, we assume that sulfur chemistry does not play a relevant role within the context of our model, and therefore it does not affect our conclusions.

Concluding, our state-of-the-art 3D magnetohydrodynamic simulations of molecular clouds formation with on-the-fly non-equilibrium chemistry indicate that the H<sub>2</sub> OPR quickly evolves with density, reaching very low values (but far from the thermalised value) already at moderate densities typical of proto-filaments. This is particularly relevant for smaller scale studies, in which the unconstrained initial OPR is either varied in the allowed range to bracket its effect on the results (see e.g. Sipilä et al. 2015; Kong et al. 2015; Bovino et al. 2020) or conservatively assumed to be around 0.1 (see e.g. Jensen et al. 2021), which is higher than ours ( $10^{-3}\text{--}10^{-2}$ ). The results in this work therefore establish that the initial conditions of the star formation process in filaments are characterised by already low OPR and typically high  $\zeta_{\text{H}_2}$ , both conditions that would dramatically boost the deuteration mechanism and shorten the corresponding timescales. We note that, even in filaments characterised by low cosmic-ray ionisation rates (Indriolo 2012), the expected OPR would be only moderately higher, and never as high as 0.1. For the first time, we have been able to determine the initial conditions of star-forming filaments from ab initio conditions, in particular the chemical abundances of important species like H<sub>2</sub> (ortho- and para-), H<sub>3</sub><sup>+</sup>, CO, and H<sub>2</sub>O, which have far-reaching implications for the deuteration process (hence for the reliability of chemical clocks), for the observed HDO/H<sub>2</sub>O ratios in planet-forming regions (which strongly depends on the OPR), and its connection with the origin of water in our Solar System.

*Acknowledgements.* We thank the anonymous referee for their useful comments. AL acknowledges funding from MIUR under the grant PRIN 2017-MB8AEZ. The computations/simulations were performed with resources provided by the Kultrun Astronomy Hybrid Cluster. This research made use of ASTRODENDRO, a PYTHON package to compute dendrograms of astronomical data (<http://www.dendrograms.org>), and PYNBODY, a PYTHON package to analyse astrophysical simulations (Pontzen et al. 2013). Part of this work was supported by the German Deutsche Forschungsgemeinschaft, DFG project number Ts17/2–1.

#### References

- Altwegg, K., Balsiger, H., Bar-Nun, A., et al. 2015, *Science*, **347**, 1261952
- André, P., Men'shchikov, A., Bontemps, S., et al. 2010, *A&A*, **518**, L102
- Arzoumanian, D., André, P., Didelon, P., et al. 2011, *A&A*, **529**, L6
- Bakes, E. L. O., & Tielens, A. G. G. M. 1994, *ApJ*, **427**, 822
- Bauer, A., & Springel, V. 2012, *MNRAS*, **423**, 2558
- Bergin, E. A., & Tafalla, M. 2007, *ARA&A*, **45**, 339
- Bergin, E. A., & van Dishoeck, E. F. 2012, *Philos. Trans. R. Soc. London Ser. A*, **370**, 2778
- Bovino, S., Grassi, T., Schleicher, D. R. G., & Caselli, P. 2017, *ApJ*, **849**, L25
- Bovino, S., Ferrada-Chamorro, S., Lupi, A., et al. 2019, *ApJ*, **887**, 224
- Bovino, S., Ferrada-Chamorro, S., Lupi, A., Schleicher, D. R. G., & Caselli, P. 2020, *MNRAS*, **495**, L7
- Brünken, S., Sipilä, O., Chambers, E. T., et al. 2014, *Nature*, **516**, 219
- Cazaux, S., Cobut, V., Marseille, M., Spaans, M., & Caselli, P. 2010, *A&A*, **522**, A74
- Ceccarelli, C., Caselli, P., Bockelée-Morvan, D., et al. 2014, in *Protostars and Planets VI*, eds. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning, 859

- Crabtree, K. N., Indriolo, N., Kreckel, H., Tom, B. A., & McCall, B. J. 2011, *ApJ*, **729**, 15
- Draine, B. T. 1978, *ApJS*, **36**, 595
- Draine, B. T., & Sutin, B. 1987, *ApJ*, **320**, 803
- Federrath, C., Banerjee, R., Clark, P. C., & Klessen, R. S. 2010a, *ApJ*, **713**, 269
- Federrath, C., Duval, J., Klessen, R. S., Schmidt, W., & Low, M. M. M. 2010b, *Highlights Astron.*, **15**, 404
- Federrath, C., Rathborne, J. M., Longmore, S. N., et al. 2016, *ApJ*, **832**, 143
- Ferrada-Chamorro, S., Lupi, A., & Bovino, S. 2021, *MNRAS*, **505**, 3442
- Furuya, K., Aikawa, Y., Hincelin, U., et al. 2015, *A&A*, **584**, A124
- Furuya, K., Aikawa, Y., Hama, T., & Watanabe, N. 2019, *ApJ*, **882**, 172
- Gavilan, L., Vidali, G., Lemaire, J. L., et al. 2012, *ApJ*, **760**, 35
- Glover, S. C. O., Federrath, C., Mac Low, M.-M., & Klessen, R. S. 2010, *MNRAS*, **404**, 2
- Grassi, T., Bovino, S., Schleicher, D. R. G., et al. 2014, *MNRAS*, **439**, 2386
- Grassi, T., Bovino, S., Haugbølle, T., & Schleicher, D. R. G. 2017, *MNRAS*, **466**, 1259
- Hama, T., & Watanabe, N. 2013, *Chem. Rev.*, **113**, 8783
- He, J., & Vidali, G. 2014, *Faraday Discuss.*, **168**, 517
- Hocuk, S., Cazaux, S., & Spaans, M. 2014, *MNRAS*, **438**, L56
- Hogerheijde, M. R., Bergin, E. A., Brinch, C., et al. 2011, *Science*, **334**, 338
- Hopkins, P. F. 2015, *MNRAS*, **450**, 53
- Hopkins, P. F. 2016, *MNRAS*, **462**, 576
- Hopkins, P. F., & Raives, M. J. 2016, *MNRAS*, **455**, 51
- Indriolo, N. 2012, *Philos. Trans. R. Soc. London Ser. A*, **370**, 5142
- Jensen, S. S., Jørgensen, J. K., Furuya, K., Haugbølle, T., & Aikawa, Y. 2021, *A&A*, **649**, A66
- Jørgensen, J. K., Belloche, A., & Garrod, R. T. 2020, *ARA&A*, **58**, 727
- Kong, S., Caselli, P., Tan, J. C., Wakelam, V., & Sipilä, O. 2015, *ApJ*, **804**, 98
- Mac Low, M.-M., & Klessen, R. S. 2004, *Rev. Mod. Phys.*, **76**, 125
- Mandal, A., Federrath, C., & Körtgen, B. 2020, *MNRAS*, **493**, 3098
- McKee, C. F., & Ostriker, E. C. 2007, *ARA&A*, **45**, 565
- Molinari, S., Swinyard, B., Bally, J., et al. 2010, *A&A*, **518**, L100
- Padoan, P., & Nordlund, Å. 2011, *ApJ*, **730**, 40
- Padoan, P., Pan, L., Haugbølle, T., & Nordlund, Å. 2016, *ApJ*, **822**, 11
- Padovani, M., Galli, D., & Glassgold, A. E. 2009, *A&A*, **501**, 619
- Padovani, M., Ivlev, A. V., Galli, D., & Caselli, P. 2018, *A&A*, **614**, A111
- Pagani, L., Lesaffre, P., Jorfi, M., et al. 2013, *A&A*, **551**, A38
- Perets, H. B., & Biham, O. 2006, *MNRAS*, **365**, 801
- Pontzen, A., Roškar, R., Stinson, G. S., et al. 2013, *Astrophysics Source Code Library* [record ascl:1305.002]
- Sipilä, O., Caselli, P., & Harju, J. 2015, *A&A*, **578**, A55
- Sonnentrucker, P., Wolfire, M., Neufeld, D. A., et al. 2015, *ApJ*, **806**, 49
- Springel, V. 2005, *MNRAS*, **364**, 1105
- Troscopnt, N., Faure, A., Maret, S., et al. 2009, *A&A*, **506**, 1243
- Tsuge, M., Namiyoshi, T., Furuya, K., et al. 2021, *ApJ*, **908**, 234
- Ueta, H., Watanabe, N., Hama, T., & Kouchi, A. 2016, *Phys. Rev. Lett.*, **116**, 253201
- Vidali, G., & Li, L. 2010, *J. Phys. Condens. Matter*, **22**, 304012
- Visser, R., van Dishoeck, E. F., & Black, J. H. 2009, *A&A*, **503**, 323
- Wakelam, V., Bron, E., Cazaux, S., et al. 2017, *Mol. Astrophys.*, **9**, 1
- Walmsley, C. M., Flower, D. R., & Pineau des Forêts, G. 2004, *A&A*, **418**, 1035
- Watanabe, N., Kimura, Y., Kouchi, A., et al. 2010, *ApJ*, **714**, L233
- Wolfire, M. G., McKee, C. F., Hollenbach, D., & Tielens, A. G. G. M. 2003, *ApJ*, **587**, 278

## Appendix A: Numerical methods

### A.1. Turbulence driving

In order to drive turbulence in the box with properties similar to observed molecular clouds, the gas distribution is stirred by a random acceleration field obtained via an Ornstein-Uhlenbeck process (Federrath et al. 2010b; Bauer & Springel 2012). The energy power spectrum normalisation is set to obtain a velocity dispersion in the cloud  $\sigma_v \sim 7 \text{ km s}^{-1}$ , and the auto-correlation time of the random process is set to  $\tau = 10 \text{ Myr}$ . Energy is injected only at large scales, according to a parabolic power spectrum peaking at  $k = 3$  (extending from  $k = 2$  up to  $k = 4$ ), assuming a half-solenoidal half-compressive driving (Bauer & Springel 2012). This assumption is reasonably consistent with recent simulations of molecular cloud formation in which turbulence was driven self-consistently via random supernova explosions (Padoan et al. 2016). During this relaxation phase, which lasts for 50 Myr in order to let turbulence fully develop, we do not include self-gravity, and we assume an isothermal equation of state with  $T = 5000 \text{ K}$  and no chemistry evolution (Mandal et al. 2020). This allows us to start the self-consistent evolution of the cloud with ab-initio chemical abundances typical of the warm neutral medium, avoiding any pre-processing of the species during the relaxation. Nevertheless, to further corroborate our results and isolate the effect of self-gravity, we also perform an additional experiment in which chemistry and proper cooling are already included during the relaxation phase.

### A.2. Sink formation

Near the end of the simulation, we expect clumps to form in filaments and a few resolution elements to reach very high densities, thus requiring very short timescales to integrate the dynamics. In order to avoid this undesired slow-down and let the simulation to evolve for longer times, we convert the gas hitting the resolution limit of the simulation into proto-stellar objects, commonly dubbed sink particles, which are able to grow via accretion of surrounding gas. In this work, we include a simple sink formation scheme. Gas particles that i) are above  $n_{\text{H,tot}} > 10^{10} \text{ cm}^{-3}$ ; ii) show negative velocity divergence; and iii) are located at a relative gravitational potential minimum are converted into sink particles unless another sink is found within their kernel volume. After a sink has formed, we allow it to accrete gas within its kernel volume, corresponding to a sphere enclosing an effective number of 32 neighbours that matches the sink formation conditions. However, since our main interest is in the pre-stellar phase, rather than in the star formation process itself, we stop our simulations after a few sinks have formed in the box, making the details of the sink formation scheme almost irrelevant.

### A.3. Cosmic-ray flux determination

The cosmic-ray ionisation rate in molecular clouds is affected by several uncertainties since it strongly depends on the physical conditions inside and outside the cloud. For this reason, most studies to date explore different values of the cosmic-ray ionisation rate  $\zeta_{\text{H}_2}$  typically in the range  $1.3 \times 10^{-17} - 1.3 \times 10^{-16}$ . In some of our simulations we also employ a constant value; instead, in a real cloud cosmic rays are attenuated as they move deeper into it, hence a more consistent modelling should take into account this aspect. However, detailed cosmic-ray propagation in 3D simulations is computationally expensive, and would

add an additional layer of complexity to our already computationally expensive simulations. For this reason in this work we opt for an effective model that, starting from the local properties of each resolution element, allows us to follow the variations of  $\zeta_{\text{H}_2}$  in the cloud, at a moderate computational cost (Padovani et al. 2018),

$$\zeta_{\text{H}_2} = \zeta_{\text{H}_2,\text{p}} + \zeta_{\text{H}_2,\text{e}}, \quad (\text{A.1})$$

where the two contributions are from protons and electrons, respectively, and are defined as

$$\zeta_{\text{H}_2,\text{p}} = \begin{cases} 6.8 \times 10^{-16} N_{20}^{-0.423} & N_{\text{eff}} < 10^{25} \text{ cm}^{-2} \\ 5.4 \times 10^{-18} \exp(-\Sigma_{\text{eff}}/38) & \text{otherwise} \end{cases}, \quad (\text{A.2})$$

$$\zeta_{\text{H}_2,\text{e}} = \begin{cases} 1.4 \times 10^{-19} N_{20}^{-0.04} & N_{\text{eff}} < 10^{25} \text{ cm}^{-2} \\ 3.3 \times 10^{-20} \exp(-\Sigma_{\text{eff}}/71) & \text{otherwise} \end{cases}, \quad (\text{A.3})$$

where  $N_{20} = N_{\text{eff}}/10^{20} \text{ cm}^{-2}$  and  $\Sigma_{\text{eff}} = 2.36 m_{\text{H}} N_{\text{eff}}$ , with  $N_{\text{eff}}$  the effective column density traversed by cosmic rays. To determine  $N_{\text{eff}}$ , we assume that the magnetic field lines are relatively not curved (as expected in this earlier star formation stage) and they have small intensity variations, allowing us to consider the total column density  $N_{\text{H}_2} = N_{\text{o-H}_2} + N_{\text{p-H}_2}$  as a reliable first-order approximation of  $N_{\text{eff}}$ . The column density  $N_{\text{H}_2}$  is then estimated using a local approximation (Grassi et al. 2017) in which  $N_x = 1.87 \times 10^{21} (n_x/10^3 \text{ cm}^{-3})^{2/3}$ , with  $x$  the chemical species considered. When the magnetic field lines are not straight, our estimate of  $N_{\text{eff}}$  represents a lower limit to the actual value, which translates into our cosmic-ray ionisation rate being an upper limit. This is why, in our suite, we chose an average model among those in the literature (Padovani et al. 2009), and also explored very different conditions (i.e. a conservative and uniform  $\zeta_{\text{H}_2} = 1.3 \times 10^{-17} \text{ s}^{-1}$ , and the locally varying value) to bracket the possible conditions of real clouds.

### A.4. Ortho-to-para H<sub>2</sub> conversion on dust

Recent experimental works both on amorphous solid water (Ueta et al. 2016) and on bare silicates (Tsuge et al. 2021), show the efficiency of the ortho-to-para H<sub>2</sub> conversion on the surface of dust grains. To consistently follow this process within our simulations and evaluate its overall impact on the OPR, we employ state-of-the-art frameworks (Bovino et al. 2017; Furuya et al. 2019) in which the conversion rates in units of  $\text{s}^{-1}$  are defined as

$$k_{\text{op}} = k_{\text{ads}}^{\text{oH}_2} \eta_{\text{op}}, \quad (\text{A.4})$$

$$k_{\text{po}} = k_{\text{ads}}^{\text{pH}_2} \eta_{\text{po}}, \quad (\text{A.5})$$

where the  $k_{\text{ads}}^i = S v_i \sigma_{\text{dust}}$  are the adsorption rates of the species on the surface of grains, with  $S$  being the sticking coefficient (here assumed to be 1 for simplicity),  $v_i$  the thermal gas speed, and  $\sigma_{\text{dust}}$  the distribution-averaged grain geometrical cross-section. The efficiency of the process is regulated by the factor  $\eta_i$  which represents the competition between the conversion process and desorption, and it is defined as (Furuya et al. 2019)

$$\eta_{\text{op}} = \frac{t_{\text{des}}}{t_{\text{des}} + \tau_{\text{conv}}} \frac{1}{1 + \gamma}, \quad (\text{A.6})$$

$$\eta_{\text{po}} = \frac{t_{\text{des}}}{t_{\text{des}} + \tau_{\text{conv}}} \frac{\gamma}{1 + \gamma}, \quad (\text{A.7})$$

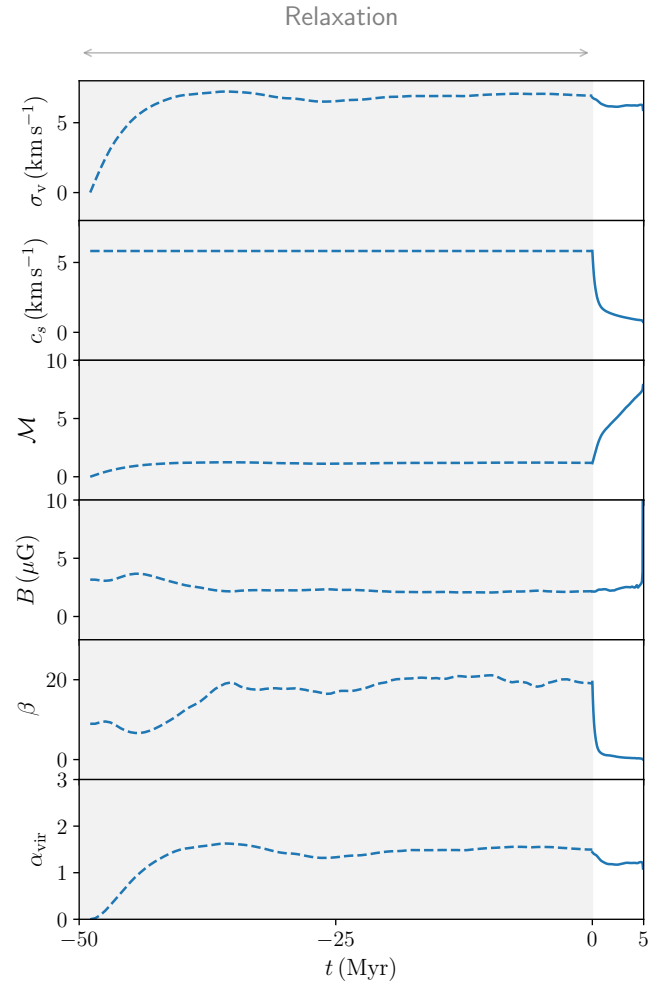
where the desorption time is calculated as the minimum between the thermal desorption time and the cosmic-ray induced desorption time,  $\tau_{\text{conv}}$  is the experimental conversion time (Tsuge et al. 2021) fitted as  $\tau_{\text{conv}} = 6.3 \times 10^4 T_{\text{dust}}^{-1.9}$ , and  $\gamma = 9 \exp(-170.5/T_{\text{dust}})$  is the thermalised value of the  $\text{H}_2$  OPR, assuming the energy difference between o- $\text{H}_2$  and p- $\text{H}_2$  on the grains is the same as that in the gas phase. This allows us to include both the ortho-to-para conversion and its inverse process (para-to-ortho) in the  $\text{H}_2$  rate equations. However, since the inverse process is, in general, negligible at low temperatures (Bovino et al. 2017), we opt for completely neglecting it in our simulations, which in practice corresponds to assuming  $\gamma = 0$  (Bovino et al. 2017). We note that our approach is based on a single average binding energy approximation, in particular  $E_b = 600$  K as reported in the literature (Perets & Biham 2006; Vidali & Li 2010; He & Vidali 2014), and refer to other recent works (Furuya et al. 2019) for a more comprehensive treatment, which also includes the thermal hopping between different adsorption sites. The effect of these processes on the evolution of the  $\text{H}_2$  OPR is discussed in Appendix C.

### A.5. HF abundance in diffuse clouds

Observations of water in diffuse clouds typically lack a direct measure of the  $\text{H}_2$  column density, using instead HF (or other hydrides) as an alternative proxy (Sonnentrucker et al. 2015). Although the correlation  $N_{\text{HF}} = \chi_{\text{HF}} N_{\text{H}_2}$  is found to be quite tight, the uncertainty in the conversion factor reaches up to a factor of  $\sim 7$ , ranging from  $0.5 \times 10^{-8}$  up to  $3.5 \times 10^{-8}$  (Indriolo 2012). For this reason, and considering that our chemical network does not include HF, no simple comparison between simulations and observations exists, and particular attention must be taken when converting  $\text{H}_2$  to HF (or vice versa). For the comparison between our runs and observations of Fig. 2, we opt therefore for a more sophisticated procedure: for each value of  $N_{\text{H}_2}$  in our simulations, we estimate the  $N_{\text{HF}}$  using 100 different values of  $\chi_{\text{HF}}$  within the observed range, so that this source of uncertainty is properly accounted for, and combine all these measures in a single 2D histogram reported in the bottom right panel of the figure, appropriately normalised to the total number of measures available.

## Appendix B: Relaxation phase

Our simulation suite is composed of four runs, three of them starting from our fiducial isothermal relaxation without chemistry, that we call *RelaxIso* and one in which proper cooling and chemistry are accounted for also during relaxation, that we call *RelaxChem*, in which  $\zeta_{\text{H}_2} = 1.3 \times 10^{-17} \text{ s}^{-1}$  is kept constant over time in the entire box. The relaxation phase is necessary to guarantee that turbulence fully develops in the box, producing initial conditions of the actual simulations that more closely represent realistic molecular clouds (Federrath et al. 2010a). To give an idea of the global evolution of our fiducial simulated cloud, we show in Fig. B.1 the main properties of the box during both the relaxation phase (reported as a grey shaded area with negative times) and the actual run (reported with positive times), i.e. the density-weighted average 3D velocity dispersion  $\sigma_v$ , sound speed, Mach number,  $\beta = P/P_B$  parameter, with  $P = \rho c_s^2$  the thermal pressure (with  $\rho$  the total gas density) and  $P_B = \langle B \rangle^2 / (8\pi)$  the magnetic pressure, and virial parameter  $\alpha_{\text{vir}} = 5\sigma_v^2 L / (3GM)$ , with  $L$  the box size and  $M$  the box total mass. During the relaxation phase,  $\sigma_v$  increases up to the desired value as a result of turbulence driving, whereas the sound speed



**Fig. B.1.** Evolution of the global properties of the turbulent cloud in our fiducial model, during the relaxation phase (shaded area with negative times) and the actual run (positive times). While turbulence leads to an initial increase of  $\sigma_v$  (reflected in all the other properties it affects), the sound speed remains constant during relaxation, producing a Mach number slightly above unity. During the run, instead,  $M$  increases quickly to the typically observed values, because of cooling. On average, the magnetic field does not change significantly, as long as the typical density interval remains small.

stays constant, because of the constant temperature assumption, which is also reflected in the virial parameter increase to about 1.5 and the Mach number staying close to unity. The modest change in density distribution results in a small decrease of the average magnetic field, and the corresponding increase of  $\beta$  by a factor of two. After relaxation,  $\sigma_v$  only modestly varies, while  $c_s$  significantly decreases because of cooling, producing a rapid increase of  $M$  to the typically observed values. These variations reflect directly on  $\beta$  and  $\alpha_{\text{vir}}$ , except for the magnetic field that suddenly rises above  $10 \mu\text{G}$  around  $t = 5$  Myr, when a large portion of the gas starts to collapse.

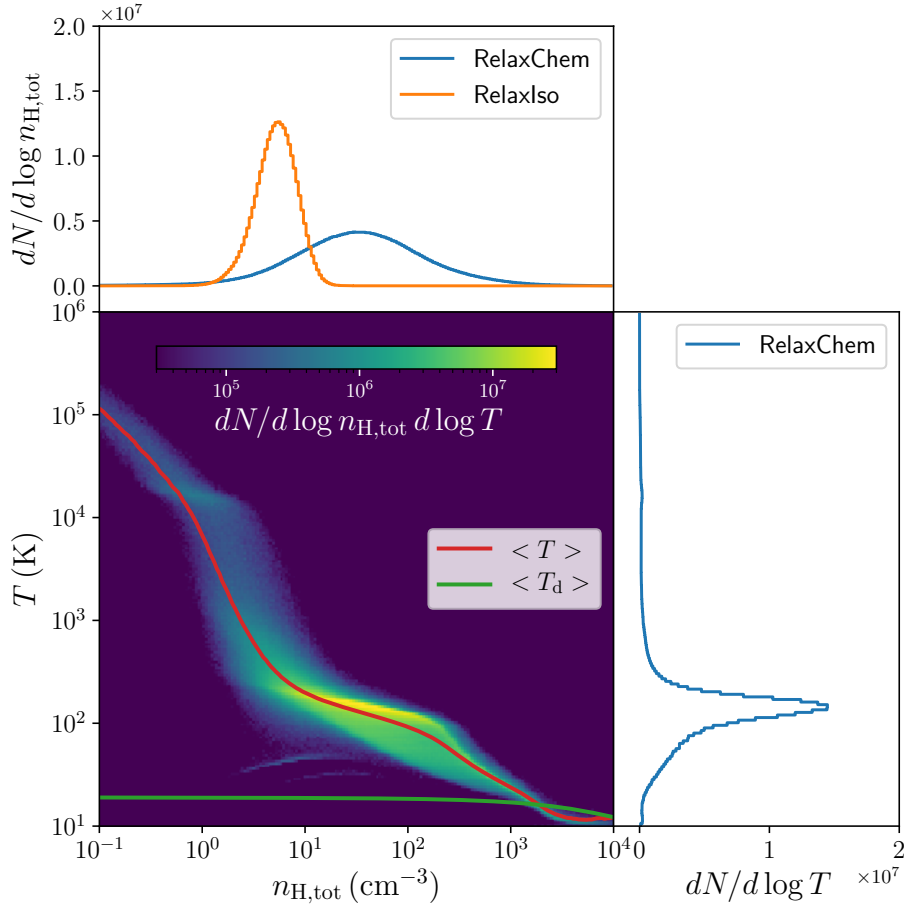
For completeness, we also report in Table B.1 the same main properties at the end of the relaxation phase for the two relaxation models we considered, in addition to the density-weighted average density and temperature. While the turbulence-driven properties are the same in both relaxation models, the thermodynamic is not, resulting in very different average temperatures and densities. To better clarify how the gas is distributed in the two cases, in Fig. B.2 we show the density and temperature



**Table B.1.** Main properties of our turbulent box after 50 Myr of relaxation for the two different considered models *RelaxIso* and *RelaxChem*.

Model	$\langle n_{\text{H,tot}} \rangle$ (cm <sup>-3</sup> )	$\langle T \rangle$ (K)	$\langle \sigma_v \rangle$ (km s <sup>-1</sup> )	$c_s$ (km s <sup>-1</sup> )	$\mathcal{M}$	$\langle \beta \rangle$	$\alpha_{\text{vir}}$
RelaxChem	123.8	128	7.525	0.663	11.35	0.318	1.761
RelaxIso	6.2	$5 \times 10^3$	7.162	5.820	1.230	2.008	1.600

**Notes.** The inclusion of cooling and chemistry during the relaxation phase allows the gas to spread over a larger density–temperature interval, resulting in very different thermal properties (typical of colder and denser gas), whereas those fully determined by turbulence remain almost identical (the 3D velocity dispersion  $\sigma_v$  and the virial parameter  $\alpha_{\text{vir}}$ ).



**Fig. B.2.** Distribution of the gas at the end of the relaxation phase. The density–temperature diagram corresponds to *RelaxChem* (this distribution does not change when gravity is included, apart from extending to higher densities as the gas collapses), with the red line representing the average temperature, and the green line the average dust temperature. The temperature-only distribution in the right-hand panel shows the typical temperature of the gas, 100–200 K, which is higher than the typical temperature assumed in isothermal molecular cloud simulations. In the top panel we show instead the density distribution for both relaxation models, where *RelaxChem*, extending to higher densities because of cooling, is shown in blue, and *RelaxIso*, more concentrated around the initial density, in orange.

distributions (the latter only for the *RelaxChem* case). Overlaid on the density-temperature plot of *RelaxChem* we also report the average temperature (red) and dust temperature (green) curves, that differ significantly at low density, while they couple at  $n_{\text{H,tot}} \sim 10^4$  cm<sup>-3</sup>. In the right panel, we show the temperature-only distribution, which peaks around 100–200 K, at which most of the diffuse gas settles. In the top panel we report instead the density-only distribution, in this case also for *RelaxIso* (orange histogram). As expected, in *RelaxChem*, gas cooling allows the gas to get denser after turbulence-induced compression, spreading over a larger density interval, that follows a Gaussian-like profile (the typically expected log-normal density probability distribution function). On the other hand, *RelaxIso* keeps the gas

to moderate densities, producing a much narrower distribution centred around the average density of the initial conditions.

### Appendix C: The full simulation suite

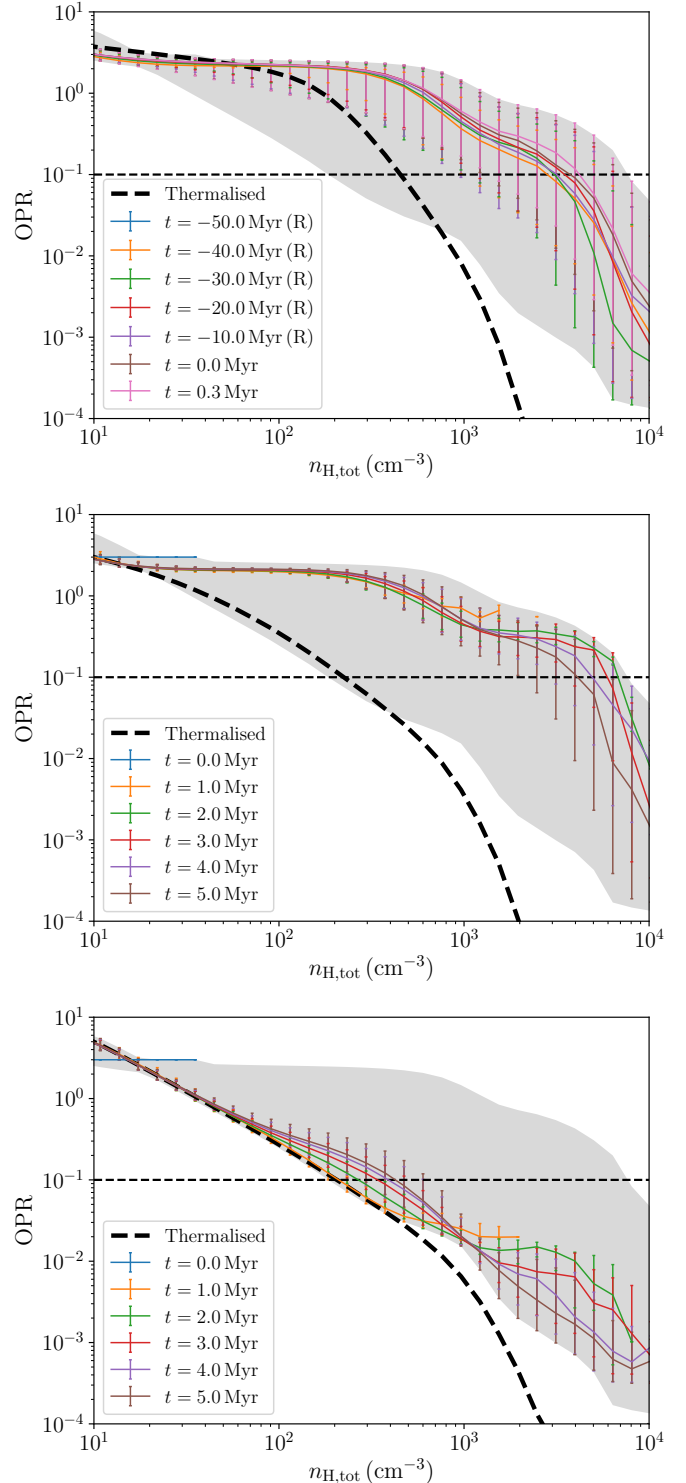
The four simulations we performed are meant to cover most of the plausible parameter space, thus properly constraining our models relative to observations. In particular, compared to our fiducial model, we consider the following: a) *RelaxChem* plus self-gravity, named *RelaxChem\_SG*; b) *RelaxIso* plus self-gravity, named *RelaxIso\_SG*; c) the fiducial model (see main text); and d) the fiducial model with the addition of ortho-to-para conversion on dust (as described in Section A), named

*RelaxIso\_CR\_OPdust*. Similar to Fig. 3 in the main text, Fig. C.1 reports the OPR evolution for models *RelaxChem\_SG* (top panel), *RelaxIso\_SG* (middle panel), and *RelaxIso\_SG\_OPdust* (bottom panel) for completeness. *RelaxIso\_CR\_OPdust* gives almost identical results to our fiducial model, suggesting that the OPR conversion on dust, which slightly accelerates the ortho-to-para conversion, does not change our picture significantly, and that the gas-phase collisions with  $\text{H}^+$  and  $\text{H}_3^+$  are already efficient enough to bring the OPR down, without the need of this additional mechanism. For *RelaxChem\_SG*, in which chemistry is also evolved during the relaxation phase, the OPR distribution in these stages is reported with negative times and an additional ‘(R)’ in the label. We immediately see that, even without self-gravity, when cooling is included, the denser gas stirred by turbulence quickly settles on a steady-state OPR distribution, and the addition of gravity<sup>2</sup> does not alter the distribution at all. Compared to the fiducial run, here the OPR is typically higher, farther from the thermalised value. Nevertheless, a clear drop can be observed around  $n_{\text{H,tot}} \sim 3\text{--}4 \times 10^3 \text{ cm}^{-3}$ , which in this case also yields an OPR well below 0.1 at the typical densities of star-forming filaments ( $n_{\text{H,tot}} \gtrsim 10^4 \text{ cm}^{-3}$ ). A similar trend can be observed in *RelaxIso\_SG*, although the high-density drop is even steeper than in the previous case, consistent with the weak time-dependence of the distribution (in this second case chemistry was not present during the initial relaxation, hence the processing time is much shorter). Combining the results of these two models, we can conclude that the higher OPR distribution relative to our fiducial model is the result of the very low cosmic-ray ionisation rate, which is reasonable for protostellar cores, but not for the typical conditions of molecular clouds (Padovani et al. 2009, 2018). Nonetheless, all our models, even the most pessimistic ones, result in low OPR ( $\leq 0.01$ ) at the typical densities of star-forming filaments, and this further corroborates our conclusions in the main text. To further support our claims about the effect of a too low  $\zeta_{\text{H}_2}$ , we compare in Fig. C.2 our *RelaxIso\_SG* with observations, as in Fig. 2. We immediately see that, with respect to our fiducial run, the results here are completely off, with the simulation yielding lower abundances than those observed. This is perfectly consistent with the discrepancy between our assumed  $\zeta_{\text{H}_2}$  with respect to the observationally inferred value (Indriolo 2012), even though these results might be compatible with the claimed non-detections. Moreover, in the bottom left panel, our results almost perfectly lie on the thermalised distribution, consistently with the fact that cosmic-ray induced ionisation of  $\text{H}_2$  becomes almost negligible with respect to the atom exchange reactions dominating in thermalised conditions.

Although it is not the focus of this study, we report in Fig. C.3 the abundance of gaseous CO in our fiducial simulation for completeness at different times. CO forms efficiently in our filaments, reaching the canonical fraction of about  $10^{-4}$ , despite the high cosmic-ray ionisation rate. We also note that CO does not exhibit any strong time-dependence in the probed density range, but this is not surprising since freeze-out on dust grains, that starts to deplete CO at  $n_{\text{H,tot}} \sim 10^4 \text{ cm}^{-3}$  as expected, still has a negligible impact.

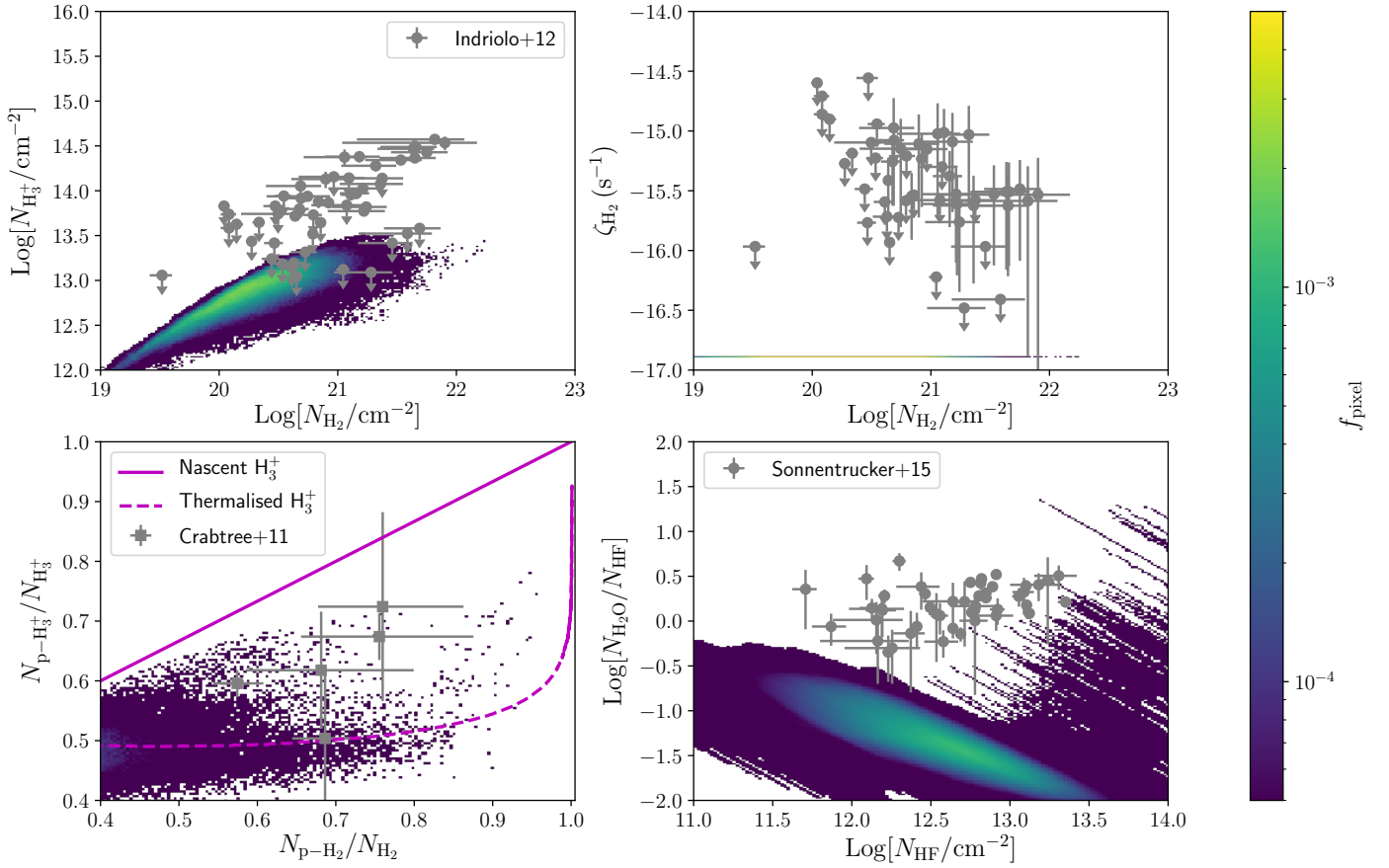
As a final example of the unprecedented level of detail of our simulations including on-the-fly complex chemistry, we also report in Fig. C.4 the column density maps of relevant chemical species that are directly tracked in our simulations (i.e. CO,  $\text{H}_2\text{O}$ ,

<sup>2</sup> In this last case, dense filaments and clumps already form during the relaxation phase, and they collapse in less than 1 Myr after self-gravity is included.

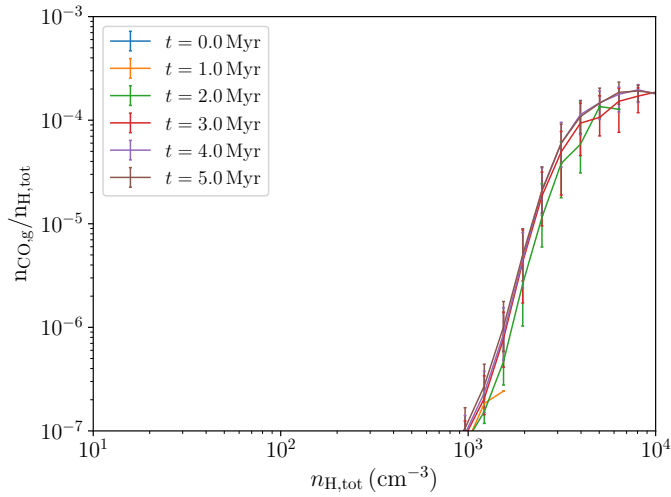


**Fig. C.1.** Same as Fig. 3, but for models *RelaxChem\_SG* (top panel), *RelaxIso\_SG* (middle panel), and *RelaxIso\_SG\_OPdust* (bottom panel). The OPR distribution shows very mild variations with time in all cases, with the values at  $n_{\text{H,tot}} \sim 10^4 \text{ cm}^{-3}$  always being well below 0.1.

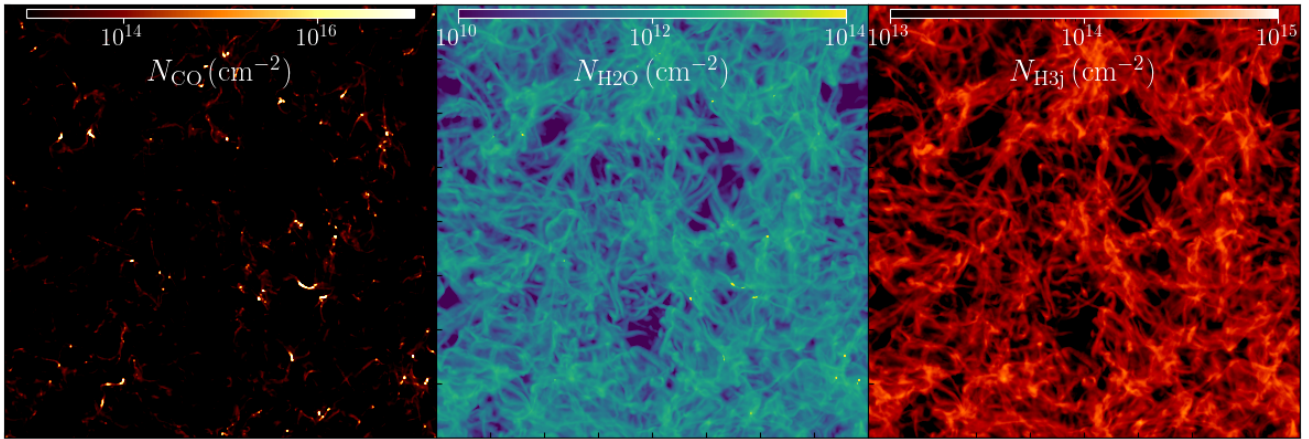
and  $\text{H}_3^+$ ). CO is the mostly concentrated species, and appears in large amounts only in dense gas (proto-filaments), with the background reaching an abundance that is at most four orders of magnitude lower. Water (and especially  $\text{H}_3^+$ ) are instead more uniformly distributed, showing mild variations across very different density conditions.



**Fig. C.2.** Same as Fig. 2, but for model *RelaxIso\_SG*. Unlike in our fiducial model, here the abundances are much lower than those observed, with the data in the bottom panel settling on the thermalised distribution.



**Fig. C.3.** Fraction of CO in gas phase in our fiducial model, as a function of total hydrogen density. CO is able to form efficiently, reaching the canonical abundance in gas above  $n_{\text{H,tot}} = 3\text{--}4 \times 10^3 \text{ cm}^{-3}$ , before freeze-out on dust grains starts to deplete it, as expected. We also find that the abundance does not significantly vary with time, similarly to the OPR, as long as depletion remains almost negligible.



**Fig. C.4.** Examples of column density maps of three important chemical species in molecular clouds, CO, H<sub>2</sub>O, and H<sub>3</sub><sup>+</sup> (from left to right). While H<sub>3</sub><sup>+</sup> is quite uniformly distributed in the box, H<sub>2</sub>O forms in larger amounts in moderately higher density gas, and CO reaches typically observed values ( $N_{\text{CO}} \gtrsim 10^{17} \text{ cm}^{-2}$ ) only in proto-filaments (where there is a huge difference, about four orders of magnitude, between the filaments and the background).