HAL
open science

Survey of ortho-H 2 D + in high-mass star-forming regions

G. Sabatini, S. Bovino, A. Giannetti, F. Wyrowski, M. Órdenes, R. Pascale, T. Pillai, M. Wienen, T. Csengeri, K. M Menten

To cite this version:

G. Sabatini, S. Bovino, A. Giannetti, F. Wyrowski, M. Órdenes, et al.. Survey of ortho-H 2 D + in high-mass star-forming regions. Astronomy and Astrophysics - A&A, 2020, 644, pp.A34. 10.1051/0004-6361/202039010 . hal-03040674

HAL Id: hal-03040674

https://hal.science/hal-03040674v1

Submitted on 17 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A&A 644, A34 (2020)
https://doi.org/10.1051/0004-6361/202039010
© ESO 2020

Astronomy & Astrophysics

Survey of ortho-H_2D^+ in high-mass star-forming regions

G. Sabatini[1,2,3,⋆], S. Bovino[2], A. Giannetti[3], F. Wyrowski[4], M. A. Órdenes[2], R. Pascale[5], T. Pillai[6], M. Wienen[4], T. Csengeri[7], and K. M. Menten[4]

[1] Dipartimento di Fisica e Astronomia, Universitá degli Studi di Bologna, Via Gobetti 93/2, 40129 Bologna, Italy
e-mail: giovanni.sabatini4@unibo.it
[2] Departamento de Astronomía, Facultad Ciencias Físicas y Matemáticas, Universidad de Concepción, Av. Esteban Iturra s/n Barrio Universitario, Casilla 160, Concepción, Chile
[3] INAF – Istituto di Radioastronomia – Italian node of the ALMA Regional Centre (It-ARC), Via Gobetti 101, 40129 Bologna, Italy
[4] Max-Planck-Institut für Radioastronomie, Auf dem Hügel, 69, 53121 Bonn, Germany
[5] INAF – Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Gobetti 93/3, 40129 Bologna, Italy
[6] Institute for Astrophysical Research, 725 Commonwealth Ave, Boston University Boston, MA 02215, USA
[7] Laboratoire d'Astrophysique de Bordeaux, Univ. Bordeaux, CNRS, B18N, allée Geoffroy Saint-Hilaire, 33615 Pessac, France

Received 23 July 2020 / Accepted 25 September 2020

ABSTRACT

Context. Deuteration has been suggested to be a reliable chemical clock of star-forming regions due to its strong dependence on density and temperature changes during cloud contraction. In particular, the H_3^+ isotopologues (e.g. ortho-H_2D^+) seem to act as good proxies of the evolutionary stages of the star formation process. While this has been widely explored in low-mass star-forming regions, in the high-mass counterparts only a few studies have been pursued, and the reliability of deuteration as a chemical clock remains inconclusive.

Aims. We present a large sample of o-H_2D^+ observations in high-mass star-forming regions and discuss possible empirical correlations with relevant physical quantities to assess its role as a chronometer of star-forming regions through different evolutionary stages.

Methods. APEX observations of the ground-state transition of o-H_2D^+ were analysed in a large sample of high-mass clumps selected from the ATLASGAL survey at different evolutionary stages. Column densities and beam-averaged abundances of o-H_2D^+ with respect to H_2, X(o-H_2D^+), were obtained by modelling the spectra under the assumption of local thermodynamic equilibrium.

Results. We detect 16 sources in o-H_2D^+ and find clear correlations between X(o-H_2D^+) and the clump bolometric luminosity and the dust temperature, while only a mild correlation is found with the CO-depletion factor. In addition, we see a clear correlation with the luminosity-to-mass ratio, which is known to trace the evolution of the star formation process. This would indicate that the deuterated forms of H_3^+ are more abundant in the very early stages of the star formation process and that deuteration is influenced by the time evolution of the clumps. In this respect, our findings would suggest that the X(o-H_2D^+) abundance is mainly affected by the thermal changes rather than density changes in the gas. We have employed these findings together with observations of $H^{13}CO^+$, DCO^+, and $C^{17}O$ to provide an estimate of the cosmic-ray ionisation rate in a sub-sample of eight clumps based on recent analytical work.

Conclusions. Our study presents the largest sample of o-H_2D^+ in star-forming regions to date. The results confirm that the deuteration process is strongly affected by temperature and suggests that o-H_2D^+ can be considered a reliable chemical clock during the star formation processes, as proved by its strong temporal dependence.

Key words. astrochemistry – stars: formation – ISM: molecules – molecular processes – surveys

1. Introduction

Crucial to the study of star formation is the measurement of timescales, which may be mass- and density-dependent (e.g. Urquhart et al. 2018; see also Motte et al. 2018 for a review). The duration of the star formation process is crucial to distinguishing between competing theories, which predict a slow or a fast evolution towards the formation of stars (e.g. Mouschovias et al. 2006; Hartmann et al. 2012), estimating its fundamental energy budget, and understanding, and thus predicting, its impact on the chemical enrichment of the interstellar medium (ISM).

Chemistry is a powerful tool for inferring how long the star-forming gas remains dense and cold, and thus is capable of actually forming stars. In addition, chemical signatures can help to characterise evolved sources, due to the change in temperature and density, and their effects on the chemistry.

In the last few decades, among the chemical processes that occur under the typical conditions prevalent in Dark Clouds, the study of freeze-out (e.g. Caselli et al. 1998; Chen et al. 2010; Hernandez et al. 2011; Giannetti et al. 2014; Feng et al. 2019; Bovino et al. 2019) and deuterium fractionation processes (e.g. Millar et al. 1989; Pagani et al. 1992a,b; Ceccarelli et al. 2007; Chen et al. 2010; Sipilä et al. 2010; Fontani et al. 2011; Kong et al. 2015; Körtgen et al. 2018) have certainly played a central role in astrochemical research. Especially during the first phases of star formation, the volume density of molecular hydrogen, $n(H_2)$, and the temperature of the clouds, T, reach the ideal conditions (i.e. $n(H_2) >$ few $\times 10^4$ cm^{-3} and $T < 20$ K) to favour freeze-out (or depletion), by which several C-, N-, and O-bearing species, in particular carbon monoxide (CO),

⋆ Marco Polo fellow – University of Bologna.

are efficiently removed from the gas phase and trapped on the surface of the dust grains (e.g. Kramer et al. 1999; Bergin et al. 2002; Caselli et al. 2008; Fontani et al. 2012; Wiles et al. 2016; Sabatini et al. 2019; see also Bergin & Tafalla 2007 for a review). The CO adsorption onto the grain surfaces also favours the progressive enrichment of deuterium atoms compared to hydrogen in the molecules, through the deuterium fractionation process (e.g. Ceccarelli et al. 2014). In this context, the main driver of deuteration, the trihydrogen cation H_3^+, leads to the enrichment of deuterated molecules like H_2D^+ and N_2D^+ in the gas phase (e.g. Pagani et al. 2009; Vastel et al. 2012; Caselli et al. 2019).

Using cold-gas chemistry to measure the age of low-mass cores (i.e. $M < 8\ M_\odot$) has been exploited via different tracers (e.g. via H_2D^+ by Brünken et al. 2014), while in the high-mass regime age estimates are much more complex, in particular due to the uncertainties in the dynamical and chemical initial conditions.

A first attempt to adopt this kind of analysis to the high-mass regime was carried out by Kong et al. (2016) for cores of ∼15–60 $M_\odot$ by measuring the deuterium fraction via N_2H^+. However, it was shown by Pillai et al. (2012) that this species may not be ideal for tracing the first stages of the high-mass star formation process. They obtained maps for part of the DR21 complex of the ortho−H_2D^+ $J_{K_a,K_c} = 1_{10}-1_{11}$ (hereafter o-H_2D^+) and N_2D^+ $J = 3$–2 transitions, observed with the *James Clerk Maxwell* Telescope (JCMT; Holland et al. 1999) and the Submillimiter Array (SMA; Ho et al. 2004), respectively. The data reveal very extended o-H_2D^+ emission, in agreement with the results of Vastel et al. (2006) for the low-mass source L1544, and that this species mainly traces gas that is not seen in dust continuum emission or in the interferometric N_2D^+ data. H_2D^+ may thus be sensitive to gas that eludes detection in the most commonly used tracers, and can represent an even earlier stage in the process of star formation.

In a pilot study to this work, with the Atacama Pathfinder EXperiment 12-meter submillimeter telescope (APEX; Güsten et al. 2006), Giannetti et al. (2019) detected o-H_2D^+ and N_2D^+ in three high-mass star-forming clumps for the first time, opening the possibility of investigating their abundance variation in sources at different evolutionary stages (G351.77–0.51 complex; e.g. Leurini et al. 2019). They considered sources ranging from a clump that is still quiescent at 70 μm to one that hosts luminous young stellar objects (YSOs), adopting the classification from Giannetti et al. (2014), König et al. (2017), and Urquhart et al. (2018). Giannetti et al. (2019) observed that the abundance of N_2D^+ progressively increases with evolution, showing a difference of a factor of ∼2 between the least and the most evolved clump. Particularly relevant is also that the o-H_2D^+ abundance decreases in the two most evolved clumps (by a factor of ∼10), likely due to the chemical conversion of H_2D^+ into D_2H^+ and D_3^+ (which would also boost the production of N_2D^+) or to the destruction by desorbed CO in the presence of a protostellar object. Similar results were also reported by Kong et al. (2016), who obtained upper limits for abundances for o-H_2D^+ of $\sim 10^{-11}$ in two high-mass star-forming regions associated with outflow activity (see Tan et al. 2016), and high abundances of N_2D^+, confirming the fact that N_2D^+ is forming at later times compared to o-H_2D^+ and under different physical conditions.

Attempts to demonstrate this hypothesis have been recently reported by Miettinen (2020), using APEX observations in three prestellar and three protostellar cores in Orion-B9. Although observing a downward trend in o-H_2D^+ abundance of about a factor of 4 with evolution, the anti-correlation with N_2D^+ was not confirmed. However, as pointed out by the authors, their spatial offset between the N_2D^+ and the o-H_2D^+ observations (∼10″ on average) might be the cause of the observed behaviour (i.e. the two molecules sample different regions).

With the present work we extend the number of high-mass star-forming regions with o-H_2D^+ detections, assembling a sample of 106 massive clumps selected from several spectral line surveys carried out on clumps as part of the ATLASGAL survey (see Sect. 2 for details). Our goal is to test the trend reported by Giannetti et al. (2019) for o-H_2D^+, on a sample that could be statistically representative of regions with ongoing massive star formation activity in order to exclude possible bias induced by particular initial chemical conditions that may have affected previous (much smaller) samples.

This paper has the following structure: we describe the sample and its selection in Sect. 2. In Sect. 3, we discuss how the spectra have been obtained and reduced, and we summarise the physical properties of the sources in our sample. The information on how the column densities and the dynamical parameters of the sources have been calculated is presented in Sect. 4. In Sect. 5, we present the comprehensive set of correlations we found between the o-H_2D^+ abundance and the other physical quantities, and we present the new estimates for the cosmic-ray ionisation rate (CRIR) for hydrogen molecules, based on the same data. Following the evolutionary stages of massive clumps, in Sect. 6 we discuss how o-H_2D^+ can be used as a powerful chemical clock to follow the evolution of high-mass star-forming regions. Finally, in Sect. 7 we summarise our conclusions.

2. Sample

The APEX Telescope Large Area Survey of the Galaxy (ATLASGAL; Schuller et al. 2009; Csengeri et al. 2014 and Li et al. 2016) provides an ideal basis for detailed studies of large numbers of massive clumps in different stages of the evolutionary sequence of high-mass star-forming regions (Molinari et al. 2008). The ATLASGAL Compact Source Catalog has delivered $\sim 10^4$ clumps (e.g. Contreras et al. 2013; Urquhart et al. 2013; Csengeri et al. 2014), with reliable estimates of kinematic distances, masses, luminosities, and dust temperatures and distribution (Urquhart et al. 2014; Wienen et al. 2015; Urquhart et al. 2018). The ATLASGAL-TOP100 sample (hereafter TOP100; see Giannetti et al. 2014) includes 111 clumps; it was selected from ATLASGAL as a flux-limited sample, using additional infrared (IR) criteria in order to include sources in different evolutionary stages (see König et al. 2017). Molecular line surveys have been carried out on the TOP100 with APEX-12 m, Mopra-22 m, and the IRAM-30 m single-dish telescopes, between 80 and 345 GHz, covering more than 120 GHz of bandwidth in three spectral windows[1] that contain a multitude of emission lines of both simple and complex molecules. A detailed analysis of a small sample of chemical species allowed us to estimate their excitation parameters, and to derive accurate column densities (e.g. Giannetti et al. 2014, 2017; Csengeri et al. 2016; Tang et al. 2018). From the analysis of this data set Giannetti et al. (2017) demonstrate that temperature, column densities of several tracers, and H_2 volume density increase with time, as a function of evolutionary class and luminosity-to-mass ratio (L/M). These trends were interpreted as the signature of the initial compression phase of the clump material and the progressive accretion onto the forming YSOs. Therefore, the evolutionary sequence defined for the TOP100 is statistically valid, and the TOP100 can be

[1] Not yet completely published.

Table 1. Summary of the physical properties of the ATLASGAL sources in our sample.

ATLASGAL-ID	$d_\odot$ (kpc)	D_{GC} (kpc)	R_{eff} [(a)] (pc)	T_{dust} (K)	ΔT_{dust} [(b)] (K)	v_{lsr} (km s^{-1})	M_{clump} (10^2 $M_\odot$)	L_{bol} (10^2 $L_\odot$)	N(H_2) log$_{10}$(cm^{-2})	n(H_2) log$_{10}$(cm^{-3})	Class [(a,b)]
G08.71–0.41 [(c)]	4.8	4.0	1.0	11.8	0.3	39.4	16.6	5.0	22.8	4.0	IRw
G13.18+0.06 [(c)]	2.4	5.9	0.5	24.2	0.8	49.9	3.7	83.2	22.9	4.3	70w
G14.11–0.57 [(c)]	2.6	6.9	0.5	22.4	0.8	20.8	3.5	31.8	22.9	4.4	IRw
G14.49–0.14 [(c)]	3.9	5.4	0.8	12.4	0.4	39.5	19.2	7.5	23.1	4.4	70w
G14.63–0.58 [(c)]	1.8	6.9	0.4	22.5	0.4	18.5	2.5	27.8	23.0	4.6	IRw
G18.61–0.07 [(c)]	4.3	5.3	0.7	13.8	0.3	46.6	8.8	5.9	22.8	4.2	IRw
G19.88–0.54 [(c)]	3.7	5.3	0.6	24.2	1.4	44.9	8.0	124.0	23.1	4.5	IRb
G28.56–0.24 [(c)]	5.5	4.8	1.3	11.7	0.1	87.3	54.1	17.7	23.1	4.2	IRw
G333.66+0.06 [(c)]	5.3	4.5	1.1	17.8	0.3	−85.1	14.2	42.7	22.7	3.9	70w
G351.57+0.76 [(c)]	1.3	7.0	0.3	17.0	0.1	−3.2	1.6	4.3	22.7	4.4	70w
G354.95–0.54 [(c)]	1.9	7.4	0.4	19.1	1.3	−5.4	1.5	4.8	22.6	4.2	70w
G12.50-0.22 [(a)]	2.6	5.9	0.4	13.0	0.2	35.5	1.2	1.6	22.8	4.4	70w
G14.23–0.51 [(a)]	1.5	6.9	0.5	17.2	2.7	19.5	7.2	14.8	23.3	4.8	HII
G15.72–0.59 [(a)]	1.8	6.6	0.2	12.1	0.5	17.8	1.7	0.3	22.8	4.7	IRw
G316.76-0.01 [(a)]	2.5	6.7	0.2	18.9	0.4	−39.9	4.7	24.2	23.1	5.0	IRb
G351.77-CL7	1.0	7.4	0.1 [(d)]	13.0	0.3 [(e)]	−3.2 [(d)]	1.2	0.2	22.9	5.2	70w

Notes. Physical properties (observed and derived) of our sample. Top row: TOP100 sources; bottom row: ATLASGAL sources not in the TOP100; T_{dust} and L_{bol} are obtained from a SED fit described in Appendix D in Urquhart et al. (2018), while their uncertainties are derived as discussed in Sect. 2; R_{eff} values are evaluated from the dust continuum at 870 μm, while N(H_2) and M_{clump} come from the 870 μm flux peaks and the integrated flux density over R_{eff}, respectively; n(H_2) values are lower limits derived here as $n(H_2) = N(H_2)/2R_{eff}$. [(a)]Distances, R_{eff}, v_{lsr}, M_{clump}, L_{bol}, N(H_2), and classification from Urquhart et al. (2018): the error associated with N(H_2) is 20% for each source; v_{lsr} values are derived from the NH_3 (1, 1) data published in Wienen et al. (2012). [(b)]Classification scheme from König et al. (2017). [(c)]$d_\odot$, M_{clump}, L_{bol}, and N(H_2) from Giannetti et al. (2017); R_{eff} and T_{dust} from König et al. (2017); D_{GC} from Urquhart et al. (2018); v_{lsr} from Giannetti et al. (2014) and derived from the $C^{17}O$ J = 3–2. [(d)]R_{eff} and v_{lsr} are from Leurini et al. (2011). v_{lsr} is derived from the $C^{18}O$ J = 2–1. [(e)]T_{dust} standard deviation computed on the clump region defined in Leurini et al. (2019).

considered representative of the Galactic protocluster population through all the evolutionary stages.

In this work, we present new observations of the o-H_2D^+ ground-state transition $J_{K_a,K_c} = 1_{10}-1_{11}$, based on two spectral line surveys of the ATLASGAL sources. The first set of observations[2] is composed of a sample of young massive clumps, selected from previous observations that showed a high degree of deuterated ammonia (Wienen et al. 2020). The second set of observations were part of a survey of [CI] $^3P_1$–$^3P_0$ fine structure line at 492 GHz (Lee et al., in prep.) carried out on the TOP100 sample. These were observed simultaneously with the dual-band FLASH345/460 GHz receiver on APEX, tuned to 372 GHz to observe the N_2H^+ J = 4–3 molecular line, which is very close to the o-H_2D^+ transition. Since the main goal of that survey was the observation of the [CI] line, the noise levels were often not adequate for an o-H_2D^+ detection. For this reason many of the sources observed in this project were used here to calculate o-H_2D^+ detection limits (we postpone further discussions to Sect. 4).

In Table 1, we summarise the physical parameters, for each clump with detection, from the dust continuum taken from Giannetti et al. (2017), König et al. (2017), and Urquhart et al. (2018). The heliocentric distances, $d_\odot$, of our sample are taken from König et al. (2017) and Urquhart et al. (2018), and range between 1 and 5.5 kpc; the galactocentric radii, D_{GC}, are between 4.5 and ∼7.5 kpc with a mean associated error of 0.3 kpc, and are taken from Urquhart et al. (2018) assuming a distance to the Galactic Center of 8.35 kpc (Reid et al. 2014); the effective radii, R_{eff}, range from 0.1 to 1.3 pc; the radial velocities, v_{lsr}, are taken from Giannetti et al. (2014) and Wienen et al. (2012), based on the $C^{17}O$ J = 3–2 and the NH_3 (1, 1) lines, respectively[3]. We note that a few TOP100 sources show an offset of ∼1 km s^{-1} between the v_{lsr} derived from $C^{17}O$ and the central velocity of the o-H_2D^+ lines. However, the v_{lsr} derived from o-H_2D^+ are consistent with the values from N_2H^+ J = 4–3 at ∼372.672 GHz. Such offsets are potential indications of strong internal motions driven by contraction, consistent with our finding that these clumps are on the verge of collapse (see Sect. 4.2).

The dust temperatures, T_{dust}, cover the typical range of values between ∼10 and ∼25 K, while ΔT_{dust} is the error derived from the covariance matrix of the Spectral Energy Distribution (SED) fit performed using a Levenberg-Marquardt least-squares minimisation (see Appendix D in Urquhart et al. 2018). We note that a wider exploration of the parameter space used to describe the dust properties, would lead to an error of ∼10% in T_{dust} (e.g. Schisano et al. 2020).

The SED fit assumes a dust absorption coefficient at 870 μm, κ_{870} = 1.85 cm^2 g^{-1}, and a dust emissivity index β = 1.75. The clump masses, M_{clump}, and bolometric luminosities, L_{bol}, are within ∼100–5400 $M_\odot$ and ∼10^{-1}–10^4 $L_\odot$, respectively. The molecular hydrogen column density, N(H_2), spans less than an order of magnitude among the different sources, with values of log$_{10}$(N(H_2) (cm^{-2})) between 22.7 and 23.3. The sample includes clumps associated with all evolutionary classes. We refer to Giannetti et al. (2017), König et al. (2017), and Urquhart

[2] Sixteen objects in total, not fully contained in the TOP100.

[3] See also the ATLASGAL Database Server at https://atlasgal.mpifr-bonn.mpg.de/cgi-bin/ATLASGAL_DATABASE.cgi

et al. (2018) for the details on how each parameter and the relative error were derived.

Since more than 2/3 of the sample in Table 1 is part of the TOP100, we employ the evolutionary classes defined in Giannetti et al. (2014), Csengeri et al. (2016), and König et al. (2017) assigned in the TOP100, where four stages were defined: (1) 70w, i.e. 70 μm weak, which represents the earliest stage of massive-star formation and includes starless or prestellar cores; (2) IRw, i.e. 24 μm weak, but bright at 70 μm, associated with sources in an early star formation stage, with clumps likely dominated by cold gas; (3) IRb, bright both at 70 and 24 μm, but not in radio continuum at 5–9 GHz and associated with the high-mass protostellar stage; (4) HII, as the compact HII region phase, bright also in radio continuum. We extended the same classification to the sources not included in the TOP100 by following the classes defined in Urquhart et al. (2018) and double-checking the 70 and 24 μm maps[4,5] according to the TOP100 classification: (*i*) Quiescent → 70w; (*ii*) Protostellar → IRw; (*iii*) YSOs → IRb; (*iv*) MSF → HII.

3. Observations and data reduction

3.1. o-H_2D^+ detections

The o-H_2D^+ $J_{K_a,K_c} = 1_{10}-1_{11}$ spectra (rest frequency 372.4214 GHz; Amano & Hirao 2005) were observed using the on-off (ONOFF) observing mode with the FLASH+ dual-frequency MPIfR principal investigator (PI) receiver (Klein et al. 2014), mounted at the APEX telescope (Güsten et al. 2006).

The complete set of observations consists of 106 sources belonging to the ATLASGAL survey. Among them, 99 targets are also contained in the TOP100 sample, observed in three projects (IDs: 0101.F-9517, M-097.F-0039-2016, and M-098.F-0013-2016; PI: F. Wyrowski), from July 2017 to December 2018, with a total of 11568 spectra (with about 20–30 s of integration time each). The receiver bandpass is covered with overlapping FFTS backends of 2.5 GHz widths (Klein et al. 2012), one of them covering from 372.250 to 374.750 GHz with a native spectral resolution of 0.038 km s^{-1} and with final mean system temperatures, $T_{\rm sys}$, ranging between ~700 and 1300 K, depending on the weather conditions. At ~372 GHz, the APEX telescope has an effective full width at half maximum (FWHM) beam size of 16.8″, with a corresponding main-beam efficiency of ~0.73[6]. The observations were reduced using a GILDAS class[7] Python interface pipeline.

To discard the low-quality spectra throughout the survey, the first step of the procedure was to evaluate the $T_{\rm sys}$ in each spectrum. We generated $T_{\rm sys}$ distributions source-by-source to avoid possible anomalies between different APEX projects. Only a few spectra (138 in total) are associated with clear $T_{\rm sys}$ outliers and were therefore discarded.

As a second step the main-beam efficiency was applied to the spectra after setting the reference frequency to 372.421 GHz. From each spectrum we subtracted a first-order polynomial baseline around the line, properly masked. In some rare cases, where a first order was not enough, a second-order (or third-order) baseline was used. In the last step, all spectra related to each source were averaged (stitched) with a spectral resolution between 0.5 and 0.6 km s^{-1} to reduce the noise, generating one final spectrum per source.

Despite the relatively high Einstein-A coefficient (i.e. $\log_{10}(A_{\rm ul}[{\rm s}^{-1}]) = -3.96$; CDMS[8]; Müller et al. 2001), the $J_{K_a,K_c} = 1_{10}-1_{11}$ transition of o-H_2D^+ is usually faint due to the low relative abundance of this species with respect to H_2. Typical observed abundances range between ~10^{-10}–10^{-12} in many low-mass star-forming regions (e.g. Vastel et al. 2006; Harju et al. 2008; Caselli et al. 2008; Friesen et al. 2010, 2014; Miettinen 2020) and few high-mass regime counterparts (e.g. Harju et al. 2006; Swift 2009; Pillai et al. 2012; Giannetti et al. 2019). For this reason, in this work we evaluated the significance of each detection considering the integrated signal-to-noise ratio of the line (iS/N) and taking each line with iS/N > 3 as a detection. For reference, the integrated main-beam temperature, $\int_{\rm ch} T_{\rm mb}\,{\rm d}\upsilon$, and the integrated root mean square noise, *i*rms[9], of each line is reported in the spectra shown in Figs. 1 and 2.

Sixteen ATLASGAL sources with o-H_2D^+ detection[10] are reported, 11 of which belong to the TOP100 sample. All the evolutionary classes are contained in our sample allowing us to study (within statistical limits) how the o-H_2D^+ emission changes through the massive star formation process, and how these changes correlate with the other observed or derived quantities.

3.2. Additional tracers: $H^{13}CO^+$, DCO^+ and $C^{17}O$

For one-half of the sources with o-H_2D^+ detection, we collected further observations of $H^{13}CO^+$, DCO^+, and $C^{17}O$ that we employ to estimate the cosmic-ray ionisation rate (explained in Sect. 5.2). We simultaneously observed the rotational transition of both $H^{13}CO^+$ and DCO^+ in 30 ATLASGAL sources, using the Eight MIxer Receiver E90 (EMIR; Carter et al. 2012) at the IRAM-30 m single-dish telescope[11] (IRAM project-id 107-15). From the study of Wienen et al. (2020), the 30 brightest clumps in deuterated ammonia were selected. The receiver was tuned at 75 GHz, allowing the simultaneous detection of $H^{13}CO^+$ and DCO^+ at 87 and 72 GHz, respectively, in 2 × 8 GHz sidebands with a velocity resolution of 0.75 km s^{-1}. The sources were observed with an average integration time of ~1 h in February 2016. The average $T_{\rm sys}$ are ~130 K for the $H^{13}CO^+$ lines and 210 K for the DCO^+ lines. The data were originally calibrated to antenna temperature, $T_{\rm A}^*$, and then converted to $T_{\rm mb}$ using the tabulated values for the IRAM-30 m forward efficiency, $\eta_{\rm eff} = 0.95$, and main-beam efficiency, $\eta_{\rm mb} = 0.74$. The final averaged rms noise was 0.02 K for both lines.

The $C^{17}O$ $J = 1-0$ observations at 112 GHz, were taken from Giannetti et al. (2014) for the TOP100 sources and from Csengeri et al. (2016) for the two sources not included in the TOP100, to which we refer for a full description of the data sets.

The full set of observations used in this work is summarised in Table 2, where the details on the observed molecules (Cols. 1–3), the telescope setup used (Cols. 4–7) and the references from which the data were selected (Col. 7) are reported.

[4] https://atlasgal.mpifr-bonn.mpg.de/cgi-bin/ATLASGAL_DATABASE.cgi
[5] http://www.alienearths.org/glimpse/glimpse.php
[6] http://www.apex-telescope.org/telescope/efficiency/
[7] https://www.iram.fr/IRAMFR/GILDAS/

[8] Cologne Database for Molecular Spectroscopy (CDMS): https://cdms.astro.uni-koeln.de/cdms/portal/
[9] *i*rms = $\sqrt{N_{\rm ch}}\,{\rm d}\upsilon\,\sigma$, where $N_{\rm ch}$ is the number of channels included in the integration, dυ is the velocity resolution, and σ is the rms per channel.
[10] One of them showing a double velocity component.
[11] IRAM is supported by INSU/CNRS (France), MPG (Germany), and IGN (Spain).

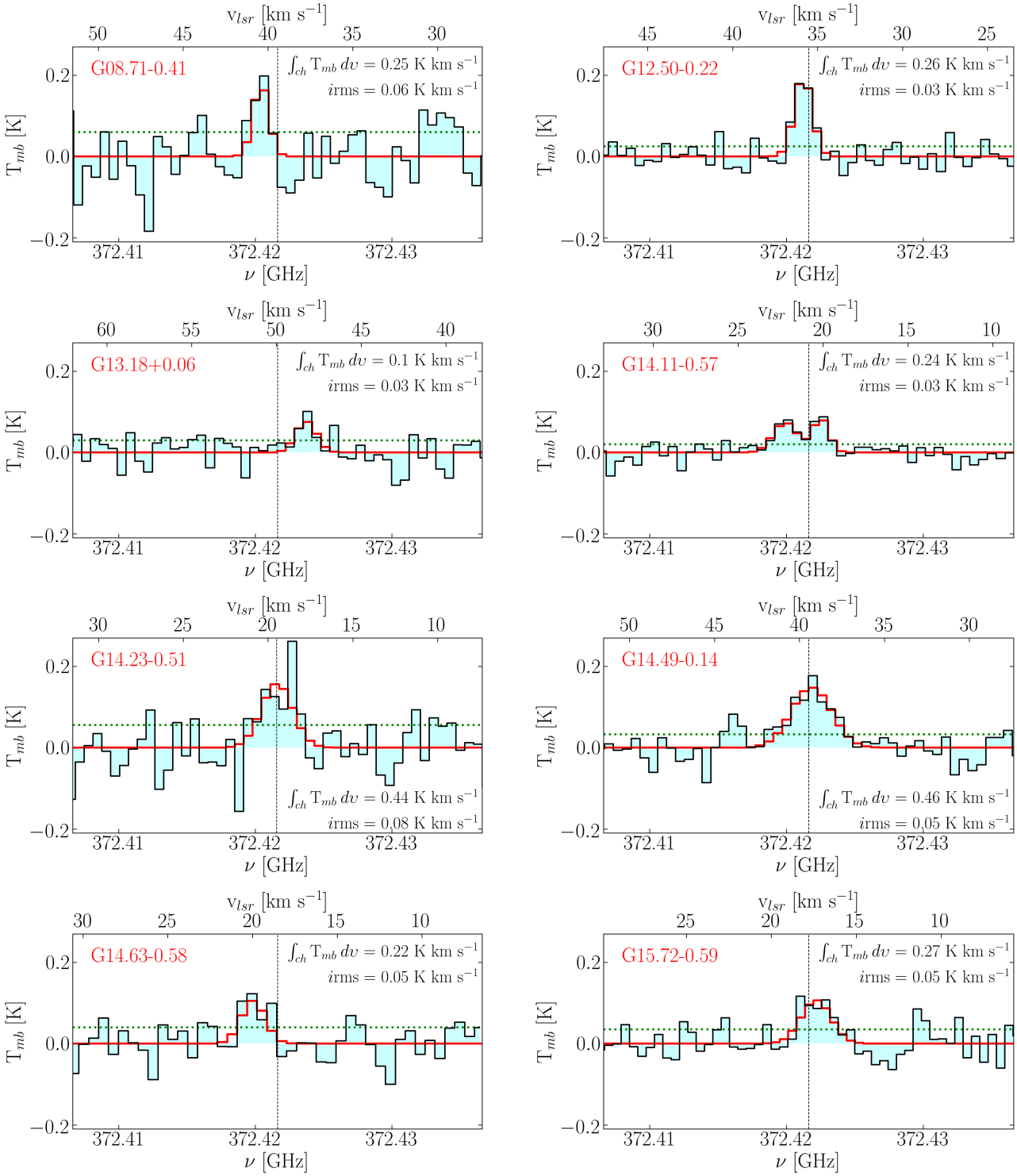

Fig. 1. o-H_2D^+ spectra for half the sources of our sample (cyan). In each panel, the source name is shown on the left in red, while the integrated main-beam temperature, $\int_{ch} T_{mb}\, dv$, and rms (*irms*) are on the right in black. The green dotted line represents the 1σ noise levels in T_{mb}. The frequency and velocity axes are reported and cut around the line, while the intensity axis is fixed to the highest T_{mb} detected (observed in G14.23–0.51). The black dashed lines indicate the v_{lsr} of each source derived from the $C^{17}O$ J = 3–2 and the NH_3 (1, 1) lines (see Table 1). The model obtained with MCWeed is superimposed in red.

4. Analysis

The spectra for all the sources with a reliable o-H_2D^+ detection are presented in Figs. 1 and 2, where the 1σ noise level is indicated by the green dotted line, and the integrated main-beam temperature and rms of each line are shown for each spectrum.

The o-H_2D^+ line emission is clearly visible with a single velocity component in all the sources, with the exception of G14.11–0.57 (see Fig. 1), which shows two line components separated by ~2.5 km s^{-1}. Both the components have iS/N $\gg$ 3, and for this reason in the following analysis they are treated separately (i.e. C1 and C2). The average noise of the spectra is almost

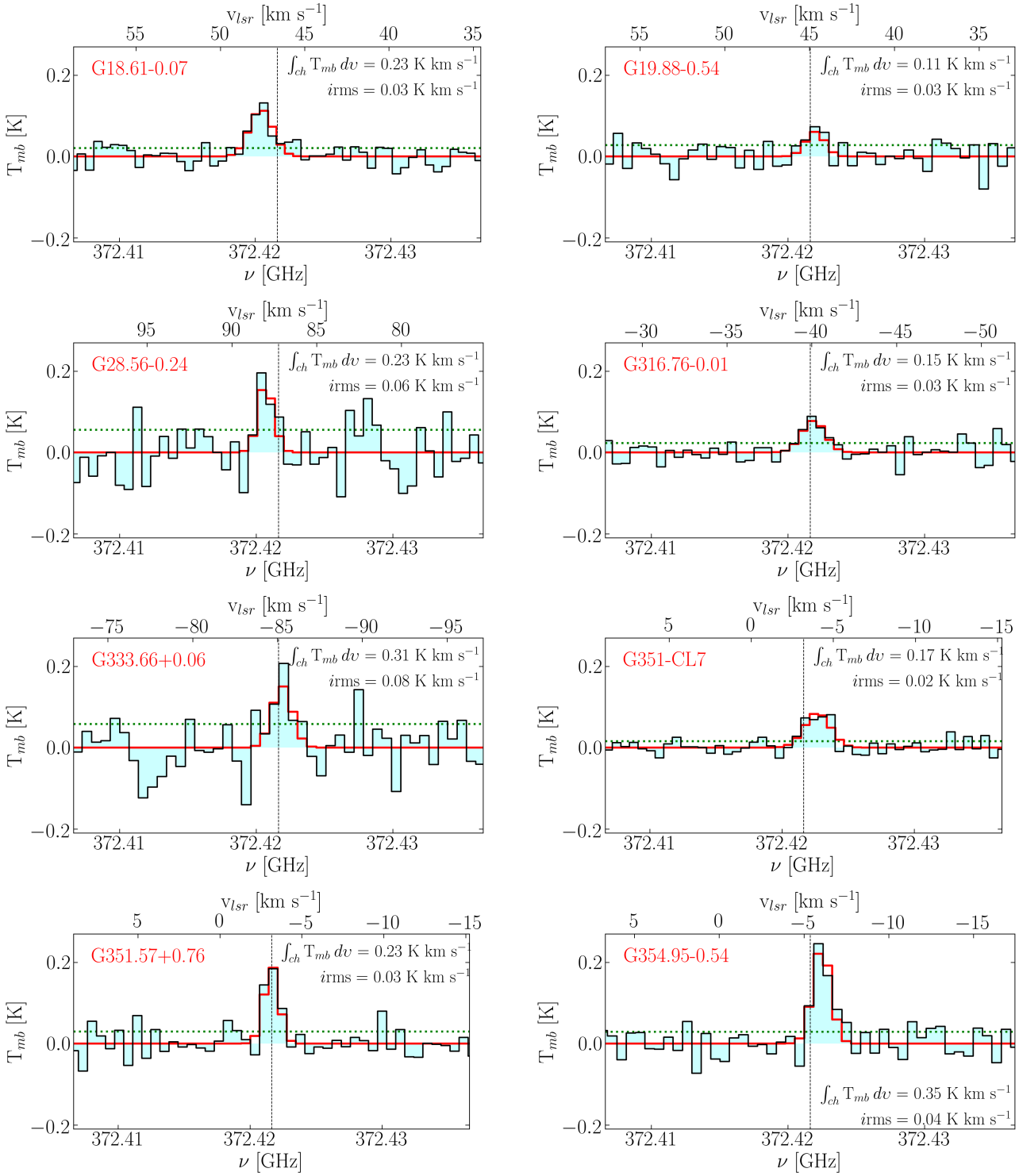

Fig. 2. Same as Fig. 1, but for the second half of the sample.

constant at ~0.02−0.03 K ($T_{\rm mb}$-scale) throughout the sample. In a few cases the noise is slightly higher (i.e. ~0.05 K). The FWHM line widths are between 1 and 3 km s^{-1}, in agreement with those reported by Giannetti et al. (2014) and Wienen et al. (2012) for the $C^{17}O$ J = 3–2 and the NH_3 (1, 1), respectively.

4.1. Column densities

Column densities are obtained by fitting the observed spectra by employing MCWeeds (Giannetti et al. 2017), an external interface between Weeds (Maret et al. 2011), simple and fast in building synthetic spectra assuming LTE, and the Bayesian statistical models and adaptation algorithms of PyMC (Patil et al. 2010). Model fits are shown in red in Figs. 1 and 2. For the 16 sources with an o-H_2D^+ detection, the results of the fit are summarised in Table 3, computed assuming that the excitation temperature of the transition $T_{\rm ex}$ = $T_{\rm dust}$ (e.g. Giannetti et al. 2019). This assumption implies that the dust and o-H_2D^+ are well mixed, which might not be the case for every source as the average volume densities (see Table 3) are slightly lower than

Table 2. Summary of the observations.

Molecules	Quantum numbers	Frequencies (GHz)	Telescopes	Beam FWHM (arcsec)	Spectral resolution (km s^{-1})	rms noise [(a)] (K)	References
o-H_2D^+	1_{10}–1_{11}	372.4	APEX	17	0.5–0.6	0.02–0.05	This work
$H^{13}CO^+$	1–0	86.8	IRAM-30 m	24	0.75	0.02	This work
DCO^+	1–0	72.0	IRAM-30 m	28	0.75	0.02	This work
$C^{17}O$ (TOP100)	1–0	112.4	IRAM-30 m	21	0.5	0.05	[1]
$C^{17}O$	1–0	112.4	IRAM-30 m	21	0.6	0.06	[2]

Notes. [(a)]The temperatures are reported on the main-beam temperature scale.

References.[1] Giannetti et al. (2014); [2] Csengeri et al. (2016).

Table 3. Summary of derived properties of the ATLASGAL-sources in our sample.

ATLASGAL-ID	L/M [(a)] $L_\odot\, M_\odot^{-1}$	$[f_D]$ [(b)] $(\chi^{obs}_{C^{17}O}/\chi^{exp}_{C^{17}O})$	N(o-H_2D^+) [(c)] $\log_{10}$(cm^{-2})	X(o-H_2D^+) $\log_{10}$(N[o-H_2D^+]/N[H_2])	$\Delta\upsilon_{obs}$ [(c)] (km s^{-1})	α [(d)]	$\mathcal{M}$ [(e)]
G08.71–0.41	0.3	8.2	$12.7^{+0.2}_{-0.2}$	–10.1 ± 0.2	$1.0^{+0.7}_{-0.7}$	0.1	1.9
G13.18+0.06	22.5	6.6	$12.2^{+0.3}_{-0.3}$	–10.6 ± 0.3	$1.4^{+1.2}_{-1.0}$	0.5	1.9
G14.11–0.57 (C1)	9.1	2.1	$12.1^{+0.3}_{-0.3}$	–10.8 ± 0.3	$1.0^{+0.7}_{-0.7}$	0.3	1.3
G14.11–0.57 (C2)	9.1	2.1	$12.3^{+0.3}_{-0.3}$	–10.6 ± 0.3	$1.6^{+1.1}_{-0.9}$	0.9	2.3
G14.49–0.14	0.4	9.0	$13.0^{+0.1}_{-0.1}$	–10.1 ± 0.1	$2.7^{+0.9}_{-0.8}$	0.6	5.4
G14.63–0.58	11.1	2.0	$12.4^{+0.2}_{-0.2}$	–10.6 ± 0.2	$1.7^{+1.0}_{-0.7}$	1.0	2.4
G18.61–0.07	0.7	4.2	$12.6^{+0.1}_{-0.1}$	–10.2 ± 0.1	$1.6^{+0.7}_{-0.6}$	0.5	3.0
G19.88–0.54	15.5	2.2	$12.1^{+0.3}_{-0.3}$	–11.0 ± 0.3	$1.3^{+1.0}_{-0.9}$	0.3	1.7
G28.56–0.24	0.3	6.3	$12.7^{+0.3}_{-0.4}$	–10.4 ± 0.4	$1.0^{+0.6}_{-0.7}$	0.1	1.9
G333.66+0.06	3.0	4.0	$12.6^{+0.3}_{-0.4}$	–10.1 ± 0.4	$1.3^{+1.0}_{-1.0}$	0.2	2.1
G351.57+0.76	2.7	2.4	$12.6^{+0.2}_{-0.2}$	–10.1 ± 0.2	$1.0^{+0.5}_{-0.4}$	0.4	1.5
G354.95–0.54	3.2	3.8	$12.7^{+0.1}_{-0.1}$	–9.9 ± 0.1	$1.2^{+0.4}_{-0.4}$	0.8	1.8
G12.50-0.22	1.3	–	$12.8^{+0.1}_{-0.1}$	–10.0 ± 0.1	$1.2^{+0.4}_{-0.3}$	1.0	2.2
G14.23–0.51	2.1	–	$12.8^{+0.2}_{-0.2}$	–10.5 ± 0.2	$2.1^{+0.8}_{-0.8}$	0.6	3.5
G15.72–0.59	0.2	–	$12.8^{+0.2}_{-0.2}$	–10.0 ± 0.2	$1.9^{+1.0}_{-1.0}$	1.0	3.8
G316.76-0.01	5.2	–	$12.3^{+0.2}_{-0.2}$	–10.8 ± 0.2	$1.6^{+0.6}_{-0.5}$	0.2	2.5
G351.77-CL7	0.7 [(f)]	3.4 [(g)]	$12.6^{+0.1}_{-0.1}$	–10.3 ± 0.1	$1.8^{+0.6}_{-0.6}$	0.5	3.5

Notes. Derived properties of our sample. Top panel: TOP100 sources; bottom panel: ATLASGAL sources not in the TOP100. The symbol "–" means or that this value is not available from the literature (for f_D) or that it could not be calculated (for G14.11–0.57). [(a)]Data from Cols. (5) and (6) in Table 1. [(b)]We associate with the CO depletion factor, f_D, a conservative error of 15%, considering recent results discussed in Sabatini et al. (2019), where we presented a detailed CO-depletion study in a local filament of massive star-forming regions. [(c)]Computed using MCweeds (Giannetti et al. 2017) assuming $T_{ex} = T_{dust}$. [(d)]Virial parameters derived following MacLaren et al. (1988), see Sect. 4.2. [(e)]Mach numbers derived following Eq. (3) in Sect. 4.2, where $m(H_2D^+) = 6.692 \times 10^{-24} g$ is the mass of the H_2D^+ molecule. [(f)]This value was calculated from Giannetti et al. (2019) as the integrated values on the APEX-beam at 230 GHz (i.e. 28 arcsec). [(g)]Data from Sabatini et al. (2019); see their Sect. 4.1 for more details on how this value is computed.

the o-H_2D^+ critical density, $n_{cr} \sim 10^5$ cm^{-3} (e.g. Caselli et al. 2008; Vastel et al. 2012). By following Caselli et al. (2008), who reported excitation temperatures up to 40% lower than the dust values, we re-performed the column density calculations taking $T_{ex} = 0.6\, T_{dust}$. We find that in most of the cases, the o-H_2D^+ column densities are well within the uncertainties reported in Table 3, so that our assumption of $T_{ex} = T_{dust}$ does not affect the final results. In this range of temperatures the line optical depths, τ, computed as in Eq. (5) of Caselli et al. (2008), are $\lesssim$0.1, implying optically thin regimes.

We also assumed extended emission relative to the FWHM APEX-beam to compute the column densities, supported by the results of Pillai et al. (2012), where the emission scale of the o-H_2D^+ was observed to be close to the typical clump size. We note that for all the clumps with a clear o-H_2D^+ detection, the angular size corresponding to the effective radii (R_{eff}; König et al. 2017; Urquhart et al. 2018) is always larger than the APEX beam sizes, except for G316.76–0.01 (Pillai et al. 2012).

Fits were performed assuming the o-H_2D^+ molecular line parameters provided by the CDMS database: an Einstein-A coefficient $\log_{10}(A_{ul}$ [s^{-1}]) = −3.96, a statistical weight of $g_{ul} = 9$, an energy gap between the two quantum levels $E_{ul} = 17.9$ K, and the partition function, $Q(T)$ for o-H_2D^+ in the relevant temperature range [9.375 K: 10.3375, 18.750 K: 12.5068, 37.500 K: 15.5054] (see also Giannetti et al. 2019).

Throughout the sample the o-H_2D^+ column density varies by less than an order of magnitude, from 1.3×10^{12} to 10^{13} cm^{-2}, while its abundance X(o-H_2D^+) = N[o-H_2D^+]/N[H_2] is in the range $\sim(1-12.6) \times 10^{-11}$. The abundances were derived with respect to the H_2 column densities reported by Giannetti et al. (2017) and Urquhart et al. (2018). The final X(o-H_2D^+) values are slightly lower than those found in many low-mass pre- and protostellar regions (e.g. Vastel et al. 2006; Harju et al. 2008; Caselli et al. 2008; Friesen et al. 2010, 2014; Miettinen 2020), but in agreement with those derived for the few high-mass star-forming regions observed by Harju et al. (2006), Pillai et al. (2012), and Giannetti et al. (2019); the only exception is the significantly lower values found by Swift (2009), who reports an abundance of $\sim(3-5) \times 10^{-13}$ in the infrared dark clouds (IRDCs) G030.88+00.13 and G028.53–00.25.

The o-H_2D^+ detection limits were calculated for sources with no detection. For each evolutionary class, we selected the observations with rms within a factor of 2 with respect to the average noise of the spectra with detection (see Appendix A and Table A.1 for details). This selection criteria leads to a loss of 44 sources. Each class has a similar number of objects (70w:8, IRw:16, IRb:9, and HII:13 sources), which makes detection limits comparable through the progressive evolution of the clumps, with a o-H_2D^+ detection rate of 47% in 70w, 27% in IRw, 18% in IRb, and 7% in HII. The detection limits were calculated for each selected source in order to obtain the column density value that would correspond to a 3σ detection. The assumptions made for the column density calculations are the same as for the sources with detection. We used the typical FWHM line width of 1.5 km s^{-1}, which corresponds to the average value of the detection sources. We note that this approach provides only qualitative information on the upper limits of column densities of o-H_2D^+ throughout the sample as it is mainly influenced by the quality of observations rather than by the physics of the sources. Nevertheless, we limit this problem by selecting spectra with noise levels comparable to those of detected sources. The (upper limit) relative abundances retrieved in this way are in the range of the typically observed values.

4.2. Estimates of dynamical quantities

In addition to the N(o-H_2D^+) value, MCWeeds also provides an estimate of the FWHMs ($\Delta\upsilon_{\rm obs}$ in Table 3) of each line, allowing us to derive two important dynamical quantities: the virial parameter, α, and the Mach number, $\mathcal{M}$. The first gives us an assessment of the dynamical state of each source, hence how far a source is from virial equilibrium (Chandrasekhar & Fermi 1953), while the second tells us about the turbulence of the gas, hence whether the clump is supported mainly by dynamical or by thermal motions.

The estimates of α were made following the classic definition of Bertoldi & McKee (1992):

$$\alpha = \frac{M_{\rm vir}}{M_{\rm clump}} \quad \text{with} \quad M_{\rm vir} = k_2 \left(\frac{R_{\rm clump}}{\rm pc}\right) \left(\frac{\Delta\upsilon_{\rm dyn}}{\rm km\ s^{-1}}\right)^2. \tag{1}$$

Here, $M_{\rm vir}$ and $M_{\rm clump}$ are the source virial mass and total mass, respectively; $R_{\rm clump}$ is the clump size in parsec, here assumed equal to the $R_{\rm eff}$; and $k_2 = 210$ for the homogeneous core assumption we made (MacLaren et al. 1988). The line width $\Delta\upsilon_{\rm dyn}$ in Eq. (1) is the combination of the thermal motions of the particle of mean mass, $\Delta\upsilon_{{\rm th},\langle m\rangle}$; the thermal gas motions, $\Delta\upsilon_{\rm th}$; and the observed FWHM. Using the values of $\Delta\upsilon_{\rm obs}$ in Table 3, we derived $\Delta\upsilon_{\rm dyn}$ by following Kauffmann et al. (2013) (see their Sect. 4):

$$\Delta\upsilon_{\rm dyn} = \sqrt{\Delta\upsilon_{\rm obs}^2 + \Delta\upsilon_{{\rm th},\langle m\rangle}^2 - \Delta\upsilon_{\rm th}^2}. \tag{2}$$

The values of α reported in Table 3 suggest that the sample in this study is heterogeneous, but mainly composed of marginal to highly unstable sources ($\alpha < 1$). We also found agreement with the results of Kauffmann et al. (2013), who discussed the implications of the low virial parameters estimated in a large sample of high-mass star-forming regions, finding that for $M_{\rm clump} \gtrsim 10\ M_\odot$, $\alpha \ll 2$ (see their Fig. 1). Such low α values might be the symptomatic consequence of a rapid collapse.

The Mach number has been computed from the o-H_2D^+ line width provided by MCWeeds as

$$\mathcal{M} = \frac{\sigma_{\rm turb}}{\sigma_{\rm H_2}}, \tag{3}$$

with

$$\sigma_{\rm turb} = \sqrt{\sigma_{\rm obs}^2 - \sigma_{\rm th}^2} \quad \text{and} \quad \sigma_{\rm H_2} = \frac{\Delta\upsilon_{{\rm th},\langle m\rangle}}{2\sqrt{2\ln 2}}. \tag{4}$$

In Eq. (4), $\sigma_{\rm obs} = \Delta\upsilon_{\rm obs}/(2\sqrt{2\ln 2})$ is the observed velocity dispersion of o-H_2D^+ lines, while $\sigma_{\rm th} = [k_{\rm B}\, T_{\rm dust}/m({\rm H_2D^+})]^{1/2}$ is the gas thermal component at $T_{\rm dust}$, where $k_{\rm B}$ is the Boltzmann constant and $m({\rm H_2D^+})$ is the mass of the H_2D^+ molecule. The Mach number range is 1.3–5.4, implying supersonic turbulent motions for the gas involved in all the sources.

5. Results

We have explored possible correlations between the o-H_2D^+ relative abundance obtained from our survey and the other physical quantities in Tables 1 and 3. In the following, we divide the discussion into three main parts: correlations with (*i*) the source parameters, to examine the possible connection between the amount of o-H_2D^+ and the quantities that characterise the clump structures and their distribution in the Galactic plane; (*ii*) the evolutionary parameters, to highlight trends during the star formation process; and (*iii*) the dynamical parameters, for possible dependencies on factors such as the concentration and the turbulence of the gas of the source.

5.1. New correlations with X(o-H_2D^+)

Figure 3 shows the linear correlations we find between the $\log_{10}X$(o-H_2D^+) and the set of physical quantities of each clump listed in Tables 1 and 3. Each linear correlation was obtained using a Bayesian approach relying on Markov chain Monte Carlo (MCMC) algorithms. While the full description of the method is left to Appendix B, where we also show an example of the posterior distributions of the model parameters, in each panel of Fig. 3 we report the best fit and the models within 3σ. In addition, to give statistical significance to the correlations, we performed a Bayesian analysis by fixing the slope of the fitting function to zero and calculating the Bayes factor $\mathcal{B}_{2,1}$ (see Table 4). According to the Jeffreys scale (Jeffreys 1961) if $\mathcal{B}_{2,1} < 1$, there is no statistical evidence to justify the use of a model with two free parameters (yellow shaded regions in Fig. 3). Alternatively, if $\mathcal{B}_{2,1} > 1$, a model with two free parameters is justified and we can consider the correlation reliable (green shaded regions in Fig. 3). We tested the same fit approach on the TOP100 subsample, always finding an agreement with the correlations shown

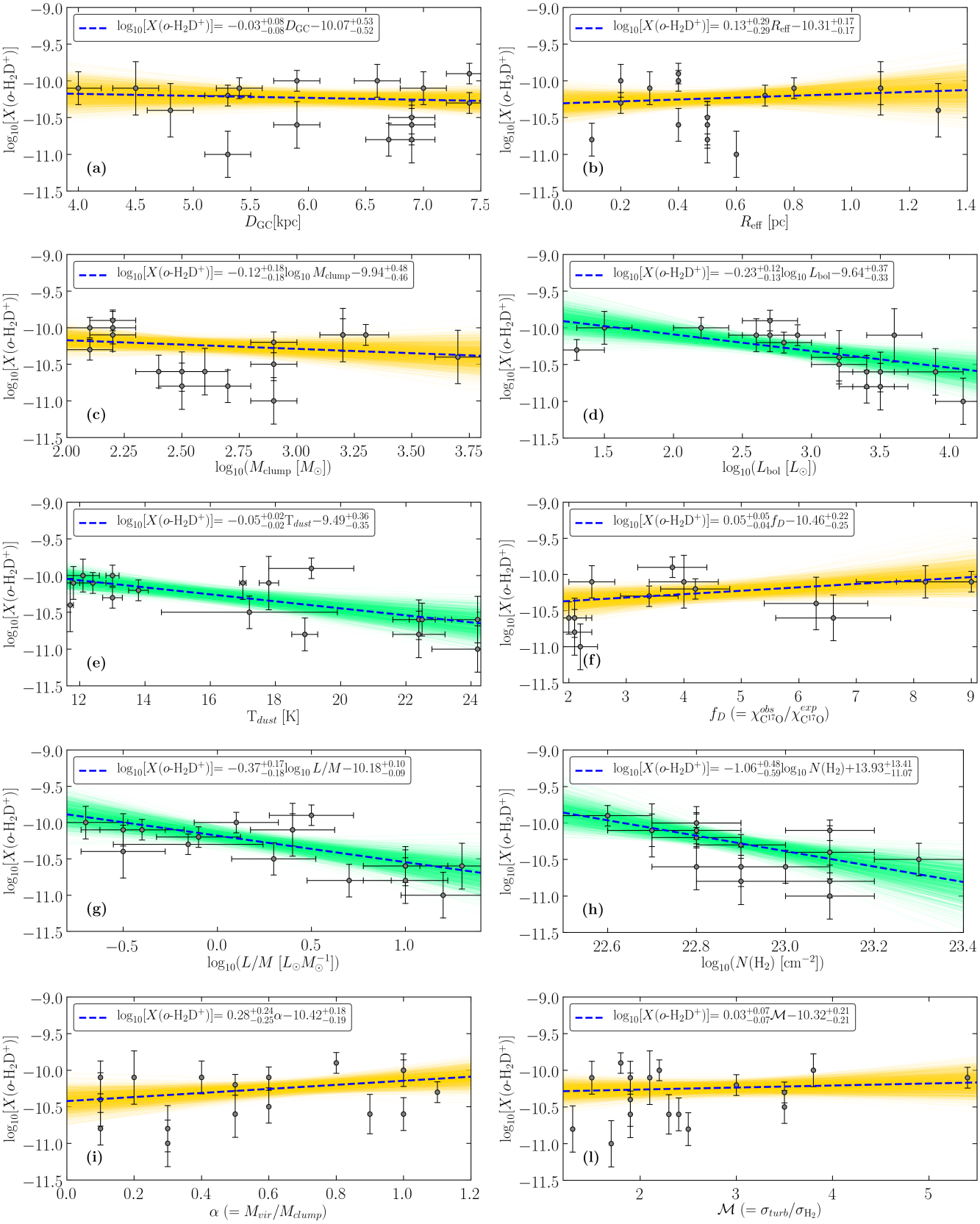

Fig. 3. Collection of the different correlations between $X(\text{o-}H_2D^+)$ and the quantities summarised in Tables 1 and 3. Grey dots are associated with each source, while uncertainties are shown as black bars. Orange and green shaded regions are the 3σ results of the MCMC linear fit respectively for the unreliable ($\mathcal{B}_{2,1} < 1$) and reliable ($\mathcal{B}_{2,1} > 1$) correlations we found for the whole sample. Blue dashed lines represent the *fiducial model* (i.e. with the highest likelihood among the ones explored by MCMC). The fit parameters are shown in the caption of each correlation. All plots are in log-linear scale, except for *panels c, d, g, and h* where the axes are set in log-log scale.

Table 4. Bayes factors, $\mathcal{B}_{2,1}$, found for each correlation in Fig. 3.

Panel	Correlation with	Bayes factors
(a)	D_{GC}	5.5×10^{-5}
(b)	R_{eff}	1.0×10^{-3}
(c)	$\log_{10}(M_{clump})$	1.0×10^{-3}
(d)	$\log_{10}(L_{bol})$	9.7
(e)	$\log_{10}(T_{dust})$	4.5
(f)	f_D	4.1×10^{-4}
(g)	$\log_{10}(L/M)$	51.1
(h)	$\log_{10}[N(H_2)]$	1.9×10^{4}
(i)	α	2.4×10^{-1}
(l)	$\mathcal{M}$	4.5×10^{-5}

Notes. The correlations are computed with respect to the abundance of o-H_2D^+ derived from the column densities in Tables 1 and 3 (i.e. N[o-H_2D^+]/N[H_2]).

in Fig. 3. This result excludes possible bias depending on the sample selection or on the procedures followed to derive the other values reported in Tables 1 and 3.

In the following paragraphs we discuss each group of correlations, paying particular attention to whether the dependence of each pair can be physically explained or if it is simply fortuitous.

5.1.1. Source parameters

The dependence of X(o-H_2D^+) on the physical properties of clumps does not seem particularly pronounced. In the first three plots in Fig. 3 (panels a, b, and c), X(o-H_2D^+) appears roughly uniform as the other quantities vary. In these cases the correlations we found are in agreement with a purely flat regime, with small fluctuations of the fiducial models slopes, probably caused by regions with sparse density of observed points.

Variations in the deuterium abundance with respect to hydrogen, [D/H], through the Galactic disk are expected considering that all deuterium atoms were formed at the birth of the Universe and then progressively consumed by thermonuclear reactions within the stellar cores (e.g. Lubowich et al. 2000; Ceccarelli et al. 2014; see also Galli & Palla 2013 for a review). Following this scenario, [D/H] should vary with location and the same would be expected for deuterated species. On the other hand, it is well known that the major boost in deuterium fractionation is driven by the star formation process itself: the temperatures, the global chemistry, and energetics of the stars can play a central role, (dis-)favouring the deuteration process (e.g. Caselli et al. 2002; Bacmann et al. 2003; Lis et al. 2006; Pillai et al. 2012; Ceccarelli et al. 2014; Giannetti et al. 2019). The last dependence seems particularly evident in panel d of Fig. 3, which shows that X(o-H_2D^+) clearly decreases as L_{bol} rises (i.e. $\mathcal{B}_{2,1} \gg 1$). The o-H_2D^+ abundance is found to change by ~0.5 orders of magnitude for a factor of ~4 change in $\log_{10}(L_{bol})$, with a fiducial model's slope of ~−0.23.

5.1.2. Evolutionary parameters

Among the quantities we defined as evolutionary tracers we have included the following: the CO-depletion factor, f_D, as the indicator of the depletion degree of each source, and defined as the ratio of the expected CO abundance with respect to H_2 to the observed value (see Caselli et al. 1999; Fontani et al. 2012); the dust temperature, which is determined by the activity of the protostars that are forming in each source; and L/M as the main indicator for the evolutionary stage of star formation processes in the high-mass regime (e.g. Saraceno et al. 1996; Molinari et al. 2008; Körtgen et al. 2017; Urquhart et al. 2018; Giannetti et al. 2019). The values for each clump are summarised in Tables 1 and 3. Depletion factors are derived from Giannetti et al. (2014) and Sabatini et al. (2019) using the $C^{17}O$ column density distributions generated by taking into account optical depth effects, while T_{dust} and L/M values are taken from Giannetti et al. (2017, 2019), König et al. (2017), and Urquhart et al. (2018), to which we refer for a more detailed discussion of how these quantities were estimated. Depletion factors are also the only incomplete values due to the lack of observational data for the sources not in the TOP100. Therefore, we note that panel f of Fig. 3 is the only one obtained with a reduced number of data points (i.e. we could not include G12.50–0.22, G14.23–0.51, G15.72–0.59, G316.76–0.01).

In panel e of Fig. 3, X(o-H_2D^+) correlates with T_{dust} with small slopes, changing by ~0.5 orders of magnitude in the range of T_{dust} between 12 and 24 K. The fit with the model with two free parameters is statistically supported by a $\mathcal{B}_{2,1} > 1$. However, we note that the scale chosen for the plot and the different ranges covered by the quantities involved may confuse the interpretation of these correlations, leading one to conclude that small slopes could mean soft correlation.

The correlation between X(o-H_2D^+) and f_D in Fig. 3 f appears similar to that found for T_{dust} (i.e. a clear correlation with a small slope). However, we find that $\mathcal{B}_{2,1} < 1$, which excludes any statistical evidence to justify the use of a model with two free parameters to establish a correlation between X(o-H_2D^+) and f_D. We note in this case that the reduced number of sources may complicate the fit with two free parameters, while a correlation between X(o-H_2D^+) and f_D is expected since the highly CO-depleted environments are those in which deuteration is favoured (e.g. Dalgarno & Lepp 1984; Caselli & Ceccarelli 2012). The log-lin scale in Fig. 3e and f was chosen to allow direct comparison with Caselli et al. (2008), where a survey of the o-H_2D^+ towards a sample of 16 low-mass sources – between starless and protostellar cores – was presented. This study gives us the opportunity to test the connection between low- and high-mass star-forming regions, with two samples statistically comparable. Caselli et al. (2008) found a progressively decreasing trend of X(o-H_2D^+) at increasing temperatures. Here, we find the same behaviour even if less pronounced, but supported by an inverse correlation with f_D (even if not statistically confirmed), also in agreement with Vastel et al. (2006).

We also find a strong anti-correlation (i.e. $\mathcal{B}_{2,1} \gg 1$) of X(o-H_2D^+) with L/M; X(o-H_2D^+) is found to decrease by one order of magnitude for one order of magnitude change in L/M, with a slope of −0.37. This confirms the hypothesis outlined by Giannetti et al. (2019) that o-H_2D^+ can be considered a good chemical clock during the star formation process (see Sect. 6.1 for a detailed discussion).

5.1.3. Dynamical parameters

We analyse here the possible correlations between X(o-H_2D^+) and the total H_2 column density, the virial parameter, and the Mach number, which we consider quantities related to the dynamical stage of the clumps. The correlation we find between X(o-H_2D^+) and N(H_2), shown in Fig. 3 h, suggests that an increased density could also play a relevant role in the deuteration process. The o-H_2D^+ abundance is found to change by one order of magnitude for less than one order of magnitude change

Table 5. Summary of the quantities to calculate the CRIR.

ATLASGAL-ID[(a)]	$N(C^{17}O)$ [(a)]	$\chi(CO)$ [(a)]	$N(H^{13}CO^+)$	$N(DCO^+)$	R_D	ζ_2^A $\ell = 0.5R_{eff}$	ζ_2^{ref} $\ell = R_{eff}$	ζ_2^B $\ell = 2R_{eff}$
	$\log_{10}(cm^{-2})$	(10^{-5})	$\log_{10}(cm^{-2})$	$\log_{10}(cm^{-2})$		$(10^{-17}\ (s^{-1}))$	$(10^{-17}\ (s^{-1}))$	$(10^{-17}\ (s^{-1}))$
G13.18+0.06	15.5	8.0	13.6 ± 0.1	12.4 ± 0.2	0.002 ± 0.001	6.92	3.46	1.73
G14.11–0.57	15.7	11.5	13.8 ± 0.1	13.6 ± 0.1	0.011 ± 0.004	3.18	1.59	0.80
G14.23–0.51	**15.7**	**4.6**	14.0 ± 0.1	13.8 ± 0.1	0.011 ± 0.004	3.67	1.84	0.92
G14.49–0.14	15.6	4.7	13.3 ± 0.1	13.0 ± 0.1	0.011 ± 0.003	6.68	3.34	1.67
G14.63–0.58	15.5	5.8	13.6 ± 0.1	12.8 ± 0.1	0.003 ± 0.001	11.62	5.81	2.91
G15.72–0.59	**15.6**	**11.1**	13.2 ± 0.1	13.2 ± 0.1	0.018 ± 0.006	3.06	1.53	0.77
G18.61–0.07	15.3	4.6	13.5 ± 0.1	12.9 ± 0.1	0.005 ± 0.002	1.33	0.67	0.34
G19.88–0.54	15.7	5.8	13.8 ± 0.1	13.2 ± 0.1	0.005 ± 0.002	3.00	1.50	0.75

Notes. ζ_2 values are calculated following the recent results of Bovino et al. (2020) and assuming three values of ℓ in Eq. (5). [(a)]Data and calculations from Giannetti et al. (2014). We assumed a constant $[^{18}O]/[^{17}O] = 4.16$, as in Wouterloot et al. (2008). Values in boldface were taken from Csengeri et al. (2016).

in $\log_{10}[N(H_2)]$, with a fiducial model slope of ∼−1.06. Frequent collisions are expected in denser environments (i.e. fast chemical kinetics) that would accelerate the formation of o-H_2D^+. On the other hand, when moving to a denser environment, the conversion of H_2D^+ in heavier isotopologues (D_2H^+ and D_3^+) or the deuterium transfer from H_2D^+ to other chemical species (e.g. N_2 and CO) can also be boosted, as suggested by Giannetti et al. (2019).

The last two panels of Fig. 3, i and l, present the correlations with α and $\mathcal{M}$ calculated from the o-H_2D^+ detected lines, as described in Sect. 4.2. Although the virial parameter and the Mach number depend on quantities for which a correlation with X(o-H_2D^+) was established, such as the line FWHM, on which both M_{vir} and σ_v depend, or the dust temperature, we do not find any particularly relevant trend with X(o-H_2D^+). We find that X(o-H_2D^+) does not show a significant variation with α and $\mathcal{M}$, which means that the slopes are in agreement (within the uncertainties) with zero in these log-linear plots.

To summarise, we find strong correlations between X(o-H_2D^+) and different physical quantities of the clumps, some expected and others less trivial. In Fig. 3 the correlations with $\mathcal{B}_{2,1} > 1$ are highlighted in green. To avoid any possible bias we repeated the MCMC fit approach looking for correlations with the o-H_2D^+ column densities, and we found that only the correlation between N(o-H_2D^+) and N(H_2) is not confirmed. All the other parameters in which a correlation is confirmed for both X(o-H_2D^+) and N(o-H_2D^+) are more or less directly connected to the evolution of the clumps. We conclude that the deuterium fractionation seems to be driven more by a temporal evolution than by a dynamical evolution of the star-forming regions, once the minimum physical conditions required to boost deuteration are fulfilled (i.e. cold, dense, and highly CO-depleted regions).

5.2. Estimates of the CRIR

Recently, a new way to estimate the cosmic-ray ionisation rate (CRIR) of hydrogen molecules, ζ_2, was presented by Bovino et al. (2020), based on H_2D^+ and other H_3^+ isotopologues. Our survey provides the opportunity to test, for the first time, this new method on a large sample of massive star-forming regions with simultaneous detections of o-H_2D^+, DCO^+, $H^{13}CO^+$, and $C^{17}O$. According to Bovino et al. (2020) the CRIR can be obtained from

$$\zeta_2 = \bar{\alpha}\, k_{CO}^{H_3^+}\, \frac{\chi(CO)\, N(H_3^+)}{\ell}, \tag{5}$$

where $\bar{\alpha} = 0.77$ is the correction factor that takes into account the errors due to approximations in the method (see Bovino et al. 2020, for details); $k_{CO}^{H_3^+}$ is the rate at which CO destroys H_3^+; ℓ is the path length over which the column densities are estimated; and χ(CO) is the relative abundance of CO with respect to H_2, which we derive by the simple formula:

$$\chi(CO) = \phi_{18/17}\, \frac{N(C^{17}O)}{N(H_2)}\, \frac{[^{16}O]}{[^{18}O]}. \tag{6}$$

Here $\phi_{18/17}$ is the isotopic ratio of $[^{18}O]/[^{17}O] = 4.16$ adopted from Wouterloot et al. (2008). For the TOP100 sources, the $C^{17}O$ column densities were taken from Giannetti et al. (2014), while for ATLASGAL sources not contained in the TOP100 we used the data described by Csengeri et al. (2016) (bold values in Table 5).

In Eq. (5), $N(H_3^+)$ is the column density of H_3^+ derived from the o-H_2D^+ and the HCO^+ deuterium fractionation, R_D, that we calculate from the DCO^+ and $H^{13}CO^+$ column densities reported in Table 5. Column densities were derived consistently with the procedure employed for o-H_2D^+ (i.e. using MCWeeds and by considering how the $^{12}C/^{13}C$ ratio varies as a function of D_{GC}; see discussion in Giannetti et al. 2014): $[^{12}C]/[^{13}C] = 6.1^{+1.1}_{-1.8}\, D_{GC} + 14.3^{+7.7}_{-7.2}$. Since the real size of the region associated with the o-H_2D^+ emission is unknown, we assume three multiples of R_{eff} as representative cases to calculate ζ_2: (a) $\ell = 0.5R_{eff}$, which approximately covers a range of ℓ comparable with the mean core size associated with the o-H_2D^+ emission reported by Pillai et al. (2012) (i.e. ∼0.2 pc); (b) $\ell = R_{eff}$ and (c) $\ell = 2R_{eff}$, which implicitly means that the o-H_2D^+ is emitting on the full clump size defined at 870 μm.

Our estimates of ζ_2 are summarised in Table 5. In Fig. 4, we present these results arranged, from left to right, according to the evolutionary stage of each source. The blue line represents the mean value of $\zeta_2 = 2.5 \times 10^{-17}\ s^{-1}$ for $\ell = R_{eff}$, while the blue shaded area is its variation with ℓ. It is worth noting that the range of values spans more or less an order of magnitude, which is in line with typical errors reported from models (see e.g. Padovani et al. 2018 and reference therein). We compare this result with recent estimates from Ivlev et al. (2019), who derive an upper limit of $\zeta_2 \sim 10^{-16}\ s^{-1}$ from a self-consistent model for the equilibrium gas temperature and size-dependent dust temperature in the prestellar core L1544. A second comparison is made with the estimate obtained by van der Tak & van Dishoeck (2000) in a sample of seven young massive stars using models

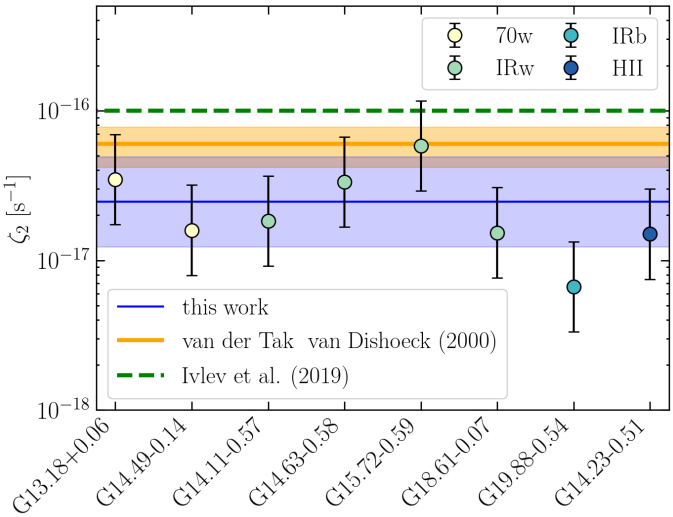

Fig. 4. Estimates of the CRIR of hydrogen molecules, ζ_2, derived for a sub-sample of sources for which H_2D^+, DCO^+, $H^{13}CO^+$, and $C^{17}O$ observations are available, and assuming $\ell = R_{\rm eff}$ in Eq. (5), representing our reference case (dots). The sources are arranged from left to right following their evolutionary phase (bottom left legend). The bars associated with each point represent the variability of our results with ℓ (i.e. $\ell = 0.5R_{\rm eff}$ and $\ell = 2R_{\rm eff}$ for the upper and lower limit, respectively). The blue line is our mean value for ζ_2 for the reference case, while the blue shaded area is its variability with ℓ. The orange line and shaded area represents the estimates of van der Tak & van Dishoeck (2000), obtained by converting the CRIR for hydrogen atom, $\zeta_{\rm H}$, via $\zeta_2 = 2.3 \times \zeta_{\rm H}$. The green line is the recent upper limit for L1544 reported by Ivlev et al. (2019).

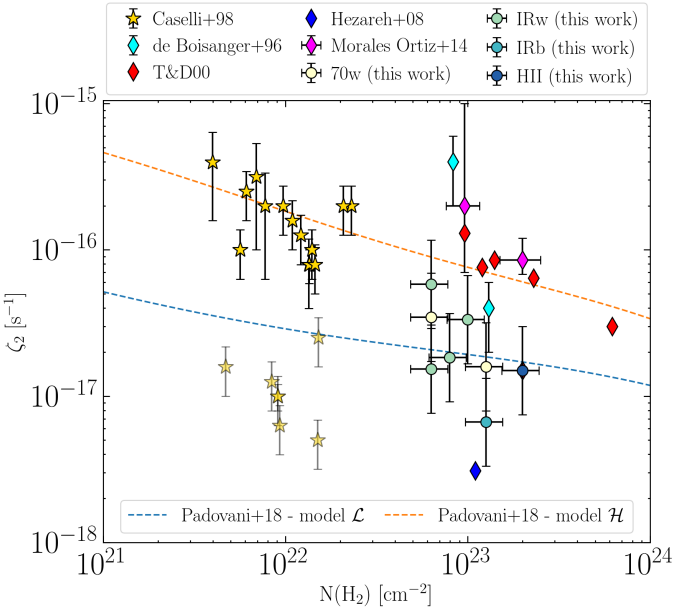

Fig. 5. ζ_2 variation as a function of N(H_2). Circles refer to our estimates of ζ_2 in high-mass star-forming regions in different evolutionary stages (same colour-coding as in Fig. 4). Yellow stars show the estimates of Caselli et al. (1998) for a large sample of low-mass cores, while diamonds represent high-mass cores (cyan from de Boisanger et al. 1996; red from van der Tak & van Dishoeck 2000 (T&D00) for ζ_2 estimates and from van der Tak et al. 2000 and Doty et al. 2002 for N(H_2) values; blue from Hezareh et al. 2008; magenta from Morales Ortiz et al. 2014). Dashed lines show the models discussed in Padovani et al. (2018) assuming different slopes for the CR proton spectrum. Different levels of transparency have been used to separate what we considered outliers (pale color) from the other estimates (full colour). This plot was readapted from Padovani et al. (2009).

based on $H^{13}CO^+$ observations. They found an average value of $\zeta_2 \sim (6.0 \pm 1.8) \times 10^{-17}$ s^{-1}. Looking at our reference case, most of the sources show values lower than the average obtained by van der Tak & van Dishoeck (2000).

The analysis we report has the advantage of being model-independent, but it is purely qualitative, first because the method proposed in Bovino et al. (2020) includes strong approximations, and second because its validity has been shown for very small regions (i.e. $R = \ell/2 \leq 0.05$ pc, where R is the distance from the centre of the clump). However, as the error here is driven by the uncertainties in the column density calculations, we can consider the CRIR estimates valid and robust within the variability shown by the different ℓ. In addition, this represents the only viable and model-independent way to estimate the CRIR in dense regions.

Figure 5 shows how the cosmic-ray ionisation rate varies as a function of N(H_2) for a sample of low- and high-mass star-forming regions. Considering the associated uncertainties, the data cover a range of ζ_2 between $\sim 3 \times 10^{-18}$ and $\sim 10^{-15}$ s^{-1}, while the molecular hydrogen column densities are in the range $\sim(0.4–80) \times 10^{22}$ cm^{-2}. The yellow stars are the ζ_2 values computed in low-mass cores from Caselli et al. (1998), while the values for N(H_2) are taken from Butner et al. (1995). Diamonds indicate massive protostellar envelopes (see the caption). The circles represent the estimates obtained in this work for the reference ℓ (same colour-coding as in Fig. 4). The dashed lines are the models discussed in Padovani et al. (2018), considering a single cosmic-ray (CR) electron spectrum and two different CR proton spectra, namely model $\mathcal{L}$ (in blue) and model $\mathcal{H}$ (in orange), and depending on the slope of the CR proton spectrum (see also Ivlev et al. 2015 for more details about the CR spectrum model).

The distribution of ζ_2 in Fig. 5, and in particular that of the full-coloured data points, seems to suggest the general trend where the CRIR decreases while the H_2 column density increases. We note that the estimates of ζ_2 in high-mass star-forming regions are smaller by a factor of ~3–4 with respect to their low-mass counterparts, which is in agreement with the fact that IRDCs are usually much denser compared to the low-mass regime. Our estimates fall between the two models of Padovani et al. (2018), while a small sub-sample of low-mass cores (~1/3 of the total sample) and a few high-mass star-forming regions find agreement with the model with lower CR proton energy. The apparent bimodality (yellow and pale yellow stars in Fig. 5) in the low-mass regime can be explained by the different $f_{\rm D}$ values assumed by Caselli et al. (1998) to estimate ζ_2. Here, by following Padovani et al. (2009), we report the best ζ_2 estimates from Caselli et al. (1998), obtained from HC_3N/CO data assuming $f_{\rm D}$ = 3–5. In particular, values of ζ_2 around $\sim 10^{-16}$ s^{-1} are produced by both of the $f_{\rm D}$ values, while those around $\sim 10^{-17}$ s^{-1} correspond to $f_{\rm D} = 3$. On the other hand, the morphology of the magnetic field lines can also play a major role in determining the CRIR. Padovani & Galli (2011) and Padovani et al. (2013) have shown that even in high-density regions, if the magnetic field lines are very concentrated around the accreting protostar, ζ_2 can be reduced by up to a factor of 10 or more.

Finally, in Fig. 6 we report the cosmic ray ionisation rate as a function of the galactocentric distances for different high-mass star-forming clumps. We found that there is no significant variation of ζ_2 with $D_{\rm CG}$ for distances between 4.7 and 11.2 kpc. This result is in agreement with those of Indriolo et al. (2015), suggesting that the CRIR is uniform for sources located at distances

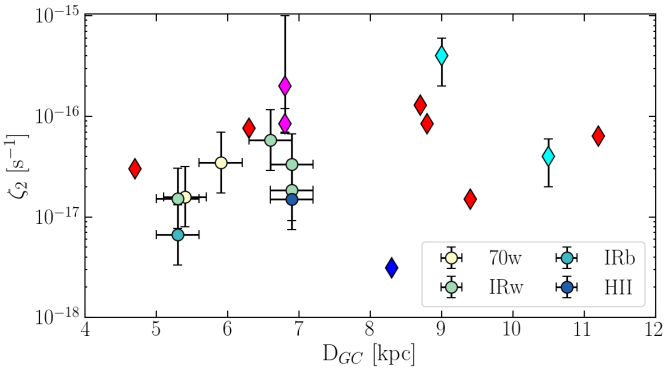

Fig. 6. ζ_2 variation as a function of D_{GC} for the high-mass cores shown in Fig. 5. Circles refer to our estimates of ζ_2 in high-mass star-forming regions in different evolutionary stages (see legend). The same colour-coding as in Fig. 5 is assumed for diamonds. The galactocentric radii are taken from Table 1 (circles), van der Tak & van Dishoeck (2000) (red), Morales Ortiz et al. 2014 (magenta), Winkel et al. 2017 (blue), Spina et al. 2017, and Yan et al. 2019 (cyan). D_{CG} are scaled assuming a distance to the Galactic Center of 8.35 kpc.

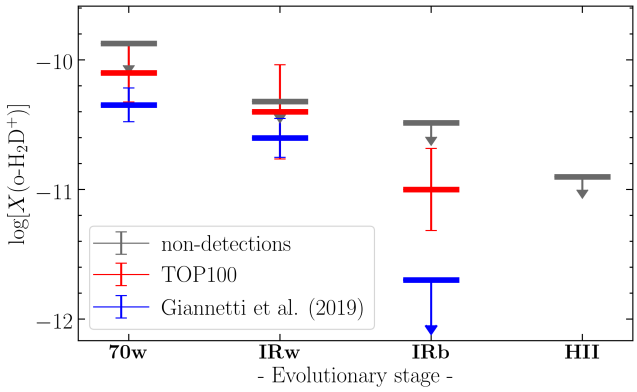

Fig. 7. Average X(o-H_2D^+) abundance as a function of the evolutionary classes in our sample. Blue markers indicate the recent estimates of Giannetti et al. (2019), while the red and grey ones indicate the mean of detection and the median X(o-H_2D^+) derived from detection limits, respectively. Uncertainties are derived as mean and median as well, considering the errors on the individual source.

above 5 kpc from the Galactic centre. Therefore, the observed spread in Fig. 5 could be the signature of local effects, for instance different magnetic field topology of each source.

6. Discussion

6.1. o-H_2D^+ as an evolutionary tracer for massive clumps

Figure 7 shows the abundance of o-H_2D^+ as a function of the evolutionary class of the TOP100 clumps (red markers and error-bars). To mitigate the possible bias caused by extreme sources with peculiar initial conditions, and to make our result as general as possible, we report the median X(o-H_2D^+) values for each evolutionary class. X(o-H_2D^+) varies by more than one order of magnitude between the least evolved and the most advanced stage of evolution, suggesting that the deuterium fractionation of H_3^+ is favoured during the initial phases of star formation, while its chemical products, in this case o-H_2D^+, slowly disappear as massive clumps evolve. We repeated this procedure for the detection limits discussed in Sect. 4 to exclude possible sensitivity effects. The results are reported in Fig. 7 as grey markers. It is clear that the trend mentioned above holds: o-H_2D^+ decreases as the evolution progresses.

Between the two solutions, we found a discrepancy (always in agreement within the error bars) of at most a factor of 6, which is expected if we consider the different noise levels of the spectra. The trend found in Fig. 7 is a different manifestation of the result already shown in panel g of Fig. 3, where the evolution of the clumps is indicated by a progressive increase of the luminosity-to-mass ratio (e.g. Saraceno et al. 1996; Molinari et al. 2008) and X(o-H_2D^+) shows the same decreasing behaviour passing from the lowest to the highest L/M values (i.e. from less to more evolved sources). The same result also emerges from Fig. 3d since the bolometric luminosity traces the evolution of clumps if we assume that during the star formation process the clump mass stays constant. This latter assumption seems reliable as X(o-H_2D^+) correlates with all the other evolutionary indicators, while no correlation was found between X(o-H_2D^+) and $M_{\rm clump}$ (Fig. 3c), in agreement with the results of König et al. (2017) and Urquhart et al. (2018), who find no trend in the clump mass between the evolutionary classes.

We compared the results found in Fig. 7 with those of Giannetti et al. (2019), where the same clear downward trend was observed in the X(o-H_2D^+) of three massive clumps associated with different evolutionary stages and harboured in the IRDC G351.77–0.51. Although less pronounced, similar results were recently reported by Miettinen (2020) for three prestellar and three protostellar low-mass cores in Orion-B9. In the less evolved sources with detection (i.e. 70w and IRw), the agreement with Giannetti et al. (2019) is within the associated error limits, while comparing our IRb detection with the upper limit provided by Giannetti et al. (2019), we found X(o-H_2D^+) higher by a factor of ~5, which might depend on different initial conditions (both physical and chemical). These results confirm that through the evolutionary sequence of massive star-forming regions a general downward trend for o-H_2D^+ is clearly observable with upper limits associated with the HII class. The same trend is also visible if the average abundances reported by Miettinen (2020) are considered.

Figure 7 suggests that the observed X(o-H_2D^+) trend is not influenced by the nature of the samples as a similar behaviour has been reported both within the same star formation complex (Giannetti et al. 2019) and within a heterogeneous sample (this work). Our new confirmation of the same trend is completely independent from the initial chemical conditions of the ISM in which the star-forming complexes were formed and from any particular or stochastic episodes during the star formation process (i.e. possible outflows or different mass accretion rates).

6.2. o-H_2D^+ in a broader scenario of star formation

The fact that o-H_2D^+ can be considered a valid chemical clock for the star formation process is a consequence of how strongly its chemistry is affected by changes in density and temperature during cloud contraction. The star formation process is conventionally assumed to start in starless or prestellar cores, that are dense, $n(H_2)$ ~10^4–10^5 cm^{-3}, and cold, T < 20 K, clouds of gas and dust in which gravitational collapse has not yet necessarily started (e.g. Bergin & Tafalla 2007). In these dense regions, dust extinction becomes higher than ~10 mag, preventing molecules from being photodissociated or photoionised by UV photons, and the only way to form positive ions is through cosmic ray-induced chemical processes. An example of this process is the formation of H_3^+, the molecular precursor of H_2D^+, produced

by the interactions of a CR particle (*c.r.*) that ionises an H_2 molecule, and followed by a fast charge exchange reaction with H_2:

$$H_2 + c.r. \rightarrow H_2^+ + e^-, \quad (7)$$

$$H_2^+ + H_2 \rightarrow H_3^+ + H. \quad (8)$$

The efficiency of H_3^+ production in reaction (8) only depends on the CRIR. The deuterium fractionation can then continue with the formation of H_2D^+ via the proton-transfer reaction:

$$H_3^+ + HD \rightleftharpoons H_2D^+ + H_2 + \Delta E. \quad (9)$$

This reaction is exothermic in the forward direction unless there is a substantial fraction of o-H_2 (e.g. Gerlich et al. 2002), with a released energy ΔE that depends on the H_3^+ and H_2D^+ isomer states involved (e.g. Hugo et al. 2007). In addition, for temperatures below 20 K, the reaction tends to favour the formation of the ortho state of H_2D^+ compared to its para form, as demonstrated by Pagani et al. (1992a). The thermal state of the clump represents a crucial parameter to understand the evolution of o-H_2D^+ along the different evolutionary classes.

According to Urquhart et al. (2018), each evolutionary class of the ATLASGAL survey shows a clear separation in temperature that increases from ~10 to ~40 K, in line with the expected evolutionary sequence. The same behaviour was also noted in the TOP100, a further confirmation of the statistical relevance of this flux-limited sub-sample as representative of the whole ATLASGAL (e.g. König et al. 2017). Urquhart et al. (2018) also pointed to positive strong correlations that link the gas temperature of the clumps to their embedded massive protostar luminosity or to the luminosity-to-mass ratio of each clump. Our results resemble a similar evolutionary behaviour in the X(o-H_2D^+) -T_{dust} correlation reported in panel e of Fig. 3, where it is visible that higher values of X(o-H_2D^+) are reached in colder environments, while as T_{dust} increases, X(o-H_2D^+) become progressively lower. The temperature is also directly responsible for the evolution of CO in these regions. Low concentrations of CO molecules in the gas phase have been observed in cold environments (e.g. Roberts & Millar 2000; Bacmann et al. 2003; Ceccarelli et al. 2014), and correlations of the CO abundance with the evolutionary stages have been reported in the low-mass regime (e.g. Caselli et al. 1998; Bacmann et al. 2002). The same behaviour was also confirmed in high-mass star-forming regions, where f_D was found to decrease with the increase of L/M, as the indirect confirmation that the clumps become warmer with time (see Giannetti et al. 2014, 2017).

In highly CO-depleted environments the formation of H_2D^+ is made more efficient by two events. On the one hand, the H_3^+ reservoir is kept available for reaction (9) as the following proton-transfer reaction is slowed down:

$$H_3^+ + CO \rightarrow HCO^+ + H_2. \quad (10)$$

On the other hand, due to the same effect, H_2D^+ is prevented from reacting with CO and, to a lesser extent, with N_2 (which normally depletes slowly) to form DCO^+ and N_2D^+, remaining abundant and observable in gas phase. The f_D-X(o-H_2D^+) correlation shown in Fig. 3f, qualitatively supports this scenario. We note that in this case the highest X(o-H_2D^+) correspond to the highest f_D values, with an opposite trend respect to what is found for T_{dust}. This anti-correlation is expected as cold (and dense) environments are those in which CO-depletion is boosted (e.g. Kramer et al. 1999; Caselli et al. 1999; Crapsi et al. 2005; Fontani et al. 2012; Wiles et al. 2016; Sabatini et al. 2019).

It is also clear, from the obtained results, that the answer to how (and if) the observed X(o-H_2D^+) are linked to the clumps' dynamical picture across their evolution is far from being exhaustive. In agreement with the results of König et al. (2017), we do not find any evident trend of α and $\mathcal{M}$ with the variation of the evolutionary stage in the TOP100 sources. We only note that for the entire sample α is found to be $\lesssim$1, which indicates dynamically unstable sources and very fast collapses (see Kauffmann et al. 2013 for a more detailed description). Interestingly, similar results have been obtained by numerical simulations (Körtgen et al. 2017, 2018) in which the influence of dynamical parameters like the Mach number ($\mathcal{M}$), the magnetic field magnitude and distribution, and the cloud gas surface density have been shown to affect the results much less than the time evolution of the clumps. This prevents us from determining whether the correlations found in panels i and l of Fig. 3 support the scenario in which X(o-H_2D^+) and the clumps' dynamical quantities are not empirically correlated, and the possibility that our sample may be affected by some type of selection bias. However, the latter is not supported by the other correlations we report, while it is more likely that the link between the emitting region associated with o-H_2D^+ and the clump dynamics is not so trivial. In addition, we should also consider that the H_2D^+ velocity dispersion cannot give us full information on the clump dynamics because close to the centre of the clumps it is quickly replaced by N_2D^+ or D_2H^+ or destroyed by the presence of freshly desorbed CO. High spectral resolution observations are necessary in order to estimate with sufficient accuracy the line FWHMs, and consequently α and $\mathcal{M}$, together with high spatial resolution (<0.2 pc; e.g. Pillai et al. 2012), to be able to distinguish the dynamics of the sub-clumps associated with the o-H_2D^+ emission.

Overall, our results suggest that the o-H_2D^+ evolution follows the physical conditions associated with the different evolutionary stages: high abundances in the cold starless or prestellar stages (70w and IRw), and a clear decrease while the evolution of the protostellar object proceeds from its young phases to more evolved situations, including HII regions. This is in particular confirmed by the correlations reported in Fig. 3, which support the scenario in which deuteration proceeds faster in the first stages favoured by a high degree of CO freeze-out and drops at later stages mainly due to the increase in temperature induced by the presence of a luminous YSO and the subsequent release of CO back into gas-phase.

7. Summary and conclusions

With the aim of confirming that o-H_2D^+ can be used as a chemical clock to follow the evolution of massive clumps, in this work we presented the first survey of o-H_2D^+ in high-mass star-forming regions. We collected more than ~10^4 spectra in 106 sources of the ATLASGAL-survey and almost entirely belonging to the TOP100 sample. We found 16 sources with reliable detections of o-H_2D^+ (with iS/N $\gg$ 3), from which we retrieved column densities and relative abundances of o-H_2D^+ with respect to H_2 by fitting the line profiles under the assumption of LTE. From the line fit outputs, we also calculated the dynamical parameters of each clump (see Tables 1 and 3). The results of this work can be summarised as follows:

1. We confirm, also in the high-mass regime of star formation, the empirical correlation between X(o-H_2D^+) and T_{dust}, and

we find new correlations with L_{bol} and L/M. $X(\text{o-}H_2D^+)$ has been found to correlate also with $N(H_2)$, but we interpret this result as less relevant for the H_3^+ deuterium fractionation process since it is not supported by an evident correlation with $N(\text{o-}H_2D^+)$, which instead has been found with f_D and with all the other quantities connected to the evolutionary stage of the clump(s). It is likely that a denser clump might have already enabled the conversion of o-H_2D^+ into other deuterated species (e.g. D_2H^+ and N_2D^+) due to fast kinetics, with the effect of reducing the abundance of o-H_2D^+;

2. From the additional detections of DCO^+, $H^{13}CO^+$, and $C^{17}O$ available in our sample, we have added six new estimates of ζ_2 in massive star-forming regions by following the analytical formulae recently reported by Bovino et al. (2020). We find a variation in the estimated ζ_2 with the H_2 column density. Nevertheless, while being connected to the morphology of the magnetic field lines of each source, as discussed by Padovani et al. (2018), we did not find any signature that the ζ_2 can be interpreted as an evolutionary indicator of the star formation activity. We estimate a mean $\zeta_2 = 2.5 \times 10^{-17}\,s^{-1}$, assuming that the region associated with the o-H_2D^+ emission is comparable to the clump's effective radius in Table 1;
3. We confirm that $X(\text{o-}H_2D^+)$ shows a general downward trend with massive clumps evolution, as pointed out by Giannetti et al. (2019). We extend this trend to the most advanced phases of HII regions through upper limit estimations. This new result, together with that of Giannetti et al. (2019), establish the role of this tracer as a chemical clock, and provides a useful reference for future observations of o-H_2D^+ in massive star-forming regions;
4. We note that the connection between $X(\text{o-}H_2D^+)$ and the dynamical parameters of clumps is not as evident as in the case of other quantities linked to the evolution of the clumps (e.g. L/M). This result is in agreement with the findings reported by Körtgen et al. (2017) and Bovino et al. (2019) suggesting that the deuterium fractionation of H_3^+ seems to be driven more by the thermal than by the dynamical evolution of the high-mass star-forming regions. It is also true that the whole process depends on the balance between the density distribution and the thermal evolution as very dense fragments will provide a larger degree of CO freeze-out with consequent larger deuterium fractionation (see Bovino et al. 2019).

In this context, it could be interesting to verify the anti-correlation between o-H_2D^+ and N_2D^+ pointed out by Giannetti et al. (2019) in the IRDC G351.77–0.51, which has not yet been confirmed for the low-mass regime (Miettinen 2020). This challenging goal, if confirmed and supported by a detailed description of a chemical model in a sample of independent sources, would provide an additional tool to follow the star formation process, especially in the case of the high-mass regime.

Acknowledgements. The authors wish to thank an anonymous Referee for his/her comments and suggestions that have helped clarify many aspects of this work, and are grateful to M. Padovani, J. Brand and E. Redaelli for fruitful scientific discussions and feedbacks. This work was partly funded by the Marco Polo program (Università di Bologna), making it possible for GS to spend three months at the Departamento de Astronomía (Universidad de Concepción) in Concepción; and also partly carried out within the Collaborative Research Council 956, sub-project A6, funded by the Deutsche Forschungsgemeinschaft (DFG). This work was also based on data acquired with the Atacama Pathfinder EXperiment (APEX). APEX is a collaboration between the Max Planck Institute for Radioastronomy, the European Southern Observatory, and the Onsala Space Observatory. GS wish to thank the GILDAS-team for their support. SB is financially supported by CONICYT Fondecyt Iniciación (project 11170268), CONICYT programa de Astronomia Fondo Quimal 2017 QUIMAL170001, and BASAL Centro de Astrofisica y Tecnologias Afines (CATA) AFB-17002. TC has received financial support from the French State in the framework of the IdEx Université de Bordeaux Investments for the future Program. This research has made use of the IRAM GILDAS software (http://www.iram.fr/IRAMFR/GILDAS), the Cologne Database for Molecular Spectroscopy (CDMS), the NASA's Astrophysics Data System Bibliographic Services (ADS), the Astropy (Astropy Collaboration 2013, 2018; see also http://www.astropy.org) and Matplotlib (Hunter 2007).

References

Amano, T., & Hirao, T. 2005, J. Mol. Spectr., 233, 7
Astropy Collaboration (Robitaille, T. P., et al.) 2013, A&A, 558, A33
Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, AJ, 156, 123
Bacmann, A., Lefloch, B., Ceccarelli, C., et al. 2002, A&A, 389, L6
Bacmann, A., Lefloch, B., Ceccarelli, C., et al. 2003, ApJ, 585, L55
Bergin, E. A., & Tafalla, M. 2007, ARA&A, 45, 339
Bergin, E. A., Alves, J., Huard, T., & Lada, C. J. 2002, ApJ, 570, L101
Bertoldi, F., & McKee, C. F. 1992, ApJ, 395, 140
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., Sipila, O., Chambers, E., et al. 2014, Nature, 516, 219
Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, A&A, 564, A125
Butner, H. M., Lada, E. A., & Loren, R. B. 1995, ApJ, 448, 207
Carter, M., Lazareff, B., Maier, D., et al. 2012, A&A, 538, A89
Caselli, P., & Ceccarelli, C. 2012, A&ARv, 20, 56
Caselli, P., Walmsley, C. M., Terzieva, R., & Herbst, E. 1998, ApJ, 499, 234
Caselli, P., Walmsley, C. M., Tafalla, M., Dore, L., & Myers, P. C. 1999, ApJ, 523, L165
Caselli, P., Walmsley, C. M., Zucconi, A., et al. 2002, ApJ, 565, 344
Caselli, P., Vastel, C., Ceccarelli, C., et al. 2008, A&A, 492, 703
Caselli, P., Sipilä, O., & Harju, J. 2019, Phil. Tran. R. Soc. London Ser. A, 377, 20180401
Ceccarelli, C., Caselli, P., Herbst, E., Tielens, A. G. G. M., & Caux, E. 2007, Protostars and Planets V, 47
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
Chandrasekhar, S., & Fermi, E. 1953, ApJ, 118, 116
Chen, H.-R., Liu, S.-Y., Su, Y.-N., & Zhang, Q. 2010, ApJ, 713, L50
Contreras, Y., Schuller, F., Urquhart, J. S., et al. 2013, A&A, 549, A45
Crapsi, A., Caselli, P., Walmsley, C. M., et al. 2005, ApJ, 619, 379
Csengeri, T., Urquhart, J. S., Schuller, F., et al. 2014, A&A, 565, A75
Csengeri, T., Leurini, S., Wyrowski, F., et al. 2016, A&A, 586, A149
Dalgarno, A., & Lepp, S. 1984, ApJ, 287, L47
de Boisanger, C., Helmich, F. P., & van Dishoeck, E. F. 1996, A&A, 310, 315
Doty, S. D., van Dishoeck, E. F., van der Tak, F. F. S., & Boonman, A. M. S. 2002, A&A, 389, 446
Feng, S., Caselli, P., Wang, K., et al. 2019, ApJ, 883, 202
Fontani, F., Palau, A., Caselli, P., et al. 2011, A&A, 529, L7
Fontani, F., Giannetti, A., Beltrán, M. T., et al. 2012, MNRAS, 423, 2342
Friesen, R. K., Di Francesco, J., Myers, P. C., et al. 2010, ApJ, 718, 666
Friesen, R. K., Di Francesco, J., Bourke, T. L., et al. 2014, ApJ, 797, 27
Galli, D., & Palla, F. 2013, ARA&A, 51, 163
Gerlich, D., Herbst, E., & Roueff, E. 2002, Planet. Space Sci., 50, 1275
Giannetti, A., Wyrowski, F., Brand, J., et al. 2014, A&A, 570, A65
Giannetti, A., Leurini, S., Wyrowski, F., et al. 2017, A&A, 603, A33
Giannetti, A., Bovino, S., Caselli, P., et al. 2019, A&A, 621, L7
Güsten, R., Nyman, L. Å., Schilke, P., et al. 2006, A&A, 454, L13
Harju, J., Haikala, L. K., Lehtinen, K., et al. 2006, A&A, 454, L55
Harju, J., Juvela, M., Schlemmer, S., et al. 2008, A&A, 482, 535
Hartmann, L., Ballesteros-Paredes, J., & Heitsch, F. 2012, MNRAS, 420, 1457
Hastings, W. K. 1970, j-BIOMETRIKA, 57, 97
Hernandez, A. K., Tan, J. C., Caselli, P., et al. 2011, ApJ, 738, 11
Hezareh, T., Houde, M., McCoey, C., Vastel, C., & Peng, R. 2008, ApJ, 684, 1221
Ho, P. T. P., Moran, J. M., & Lo, K. Y. 2004, ApJ, 616, L1
Holland, W. S., Robson, E. I., Gear, W. K., et al. 1999, MNRAS, 303, 659
Hugo, E., Asvany, O., Harju, J., & Schlemmer, S. 2007, in Molecules in Space and Laboratory, 119

Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90
Indriolo, N., Neufeld, D. A., Gerin, M., et al. 2015, ApJ, 800, 40
Ivlev, A. V., Padovani, M., Galli, D., & Caselli, P. 2015, ApJ, 812, 135
Ivlev, A. V., Silsbee, K., Sipilä, O., & Caselli, P. 2019, ApJ, 884, 176
Jeffreys, H. 1961, Theory of Probability, 3rd edn. (Oxford, England: Oxford University Press)
Kauffmann, J., Pillai, T., & Goldsmith, P. F. 2013, ApJ, 779, 185
Klein, B., Hochgürtel, S., Krämer, I., et al. 2012, A&A, 542, L3
Klein, T., Ciechanowicz, M., Leinz, C., et al. 2014, IEEE Trans. Terahertz Sci. Technol., 4, 588
Kong, S., Caselli, P., Tan, J. C., Wakelam, V., & Sipilä, O. 2015, ApJ, 804, 98
Kong, S., Tan, J. C., Caselli, P., et al. 2016, ApJ, 821, 94
König, C., Urquhart, J. S., Csengeri, T., et al. 2017, A&A, 599, A139
Körtgen, B., Bovino, S., Schleicher, D. R. G., Giannetti, A., & Banerjee, R. 2017, MNRAS, 469, 2602
Körtgen, B., Bovino, S., Schleicher, D. R. G., et al. 2018, MNRAS, 478, 95
Kramer, C., Alves, J., Lada, C. J., et al. 1999, A&A, 342, 257
Leurini, S., Pillai, T., Stanke, T., et al. 2011, A&A, 533, A85
Leurini, S., Schisano, E., Pillai, T., et al. 2019, A&A, 621, A130
Li, G.-X., Urquhart, J. S., Leurini, S., et al. 2016, A&A, 591, A5
Lis, D. C., Gerin, M., Roueff, E., Vastel, C., & Phillips, T. G. 2006, ApJ, 636, 916
Lubowich, D. A., Pasachoff, J. M., Balonek, T. J., et al. 2000, Nature, 405, 1025
MacLaren, I., Richardson, K. M., & Wolfendale, A. W. 1988, ApJ, 333, 821
Maret, S., Hily-Blant, P., Pety, J., Bardeau, S., & Reynier, E. 2011, A&A, 526, A47
Metropolis, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, J. Chem. Phys., 21, 1087
Miettinen, O. 2020, A&A, 634, A115
Millar, T. J., Bennett, A., & Herbst, E. 1989, ApJ, 340, 906
Molinari, S., Pezzuto, S., Cesaroni, R., et al. 2008, A&A, 481, 345
Morales Ortiz, J. L., Ceccarelli, C., Lis, D. C., et al. 2014, A&A, 563, A127
Motte, F., Bontemps, S., & Louvet, F. 2018, Ann. Rev. Astron. Astrophys., 56, 41
Mouschovias, T. C., Tassis, K., & Kunz, M. W. 2006, ApJ, 646, 1043
Müller, H. S. P., Thorwirth, S., Roth, D. A., & Winnewisser, G. 2001, A&A, 370, L49
Padovani, M., & Galli, D. 2011, A&A, 530, A109
Padovani, M., Galli, D., & Glassgold, A. E. 2009, A&A, 501, 619
Padovani, M., Hennebelle, P., & Galli, D. 2013, A&A, 560, A114
Padovani, M., Ivlev, A. V., Galli, D., & Caselli, P. 2018, A&A, 614, A111
Pagani, L., Salez, M., & Wannier, P. G. 1992a, A&A, 258, 479
Pagani, L., Wannier, P. G., Frerking, M. A., et al. 1992b, A&A, 258, 472
Pagani, L., Vastel, C., Hugo, E., et al. 2009, A&A, 494, 623
Patil, A., Huard, D., & Fonnesbeck, C. 2010, J. Stat. Softw., 35, 1
Pillai, T., Caselli, P., Kauffmann, J., et al. 2012, ApJ, 751, 135
Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2014, ApJ, 783, 130
Roberts, H., & Millar, T. J. 2000, A&A, 361, 388
Sabatini, G., Giannetti, A., Bovino, S., et al. 2019, MNRAS, 490, 4489
Saraceno, P., Andre, P., Ceccarelli, C., Griffin, M., & Molinari, S. 1996, A&A, 309, 827
Schisano, E., Molinari, S., Elia, D., et al. 2020, MNRAS, 492, 5420
Schuller, F., Menten, K. M., Contreras, Y., et al. 2009, A&A, 504, 415
Sipilä, O., Hugo, E., Harju, J., et al. 2010, A&A, 509, A98
Spina, L., Randich, S., Magrini, L., et al. 2017, A&A, 601, A70
Swift, J. J. 2009, ApJ, 705, 1456
Tan, J. C., Kong, S., Zhang, Y., et al. 2016, ApJ, 821, L3
Tang, X. D., Henkel, C., Wyrowski, F., et al. 2018, A&A, 611, A6
Urquhart, J. S., Figura, C. C., Moore, T. J. T., et al. 2013, MNRAS, 437, 1791
Urquhart, J. S., Moore, T. J. T., Csengeri, T., et al. 2014, MNRAS, 443, 1555
Urquhart, J. S., König, C., Giannetti, A., et al. 2018, MNRAS, 473, 1059
van der Tak, F. F. S., & van Dishoeck, E. F. 2000, A&A, 358, L79
van der Tak, F. F. S., van Dishoeck, E. F., Evans, Neal J., I., & Blake, G. A. 2000, ApJ, 537, 283
Vastel, C., Caselli, P., Ceccarelli, C., et al. 2006, ApJ, 645, 1198
Vastel, C., Caselli, P., Ceccarelli, C., et al. 2012, A&A, 547, A33
Wienen, M., Wyrowski, F., Menten, K. M., et al. 2015, A&A, 579, A91
Wienen, M., Wyrowski, F., Schuller, F., et al. 2012, A&A, 544, A146
Wienen, M., Wyrowski, F., Walmsley, C. M., et al. 2020, A&A, submitted
Wiles, B., Lo, N., Redman, M. P., et al. 2016, MNRAS, 458, 3429
Winkel, B., Wiesemeyer, H., Menten, K. M., et al. 2017, A&A, 600, A2
Wouterloot, J. G. A., Henkel, C., Brand, J., & Davis, G. R. 2008, A&A, 487, 237
Yan, Y. T., Zhang, J. S., Henkel, C., et al. 2019, ApJ, 877, 154

Appendix A: Details on the o-H_2D^+ detection limits

In this section, we report the details of the sources used to calculate upper limits for the o-H_2D^+ detection limits, following the procedure described in Sect. 4.1. All data related to those sources are summarised in Table A.1. Cols. 2, 3, and 7 respectively give T_{dust}, N(H_2), and the evolutionary class; the data are from Giannetti et al. (2017) and Urquhart et al. (2018) (see Tables 1 and 3). In Col. 4, we report the average rms noise of each spectrum provided by GILDAS-class, while Cols. 5 and 6 respectively give N(o-H_2D^+) and X(o-H_2D^+) upper limits (derived as in Sect. 4.1). The resulting median X(o-H_2D^+) upper limits for each evolutionary class are shown as grey markers in Fig. 7.

Table A.1. Summary of the ATLASGAL sources used to compute the o-H_2D^+ detection limits.

ATLASGAL-ID	T_{dust} (K)	N(H_2) $\log_{10}$(cm^{-2})	rms noise [a] (K)	N(o-H_2D^+) [b] $\log_{10}$(cm^{-2})	X(o-H_2D^+) [b] $\log_{10}$(N[o-H_2D^+]/N[H_2])	Class
G30.85-0.08 [c]	16.7	22.72	0.07	<12.77	<−9.95 ± 0.14	70w
G320.88-0.40 [c]	16.8	22.69	0.09	<12.86	<−9.83 ± 0.13	70w
G337.28+0.01 [c]	10.7	22.92	0.09	<13.07	<−9.86 ± 0.12	70w
G338.07+0.01 [c]	18.5	22.51	0.07	<12.75	<−9.76 ± 0.15	70w
G338.78+0.48 [c]	12.2	22.85	0.07	<12.90	<−9.95 ± 0.13	70w
G351.13+0.77 [c]	18.6	22.49	0.06	<12.69	<−9.80 ± 0.14	70w
G351.57+0.76 [c]	17.0	22.67	0.03	<12.40	<−10.27 ± 0.14	70w
G353.42-0.08 [c]	17.1	22.38	0.04	<12.49	<−9.89 ± 0.14	70w
G08.68-0.37 [c]	24.2	22.94	0.07	<12.67	<−10.27 ± 0.16	IRw
G10.45-0.02 [c]	20.7	22.70	0.05	<12.58	<−10.12 ± 0.15	IRw
G18.73-0.23 [c]	21.9	22.90	0.06	<12.66	<−10.24 ± 0.15	IRw
G18.89-0.47 [c]	14.4	23.17	0.02	<12.25	<−10.92 ± 0.13	IRw
G23.21-0.38 [c]	22.1	23.20	0.03	<12.36	<−10.84 ± 0.15	IRw
G24.63+0.17 [c]	18.1	22.57	0.06	<12.71	<−9.86 ± 0.14	IRw
G317.87-0.15 [c]	19.3	23.03	0.06	<12.68	<−10.35 ± 0.14	IRw
G318.78-0.14 [c]	24.9	22.60	0.07	<12.68	<−9.92 ± 0.15	IRw
G331.71+0.60 [c]	21.0	22.88	0.05	<12.56	<−10.32 ± 0.15	IRw
G335.79+0.17 [c]	24.7	23.18	0.06	<12.61	<−10.57 ± 0.16	IRw
G337.26-0.10 [c]	21.7	22.72	0.06	<12.62	<−10.10 ± 0.16	IRw
G338.92+0.56 [c]	24.2	23.48	0.07	<12.68	<−10.80 ± 0.16	IRw
G340.78-0.10 [c]	26.2	22.82	0.08	<12.70	<−10.12 ± 0.16	IRw
G342.48+0.18 [c]	23.6	22.82	0.05	<12.50	<−10.32 ± 0.16	IRw
G343.75-0.16 [c]	24.3	23.30	0.06	<12.57	<−10.73 ± 0.16	IRw
G351.45+0.66 [c]	21.4	23.88	0.03	<12.34	<−11.54 ± 0.16	IRw
G19.88-0.54 [c]	24.2	23.12	0.03	<12.30	<−10.82 ± 0.16	IRb
G34.41+0.23 [c]	26.1	23.25	0.05	<12.50	<−10.75 ± 0.16	IRb
G37.55+0.20 [c]	28.4	22.73	0.05	<12.50	<−10.23 ± 0.16	IRb
G53.14+0.07 [c]	25.4	22.90	0.04	<12.42	<−10.48 ± 0.16	IRb
G332.09-0.42 [c]	30.8	22.94	0.05	<12.46	<−10.48 ± 0.16	IRb
G333.31+0.11 [c]	25.9	22.76	0.04	<12.44	<−10.32 ± 0.16	IRb
G341.22-0.21 [c]	27.0	22.86	0.06	<12.55	<−10.31 ± 0.16	IRb
G351.16+0.70 [c]	21.9	23.67	0.06	<12.62	<−11.05 ± 0.15	IRb
G351.25+0.67 [c]	32.5	23.34	0.06	<12.56	<−10.78 ± 0.16	IRb
G10.62-0.38 [c]	34.5	23.57	0.08	<12.67	<−10.90 ± 0.17	HII
G12.81-0.20 [c]	35.1	23.56	0.07	<12.60	<−10.96 ± 0.17	HII
G14.33-0.64 [d]	21.6	23.51	0.06	<12.65	<−10.86 ± 0.15	HII
G31.41+0.31 [c]	26.3	23.57	0.06	<12.62	<−10.95 ± 0.16	HII
G34.26+0.15 [c]	31.0	23.85	0.07	<12.66	<−11.19 ± 0.16	HII
G34.40+0.23 [c]	22.8	23.11	0.06	<12.63	<−10.48 ± 0.15	HII
G330.95-0.18 [c]	33.0	23.76	0.04	<12.37	<−11.39 ± 0.16	HII
G333.13-0.43 [c]	35.2	23.40	0.05	<12.50	<−10.90 ± 0.16	HII
G333.28-0.39 [c]	30.4	23.24	0.05	<12.51	<−10.73 ± 0.16	HII
G333.60-0.21 [c]	41.1	23.49	0.07	<12.60	<−10.89 ± 0.17	HII
G337.40-0.40 [c]	31.8	23.36	0.07	<12.70	<−10.65 ± 0.16	HII
G337.70-0.02 [c]	25.6	23.35	0.08	<12.66	<−10.70 ± 0.16	HII
G343.13-0.06 [c]	30.9	23.39	0.05	<12.48	<−10.91 ± 0.16	HII

Notes. Derived physical properties of the sample of massive clumps used to derive the o-H_2D^+ detection limits described in Sect. 4.1. The sources are separated, from top to bottom, into their evolutionary classes. The error associated with N(H_2) is 20% for each source. [a]The temperatures are reported on the main-beam temperature scale; [b]derived in this work. [c]T_{dust}, N(H_2), and Class from Giannetti et al. (2017). [d]T_{dust}, N(H_2), and Class from Urquhart et al. (2018).

Appendix B: Details on linear fit

Each data set comprises N measurements $\mathcal{D} \equiv \{x_k, \delta x_k, y_k, \delta y_k\}$, with $k = 1, \ldots, N$. Each point k of a data set corresponds to a different massive star-forming region of the ATLASGAL survey; x_k and y_k are the quantities over which we look for a linear correlation, while δx_k and δy_k are the corresponding errors. We test our data sets against linear correlations of the form $y = mx + q$, where m and q are, respectively, the slope and the normalisation. In our specific case we always consider the y_k as measures of $\log_{10} X(\text{o-H}_2\text{D}^+)$, while the x_k can be the target's galactocentric radius, $D_{\rm GC}$; effective radius, $R_{\rm eff}$; clump total mass, $\log_{10} M_{\rm clump}$; bolometric luminosity, $\log_{10} L_{\rm bol}$; dust temperature, $T_{\rm dust}$; CO-depletion factor, $f_{\rm D}$; H_2 column density, $\log_{10} H_2$; luminosity-to-mass ratio, $\log_{10} L/M$; virial factor, α; or the Mach number, $\mathcal{M}$ (see Tables 1 and 3 and Sect. 2).

We assume that the errors δx_k and δy_k follow a Gaussian distribution and we define the model's likelihood $\ln \mathcal{L}(\boldsymbol{\theta}|\mathcal{D})$, defined by the parameters $\boldsymbol{\theta} \equiv \{m, q\}$, given the data $\mathcal{D}$

$$\ln \mathcal{L} \equiv \frac{1}{2} \sum_{k=1}^{N} \left[-\left(\frac{y_k - m x_k - q}{\sigma_k} \right)^2 - \ln(2\pi\sigma_k^2) \right], \tag{B.1}$$

where $\sigma = \sqrt{m^2 \delta x_k^2 + \delta y_k^2}$.

We rely on a Bayesian approach and we run a parameter space search using a Markov chain Monte Carlo (MCMC) algorithm, using uninformative flat priors over the models' free parameters. For each correlation, we run ten chains, each evolved for 10000 steps, using a classical Metropolis-Hasting sampler to sample from the posterior (Metropolis et al. 1953; Hastings 1970). All the chains converge quickly towards the highest probability region of the parameter space and we eliminate, as a conservative choice, the first 3000 steps from each chain, considered to be a reliable burn-in. The remaining steps are used to compute the posterior distribution over the model's free parameters. We test the correlations we obtain from the MCMC also considering smaller burn-in and we note that our results do not depend significantly on this choice as long as we eliminate at least the first 1000 steps.

To test the performances of our models in fitting data, for each correlation we evaluate the Bayesian evidence $\mathcal{E}$, defined as the average of the likelihood (Eq. (B.1)) under the considered priors. Specifically,

$$\mathcal{E} = \int \mathcal{L}(\boldsymbol{\theta}|\mathcal{D}) P(\boldsymbol{\theta})\, \mathrm{d}\boldsymbol{\theta}, \tag{B.2}$$

where P denotes the priors that we used. Afterwards, we compute the Bayesian evidence after fitting each data set with a model where the slope has been fixed to zero. In $\mathcal{M}_1$ and $\mathcal{M}_2$, which are respectively the models with $m = 0$ and the general model with two free parameters, the Bayes factor,

$$\mathcal{B}_{2,1} = \frac{\mathcal{E}_2}{\mathcal{E}_1}, \tag{B.3}$$

describes whether $\mathcal{M}_1$ provides a better description of the data than $\mathcal{M}_2$ ($\mathcal{B}_{2,1} < 1$), or the other way around ($\mathcal{B}_{2,1} > 1$). The Bayesian evidence is computed using the software package PyMultiNest (Buchner et al. 2014), as implemented in its publicly available version.

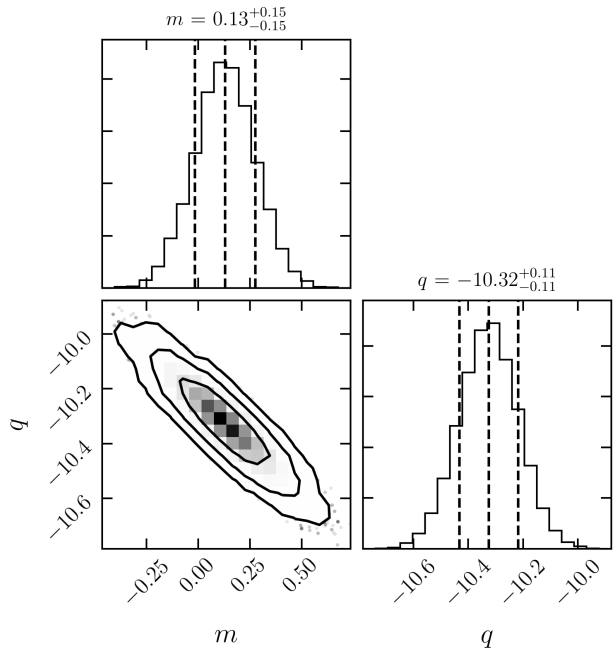

Fig. B.1. Example of one- and two-dimensional marginalized posterior distributions of the free parameters in our model. The black curves in the two-dimensional distribution correspond to regions enclosing 68%, 95%, and 99.7% of the total probability, while the black vertical lines in the one-dimensional marginalised distributions correspond to the 16th, 50th, and 84th percentiles, used to estimate the uncertainties on the models' parameters.

Figure B.1 shows an example of the posterior distribution obtained for each correlation. The black curves in the two-dimensional distributions indicate regions enclosing 68%, 95%, and 99.7% of the total probability (i.e. 1σ, 2σ, and 3σ, respectively). The black vertical lines in the one-dimensional marginalised distributions instead indicate the 16th, 50th, and 84th percentiles, used to estimate the 1σ error bars. To derive the 2σ error bars we use instead the interval between the 5th and 95th percentiles of the corresponding one-dimensional marginalised posterior distributions.