OH in the diffuse interstellar medium: physical modelling and prospects with upcoming SKA precursor/pathfinder surveys

S. A. Balashev [ID],[1]★ N. Gupta[2] and D. N. Kosenko[1]

[1]*Ioffe Institute, 26 Politekhnicheskaya st, St. Petersburg 194021, Russia*
[2]*Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, 411 007 Pune, India*

ABSTRACT

Hydroxyl (OH) is known to form efficiently in cold gas ($T \sim 100\,$K) along with the molecule H_2 and can be used as an efficient tracer of the diffuse molecular gas in the interstellar medium (ISM). Using a simple formalism describing the $H\,I/H_2$ transition and a reduced network of major chemical reactions, we present a semi-analytical prescription to estimate the abundances of O-bearing molecules in the diffuse ISM. We show that predictions based on our prescription are in good agreement with the estimates obtained using the MEUDON PDR code which utilizes the full reaction network. We investigate the dependence of the relative abundances of $OH/H\,I$ and OH/H_2 on the variations of physical conditions i.e. the metallicity, number density (n), cosmic ray ionization rate (ζ), and strength of UV field (χ) in the medium. We find that the $OH/H\,I$ abundances observed in the Galactic ISM can be reproduced by models with $n \sim 50\,\mathrm{cm}^{-3}$, $\chi \sim 1$ (Mathis field), and $\zeta \sim 3 \times 10^{-17}\,\mathrm{s}^{-1}$, with a variation of about 1 dex allowed around these values. Using the constrained H_2 column density distribution function at $z \sim 3$, we estimate the OH column density distribution function and discuss future prospects with the upcoming large radio absorption line surveys.

Key words: ISM: abundances – ISM: molecules – quasars: absorption lines.

1 INTRODUCTION

The cold atomic ($H\,I$) and molecular (H_2) phases of the interstellar medium (ISM) are basic fuel for star formation in galaxies. For this reason, understanding the physical properties and evolution of these two phases is of great interest and crucial to understand many key observables – such as the cosmic evolution of star formation rate density (SFRD; Madau & Dickinson 2014) – related to the galaxy evolution. In a diffuse molecular cloud, these two phases are intimately linked to each other through an important ingredient of the ISM physics – the $H\,I/H_2$ transition. This transition occurs in a very narrow region of the medium with typical temperatures, $T \sim 100\,$K, and number densities, $n \sim 100\,\mathrm{cm}^{-3}$.

The cold atomic phase which is classically referred to as the cold neutral medium (CNM; Heiles & Troland 2003b) can be easily observed through $H\,I$ 21-cm line in emission and absorption. However, due to the weakness of the transition the emission lines are mostly detected in the Galaxy and the local Universe. In turn, the major constituent of the molecular ISM i.e. H_2 is extremely hard to observe in emission. Consequently, it is mostly studied indirectly, through other molecular species such as CO and HCO^+, whose production is coupled with the presence of substantial amount of H_2 in the medium (Dame, Hartmann & Thaddeus 2001; Leroy et al. 2009; Tacconi et al. 2018; Freundlich et al. 2019). However, these molecular emission line observations mostly trace the dense molecular gas, which generally represent clumps of the cold ISM embedded in the elongated envelopes of diffuse gas. In the local

Universe, the diffuse molecular gas can be detected via far-UV Lyman- and Werner-band absorption lines of H_2. At $z \gtrsim 2$, these lines are redshifted to optical wavelengths, which has enabled sensitive observations of diffuse molecular gas in distant galaxies using large ground-based telescopes. It has been recently established, that this diffuse molecular gas, that is not traced by CO (Balashev et al. 2017), can be present in a significant amount in the ISM (e.g. $\sim$30 per cent of the molecular gas in our Galaxy; Grenier, Casandjian & Terrier 2005; Pineda et al. 2013).

One of the promising tracer of the diffuse gas is the OH molecule. Indeed it is found to be significantly widespread with respect to CO, and can be used to observe ISM phases that are partly atomic, partly molecular i.e. envelopes surrounding dense molecular clouds, which are neither detectable in CO nor HI emission. In fact, OH was the first molecule detected at radio wavelengths (Weinreb et al. 1963). Since then, along with HCO^+, it has emerged as one of the best tracers of H_2 gas (Liszt & Lucas 1999). It is quite efficient to observe it through the four 18-cm ground-state Λ-doubling transitions which occur at rest frequencies of 1612.231, 1665.402, 1667.359, and 1720.530 MHz. The relative strengths of these lines in the local thermodynamic equilibrium is 1612:1665:1667:1720 MHz = 1:5:9:1. But the line ratios are seldom found to be in LTE and are sensitive to the local physical conditions in the ISM.

In the Galaxy, OH is routinely detected in a broad range of astrophysical environments. This includes OH maser emission arising from the star-forming regions (Caswell 1999), the envelopes of the late-type stars (te Lintel Hekkert et al. 1989; Engels & Bunzel 2015), the proto-planetary nebulae (te Lintel Hekkert & Chapman 1996), and the supernovae remnants (Brogan et al. 2013). The non-amplified OH emission and OH absorption are detected in the diffuse

★ E-mail: s.balashev@gmail.com

molecular clouds in the Galactic plane (e.g. Dawson et al. 2014), the individual Giant Molecular Cloud complexes (as W43 Walsh et al. 2016), the clouds at the Galactic Centre (Boyce & Cohen 1994), the high-latitude diffuse and translucent clouds (Grossmann et al. 1990; Cotten et al. 2012; Donate, White & Magnani 2019), and cirrus clouds of our Galaxy (Barriault et al. 2010). Notably, Li et al. (2018) recently published OH absorption line measurements from the Millenium survey (Heiles & Troland 2003a). They find that most of the detections are associated with the diffuse and translucent clouds, i.e. probing the so-called 'CO-dark' gas, and the peak of the lognormal function fitted to the excitation temperature distribution is 3.5 K. The latter explains the general difficulty in detecting OH emission from diffuse clouds.

Outside the Galaxy, OH is detected in many luminous infrared galaxies locally and up to redshifts of 0.265 (Darling & Giovanelli 2002; Fernandez et al. 2010), not only as a megamaser emission, but also in absorption (McBride, Alatalo & Nyland 2015). However, in normal star-forming galaxies, the detections are sparse. To date only four intervening OH absorbers are known at $z > 0$ (Chengalur, de Bruyn & Narasimha 1999; Chengalur & Kanekar 2003; Kanekar et al. 2005; Gupta et al. 2018b). Three of these originate from the dense molecular gas in gravitationally lensed systems, that pre-select the line of sights (LOSs) with high dust-extinction passing through the regions close to the centre of galaxies. Only recently the OH main lines have been systematically searched in a sample of quasar sightlines tracing the diffuse cold atomic gas (Gupta et al. 2018b). This led to the first successful extragalactic detection of OH originating from the diffuse gas (see also Combes et al. 2019). The upcoming large blind H I 21-cm absorption line surveys, particularly the MeerKAT Absorption Line Survey (MALS; Gupta et al. 2016) and the the First Large Absorption Survey in H I (FLASH; Allison et al. 2020) with the Square Kilometre Array (SKA) precursors and pathfinders will have the sensitivity and redshift paths to search OH 18-cm lines in a wide-range of environments in the Galaxy and distant galaxies at $0 < z < 2$. A summary of these surveys is provided in table 1 of Gupta et al. 2016. Additionally, extragalactic OH may also be searched with Five-hundred-meter Aperture Spherical Telescope (FAST; Zheng et al. 2020) and upgraded Giant Metrewave Radio Telescope (uGMRT; Gupta et al. 2020). Since the diffuse gas has much higher cross-section than the dense gas, one may expect that the former will be preferentially represented in these surveys. It is therefore of great interest to obtain the expected detection rates and the physical conditions in diffuse gas that will be probed by these surveys. This requires a detailed understanding of the production of OH and the influence of external factors on the populations of levels.

The production of OH molecule through the O-bearing cycle has been studied in detail for different phases of ISM by many groups (e.g. Black & Dalgarno 1973; Herbst & Klemperer 1973; Prasad & Huntress 1980; Le Petit, Roueff & Le Bourlot 2002; Hollenbach et al. 2012; Bialy & Sternberg 2015). Most of the theoretical studies of OH production either assumed some abundances of H_2 in the medium or calculated abundances of various species using complex chemical reaction network and radiative transfer modelling. A key component of the modelling is H I/H_2 transition, since it is crucial to determine H_2 abundance, which is an important ingredient of the production cycle of O-bearing molecules. Recently, Sternberg et al. (2014) and Bialy & Sternberg (2016) presented a simple analytical formalism to describe H I/H_2 transition in the cold medium. In Balashev & Kosenko (2020), we successfully applied this formalism to describe the production of the HD molecule in a semi-analytical manner. In this paper, we expand this semi-analytical formalism to describe the production

cycle of O-bearing molecules in diffuse ISM. We show that the galactic measurements by Li et al. (2018) can be well explained by our calculations and provide an estimate of the expected physical conditions in such diffuse molecular clouds. More specifically, we find that the measurements in high-latitude clouds, probing the cold ISM, with H_2 column densities $\log N \gtrsim 19$ (where N in cm^{-2}). This, together with recent constraints of the H_2 column density distribution function by Balashev & Noterdaeme 2018, allow us to estimate the expected detection rates of OH in the blind absorption line surveys.

This paper is organized as follows. In Section 2, we describe the semi-analytical model to calculate the abundances of the O-bearing molecules in the diffuse ISM. In Section 3, we present results from the calculation and show that these are in a good agreement with the estimates from sophisticated MEUDON PDR numerical code. We also examine how the relative abundances of OH/H I and OH/H_2 depend on the variation of the physical conditions in the medium. In Section 4, we show that our prescription explains the observations in the Galaxy. Subsequently, we use it to estimate the column density distribution function of OH and hence the expected incidence rate of OH detection from the upcoming blind absorption line surveys with a particular focus on MALS. We summarize our results in Section 5. Throughout this paper, we assume a flat ΛCDM cosmology model with $\Omega_\Lambda = 0.7$ and $H_0 = 68\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

2 ANALYTICAL DESCRIPTION

We consider the homogeneous and isothermal medium with following physical parameters that determine the chemical abundances: the total hydrogen number density n_H^{tot} (in the following, we will use $n_2 = n_H^{tot}/100\,cm^{-3}$), kinetic temperature T (in K), the metallicity Z (relative to solar), the incidence UV field of strength χ (in the units of Mathis field; Mathis, Mezger & Panagia 1983) and cosmic rays ionization rate (CRIR), ζ (in units of $10^{-17}\,s^{-1}$). For the latter, we will implicitly assume that it is the primary ionization rate per hydrogen atom.

In this work, we consider diffuse ISM, with typical number densities $n \sim 10-1000\,cm^{-3}$, exposed by relatively mild UV fields $\chi < 100$ and CRIR values $\zeta < 100$. We also explicitly study the dependence of the abundances of various species on the metallicity, which we varied in the range: $Z \sim 0.01-1$. We fixed the temperature to $T = 100\,K$, i.e. we did not consider it as an additional parameter. This is reasonable because the relevant reaction rates have little temperature dependence in the range (50–200 K) corresponding to the diffuse ISM (Balashev et al. 2019). Further, in the following for a species denoted by X, we will use n_X, x_X, d_X, and N_X to describe the number density, relative abundance, depletion, and column density, respectively. We will also explicitly denote atomic hydrogen by H I and express all column densities in cm^{-2}.

The ionization potentials of C and O (metals that are relevant for this study) are 11.2 and 13.6 eV, respectively. Therefore, in the diffuse medium, they are predominantly in the ionized and neutral form, respectively. Hence, their number densities can be defined as

$$n_{C^+} \approx n_C^{tot} \equiv x_C n_H^{tot} \approx 2.7 \times 10^{-4} Z d_C n_H^{tot}, \qquad (1)$$

$$n_O \approx n_O^{tot} \equiv x_O n_H^{tot} = 4.9 \times 10^{-4} Z d_O n_H^{tot}. \qquad (2)$$

Further, we assume the number density of electrons, $n_e = n_{H^+} + n_{C^+}$. Since in the CNM, He, and O are predominantly in the neutral form, we will neglect their contributions to n_e. Following Bialy &

Sternberg ([2019](#)) for the depletions d_C and d_O, we use

$$d_X = \begin{cases} 1 - \delta_X & Z \geq Z_0 \\ 1 - \delta_X(Z/Z_0)^2 & Z < Z_0, \end{cases} \quad (3)$$

where $Z_0 = 0.2$ and δ_X is the specified depletion at solar metallicity. We used $\delta_C = 0.53$ and $\delta_O = 0.41$, which recover typical ISM abundances of carbon and oxygen in our Galaxy. We also considered He and Ar assuming that these element are not depleted and have abundances of 0.085 and 3.3×10^{-6}, respectively.

In Table [1](#), we summarize reactions and their rates involved in the OH production. The focus is to write the (semi-)analytical description of the abundances of O-bearing molecules. Therefore, we have restricted ourselves to the hydrogen and oxygen chemical networks and consider only two reactions of O-bearing species with C^+, which are necessary to obtain correct destruction rates of OH and H_2O molecules in the diffuse ISM. In Fig. [1](#), we present a simplified visual representation of the major reactions involved in the OH production cycle. Such representations have been presented in the literature (see for example Prasad & Huntress [1980](#); Bialy & Sternberg [2015](#)). Most of the reaction rates have been taken from the UMIST data base (McElroy et al. [2013](#)). Note reaction 11 is missed in the data base. For O-H charge exchange reactions, we used rates from Draine ([2011](#)). Compared to UMIST, these provide a better fit to the detailed calculation by Stancil et al. ([1999](#)). For the grain recombination reaction, we used the fits to rates provided by Weingartner & Draine ([2001](#)). We did not study the grain and PAH charge distributions explicitly. For uniformity, the photodissociation rates and dust shielding have been taken from the recent extensive study by Heays, Bosman & van Dishoeck ([2017](#)). These assume the plane–parallel geometry and the one-sided isotropic radiation field is of the spectral shape given by Mathis et al. ([1983](#)).

2.1 H_2 molecular fraction

Our description of the abundances of the species starts with defining the H_2/H I transition, which essentially governs the gross structure of the cloud. Following the formalism of Sternberg et al. ([2014](#)), we express the H_2 molecular fraction f_{H_2} as a function of N_{H_2} through

$$f_{H_2} \equiv \frac{2n_{H_2}}{n_H^{tot}} \equiv \frac{2n_{H_2}}{n_H + 2n_{H_2}} = \frac{2}{\alpha S_{H2}(N_{H_2})e^{-\sigma_g(N_H + 2N_{H_2})} + 2}, \quad (55)$$

where S_{H2} is the self-shielding function for H_2 (Draine & Bertoldi [1996](#)) and $\sigma_g \approx 1.9 \times 10^{-21} Z'$ cm^2 is the dust LW-photon absorption cross-section per hydrogen atom. The ratio of free space H_2 photo-dissociation and H_2 formation on the dust grains, i.e. α is defined as

$$\alpha \equiv \frac{\chi D_0^{H2}}{R^{H2}n_H^{tot}} = 0.6 \times 10^4 n_2^{-1} \chi Z^{-1}, \quad (56)$$

where D_0^{H2} is the unattenuated photodissociation rate of H_2 in Mathis UV field and R^{H2} is the formation rate of H_2 on the dust grains. To be consistent with MEUDON PDR models, we used $D_0^{H2} = 4.8 \times 10^{-11}$ s^{-1} and $R^{H2} = 8 \times 10^{-17}$ cm^3 s^{-1} (see Section 3.3.2). Sternberg et al. ([2014](#)) also found that N_{HI} can be expressed as a function of N_{H_2} through

$$N_{HI}^{Stern} = \frac{1}{\sigma_g} \ln\left(\frac{\alpha G(N_{H_2})}{2} + 1\right), \quad (57)$$

where $G(N_{H_2})$ is cloud-averaged H_2 self-shielding function. Here, we use the analytical fit to it from the paper by Bialy & Sternberg ([2016](#)).

In this formalism, the atomic hydrogen abundance drops to vanishingly small values in the H_2 dominated core of the cloud, and hence, the H I column density has a fixed asymptotic value for each set of the α and σ_g parameters. However, note that even at the cloud centre, due to dissociation of H_2 by cosmic rays, and/or desorption/adsorption reactions on the surface of dust grains the H I can be still present at reasonable levels. Based on the MEUDON PDR calculations (see Section 3.1), we roughly estimated the x_{HI} to be $10^{-3}/Z$ in H_2-dominated regions. This gives the correction to the N_{HI} provided in equation (57) as

$$N_{HI} = N_{HI}^{Stern}(N_{H_2}) + \frac{10^{-3}}{Z} N_{H_2}. \quad (58)$$

2.2 O^+, He^+, HeH^+, and ArH^+

The abundance of O^+ is determined by the destruction with H_2 (24), and O-H charge exchange reactions (5) and (6). Using these reactions, we get

$$n_{O^+} = \frac{k_5 n_{H^+} n_O}{k_6 n_H + k_{24} n_{H_2}}. \quad (59)$$

The abundance of He^+ is determined by cosmic ray ionization of He (47) and destruction in reactions (43), (7), (8), (9), (10), and (11). Hence

$$n_{He^+} = \frac{k_\zeta^{He} n_{He}}{n_e \alpha_{He}^{rr} + \alpha_{He}^{gr} n_H^{tot} + (k_8 + k_9)n_{H_2} + (k_{10} + k_{11})n_H}. \quad (60)$$

The He^+ provides the main formation channel for HeH^+, which is important for the formation H_2^+ (in reaction 12) and H_3^+ (in reaction 13). The abundance of HeH^+ can be obtained as

$$n_{HeH^+} = \frac{k_{11} n_{He^+} n_H}{k_{12} n_H + k_{13} n_{H_2}}. \quad (61)$$

The production of Ar^+ is important for the H_3^+ formation in diffuse atomic medium. Ar^+ is formed in cosmic ray ionization and destructed mainly with H_2 leading to production of ArH^+, which in turn is destructed in reactions (18) and (19) with e and O, respectively, and with H_2 in reaction (20). The latter results in H_3^+ formation. Therefore, we get

$$n_{ArH^+} = \frac{k_\zeta^{Ar} n_{Ar}}{k_{18} n_e + k_{19} n_O + k_{20} n_{H_2} + D^{ArH^+}}. \quad (62)$$

2.3 H_2^+ and H_3^+

The formation of H_2^+ proceeds through two channels: one is due to cosmic rays ionization of H_2 (46), and the other is due to HeH^+ through reaction (12) Note, the reaction (9) can be neglected. The destruction of H_2^+ proceeds through reactions (14), (15), (16), and photodissociation, which typically can be neglected. Hence

$$n_{H_2^+} = \frac{n_{H_2} k_\zeta^{H_2} + k_{12} n_{HeH^+} n_H}{k_{14} n_H + k_{15} n_e + k_{16} n_{H_2} + D^{H_2^+}}. \quad (63)$$

In the medium where H_2 abundance is high, H_3^+ molecule is preferentially formed from H_2^+ through reaction (16), and in the atomic envelope of the cloud from ArH^+ and HeH^+ through reactions (20) and (13), respectively. H_3^+ is destroyed in reactions (21), (22), (23), and (27). Therefore

$$n_{H_3^+} = \frac{k_{16} n_{H_2^+} n_{H_2} + k_{20} n_{ArH^+} n_{H_2} + k_{13} n_{HeH^+} n_{H_2}}{(k_{21} + k_{22})n_e + (k_{23} + k_{27})n_O}. \quad (64)$$

Table 1. Gas-phase reactions.

Reaction[a]	Notation	α (cm^3 s^{-1})	β	γ	References
$H^+ + e \rightarrow H + h\nu$	(4)	α_H^{rr}	3.5×10^{-12}	-0.7	–	(Prasad & Huntress 1980)
$H^+ + O \rightarrow H + O^+$ [b]	(5)	k_5	6.9×10^{-10}	0.26	224	(McElroy et al. 2013)
$O^+ + H \rightarrow H^+ + O$ [b]	(6)	k_6	5.7×10^{-10}	0.36	-8.6	(McElroy et al. 2013)
$He^+ + e \rightarrow He + h\nu$	(7)	α_{He}^{rr}	4.5×10^{-12}	-0.67	–	(Prasad & Huntress 1980)
$He^+ + H_2 \rightarrow He + H + H^+$	(8)	k_8	3.7×10^{-14}	–	35	(McElroy et al. 2013)
$He^+ + H_2 \rightarrow He + H_2^+$	(9)	k_9	7.2×10^{-15}	–	–	(McElroy et al. 2013)
$He^+ + H \rightarrow He + H^+$	(10)	k_{10}	1.2×10^{-15}	0.25	–	(Stancil, Lepp & Dalgarno 1998)
$He^+ + H \rightarrow HeH^+ + h\nu$	(11)	k_{11}	1.7×10^{-15}	-0.37	–	(Le Petit et al. 2006)
$HeH^+ + H \rightarrow He + H_2^+$	(12)	k_{12}	8.2×10^{-10}	0.11	31.5	(Bovino et al. 2011)
$HeH^+ + H_2 \rightarrow He + H_3^+$	(13)	k_{13}	1.5×10^{-9}	–	–	(McElroy et al. 2013)
$H_2^+ + H \rightarrow H_2 + H^+$	(14)	k_{14}	6.4×10^{-10}	–	–	(Karpas, Anicich & Huntress 1979)
$H_2^+ + e \rightarrow H + H$	(15)	k_{15}	1.6×10^{-8}	-1.18	7.12	(Epée Epée et al. 2016)
$H_2^+ + H_2 \rightarrow H_3^+ + H$	(16)	k_{16}	2.1×10^{-9}	–	–	(Theard & Huntress 1974)
$Ar^+ + H_2 \rightarrow ArH^+ + H$	(17)	k_{17}	8.4×10^{-10}	0.16	–	(Rebrion, Rowe & Marquette 1989)
$ArH^+ + e \rightarrow H + Ar$	(18)	k_{18}	5×10^{-10}	-0.5	–	(Mitchell et al. 2005)
$ArH^+ + O \rightarrow Ar + OH^+$	(19)	k_{19}	8×10^{-10}	–	–	(Schilke et al. 2014)
$ArH^+ + H_2 \rightarrow H_3^+ + Ar$	(20)	k_{20}	8×10^{-10}	–	–	(Villinger et al. 1982)
$H_3^+ + e \rightarrow H + H + H$	(21)	k_{21}	4.4×10^{-8}	-0.52	–	(McCall et al. 2004)
$H_3^+ + e \rightarrow H_2 + H$	(22)	k_{22}	2.3×10^{-8}	-0.52	–	(McCall et al. 2004)
$H_3^+ + O \rightarrow OH^+ + H_2$	(23)	k_{23}	8.0×10^{-10}	-0.16	1.4	(Bettens, Hansen & Collins 1999)
$H_2 + O^+ \rightarrow OH^+ + H$	(24)	k_{24}	1.7×10^{-9}	–	–	(Smith, Adams & Miller 1978)
$OH^+ + e \rightarrow H + O$	(25)	k_{25}	3.75×10^{-8}	-0.50	–	(Mitchell 1990)
$OH^+ + H_2 \rightarrow H_2O^+ + H$	(26)	k_{26}	1.0×10^{-9}	–	–	(Jones, Birkinshaw & Twiddy 1981)
$H_3^+ + O \rightarrow H_2O^+ + H$	(27)	k_{27}	3.4×10^{-10}	-0.16	1.4	(Bettens et al. 1999)
$H_2O^+ + e \rightarrow O + H_2$	(28)	k_{28}	3.9×10^{-8}	-0.50	–	(Rosen et al. 2000)
$H_2O^+ + e \rightarrow O + H + H$	(29)	k_{29}	3.1×10^{-7}	-0.50	–	(Rosen et al. 2000)
$H_2O^+ + e \rightarrow OH + H$	(30)	k_{30}	8.6×10^{-8}	-0.50	–	(Rosen et al. 2000)
$H_2O^+ + H_2 \rightarrow H_3O^+ + H$	(31)	k_{31}	6.4×10^{-10}	–	–	(McElroy et al. 2013)
$H_3O^+ + e \rightarrow OH + H_2$	(32)	k_{32}	5.4×10^{-8}	-0.50	–	(Novotný et al. 2010)
$H_3O^+ + e \rightarrow OH + H + H$	(33)	k_{33}	3.1×10^{-7}	-0.50	–	(Novotný et al. 2010)
$H_3O^+ + e \rightarrow H_2O + H$	(34)	k_{34}	7.1×10^{-8}	-0.50	–	(Novotný et al. 2010)
$H + O \rightarrow OH + h\nu$	(35)	k_{35}	9.9×10^{-19}	-0.38	–	(Le Petit et al. 2006)
$OH + H \rightarrow H_2O + h\nu$	(36)	k_{36}	4.0×10^{-18}	-2.00	–	(Le Petit et al. 2006)
$OH + H^+ \rightarrow H + OH^+\nu$	(37)	k_{37}	2.1×10^{-9}	-0.50	–	(Prasad & Huntress 1980)
$OH + C^+ \rightarrow CO^+ + H$	(38)	k_{38}	7.7×10^{-10}	-0.50	–	(Prasad & Huntress 1980)
$H_2O + C^+ \rightarrow HCO^+ + H$	(39)	k_{39}	7.0×10^{-10}	-0.50	–	(Martinez et al. 2008)
$H_2O + C^+ \rightarrow HOC^+ + H$	(40)	k_{40}	1.4×10^{-9}	-0.50	–	(Martinez et al. 2008)
$OH + O \rightarrow O_2 + H$	(41)	k_{41}	3.7×10^{-11}	-0.25	12.9	(Prasad & Huntress 1980)
Grain recombination:						
$H^+ + Gr^{+k} \rightarrow H + Gr^{k+1}$	(42)	α_H^{gr}		Approximation from		(Weingartner & Draine 2001)
$He^+ + Gr^{+k} \rightarrow He + Gr^{k+1}$	(43)	α_{He}^{gr}		Approximation from		(Weingartner & Draine 2001)
Cosmic ray ionization[c]						
$H + c.r \rightarrow H^+ + e$	(44)	k_ζ^H	$1.7 \times \zeta$	–	–	(Draine 2011)
$H_2 + c.r \rightarrow H^+ + H + e$	(45)	$\tilde{k}_\zeta^{H_2}$	$0.17 \times \zeta$	–	–	(Le Petit et al. 2002)
$H_2 + c.r \rightarrow H_2^+ + e$	(46)	$k_\zeta^{H_2}$	$3.4 \times \zeta$	–	–	(Le Petit et al. 2002)
$He + c.r \rightarrow He^+ + e$	(47)	k_ζ^{He}	$1.4 \times \zeta$	–	–	(Jenkins 2013)
$Ar + c.r \rightarrow Ar^+ + e$	(48)	k_ζ^{Ar}	$12.5 \times \zeta$	–	–	(Schilke et al. 2014)

| **Photodissociation reactions** | | | | | |
Reaction	Notation	Rate, D_0 (s^{-1})[d]	γ_{ds}[e]	References
$H_2 + h\nu$	(49)		4.8×10^{-11}	4.18	from Meudon
$H_2^+ + h\nu$	(50)	$D^{H_2^+}$	3.9×10^{-10}	2.78	(Heays et al. 2017)
$OH + h\nu$	(51)	D^{OH}	2.5×10^{-10}	2.66	(Heays et al. 2017)
$OH^+ + h\nu$	(52)	D^{OH^+}	1.1×10^{-11}	3.97	(Heays et al. 2017)
$H_2O + h\nu$	(53)	D^{H_2O}	5.3×10^{-10}	2.63	(Heays et al. 2017)
$ArH^+ + h\nu$	(54)	D^{ArH^+}	1.0×10^{-11}	2.63	(Schilke et al. 2014)

[b] We adopted approximation of these reaction rates from UMIST2012 data base, which is based on Stancil et al. (1999) calculations, while Spirko, Zirbel & Hickman (2003) obtained different rates. See also discussion in Neufeld & Wolfire (2017).

[c] Most of the cosmic ray ionization rates (except for He) depend on the the fraction of secondary electrons which was calculated based on the physical parameters of models (Draine 2011; Jenkins 2013).

[d] D_0 is a free-space photodissociation rate in Mathis field (Mathis et al. 1983), i.e. neglecting any extinction and self-shielding.

[e] γ_{ds} is a constant in the factor related to the shielding by dust. Following Heays et al. (2017), $D = D_0 e^{-\gamma_{ds} A_V}$.
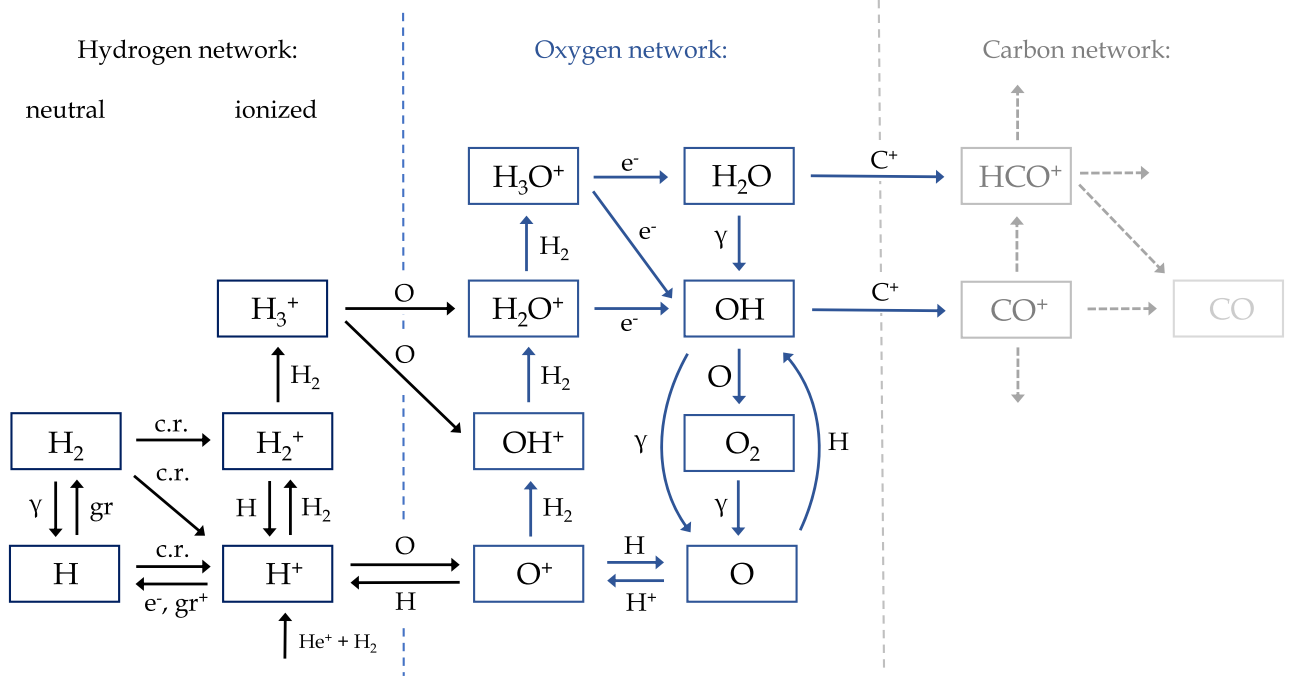
Figure 1. Simplified representation of chemical reaction network used in this paper. We plot only major reactions from the Table. 1. To keep the representation simple, we also do not show the channels that include He and Ar chemistry used in the model.

The chemistry of H_2^+ and H_3^+ was recently extensively discussed by Neufeld & Wolfire (2017) in the context of cosmic ray ionization rate in the diffuse medium. However, we note that to derive their equations, they did not take into account the formation of H_2^+ and H_3^+ through HeH^+ and ArH^+, which resulted in smaller abundances of H_2^+ and H_3^+ in diffuse atomic part of the clouds, where H_2 abundance is low.

2.4 Ionization fraction

After determining He^+ abundance, and knowing how to relate O^+ and H^+ abundance, the next step is to determine hydrogen ionization fraction. This is an important part of our prescription to calculate the abundances of O-bearing molecules. This is because the ionization fraction determines O^+ abundance, and hence has a crucial importance on some of the channels for the production of O-bearing molecules. We will show that once ionization fraction is determined, we can write explicit analytical equations for abundances of O-bearing molecules.

The ionization fraction in the diffuse ISM, $f_{H^+} = n_{H^+}/n_H^{tot}$ can be obtained assuming that H^+ is mainly produced by ionization by cosmic rays (44 and 45), charge exchange reaction with O^+ (6), and from He^+ and H_2^+ in reactions (8) and (14), respectively. In turn, H^+ is destroyed by radiative recombination (42), charge-exchange reaction between H^+ and oxygen (5), and neutralization by grains (42). We neglect here the channels of the formation from D^+ by reaction with H and H_2 and destruction with D, since they precisely equate each other, so they can be omitted from the balance equation for H^+. Therefore

$$k_\zeta^H n_H + \tilde{k}_\zeta^{H_2} n_{H_2} + k_6 n_{O^+} n_H + k_{14} n_{H_2^+} n_H + k_8 n_{He^+} n_{H_2}$$

$$= \alpha^{rr} n_e n_{H^+} + \alpha_H^{gr} n_H^{tot} n_{H^+} + k_5 n_O n_{H^+}. \tag{65}$$

From the balance equation (59) for n_{O^+}, we can write: $k_5 n_O n_{H^+} - k_6 n_{O^+} n_H = k_{24} n_{O^+} n_{H_2}$. Using this and the expressions (59), (60) and (63) for n_{O^+}, n_{He^+} and $n_{H_2^+}$, respectively, the equation (65) is modified to

$$k_\zeta^H n_H + \tilde{k}_\zeta^{H_2} n_{H_2} + k_\zeta^{H_2} n_{H_2} \frac{k_{14} n_H}{k_{14} n_H + k_{15} n_e + k_{16} n_{H_2}}$$

$$+ k_\zeta^{He} n_{He} \frac{k_8 n_{H_2} + \dfrac{k_{11} n_H k_{14} n_H}{k_{14} n_H + k_{15} n_e + k_{16} n_{H_2}} \dfrac{k_{12} n_H}{k_{12} n_H + k_{13} n_{H_2}}}{n_e \alpha_{He}^{rr} + \alpha_{He}^{gr} n_H^{tot} + (k_8 + k_9) n_{H_2} + (k_{10} + k_{11}) n_H}$$

$$= (\alpha^{rr} n_e + \alpha_H^{gr} n_H^{tot}) n_{H^+} + \frac{k_5 k_{24} n_O n_{H_2}}{k_6 n_H + k_{24} n_{H_2}} n_{H^+}. \tag{66}$$

This equation can be solved numerically, taking into account that the ion neutralization coefficients α_H^{gr} and α_{He}^{gr} are functions of the physical parameters. However, one can also write an approximate solution of this equation. Indeed, using Weingartner & Draine (2001) for $T = 100$ K, we can express α_H^{gr} as

$$\alpha_H^{gr} \approx 2.7 \times 10^{-11} \psi^{-0.9} \text{cm}^3 \text{ s}^{-1}$$

$$\approx 1.6 \times 10^{-10} \chi^{-1} n_2 Z x_e \text{ cm}^3 \text{ s}^{-1} \equiv \tilde{\alpha}_H^{gr} x_e, \tag{67}$$

where $\psi = G_0 \sqrt{T/K}/(n_e/\text{cm}^{-3})$ and $x_e = n_e/n_H^H = x_C + f_{H^+}$. Here, G_0 is the radiation field intensity relative to Habing (1968), which is a factor ≈ 1.3 lower than Mathis field in 912–1100 Å range. This approximation, where we replaced power -0.9 by -1, works well till the following inequality, $\chi n_2^{-1}(10^{-4}/x_e) \gg 0.23$, valid. This is indeed the case for physical parameter ranges relevant for the cold ISM.

Next, we neglect the $k_{15} n_e$ and $\alpha_{He}^{rr} n_e$ terms in the left-hand part of the equation (66). Then, substituting $n_H = n_H^{tot}(1 - f_{H_2})(1 - f_{H^+})$,

$n_{H_2} = n_H^{tot} f_{H_2}(1 - f_{H^+})/2$, and $n_{H^+} = f_{H^+} n_H^{tot}$, and introducing

$$k_\zeta^{eff} = k_\zeta^H(1 - f_{H_2}) + \tilde{k}_\zeta^{H_2} f_{H_2}/2 + k_\zeta^{H_2} \frac{f_{H_2}}{2}\left(1 + \frac{k_{16}f_{H_2}}{2k_{14}(1 - f_{H_2})}\right)^{-1}$$

$$+ k_\zeta^{He} f_{He} \frac{k_8 \frac{f_{H_2}}{2} + k_{12}\left(1 + \frac{k_{16}f_{H_2}}{2k_{14}(1 - f_{H_2})}\right)^{-1}}{\alpha_{He}^{gr} + (k_8 + k_9)f_{H_2}/2 + (k_{10} + k_{11})(1 - f_{H_2})}, \quad (68)$$

we obtain the equation for f_{H^+} as

$$f_{H^+}^2 + f_{H^+}\left(x_C + \frac{k_5 x_O A}{(\alpha^{rr} + \tilde{\alpha}^{gr})}\right) = \frac{k_\zeta^{eff}}{(\alpha^{rr} + \tilde{\alpha}^{gr})n_H^{tot}}, \quad (69)$$

where $A = \left(1 + \frac{k_6 n_H}{k_{24}n_{H_2}}\right)^{-1} \approx \frac{2f_{H_2}}{1 + f_{H_2}}$ (at $T = 100$ K) and we take into account that $(\alpha^{rr} + \tilde{\alpha}^{gr})x_C \gg k_\zeta^{eff}/n_H^{tot}$. Note that a substitution similar to (68) was used by Glassgold & Langer ([1974]).

Based on the quadratic equation (69), we can write the solution for f_{H^+} as

$$f_{H^+} = \tilde{x}\left(\sqrt{\frac{k_\zeta^{eff}}{(\tilde{\alpha}^{gr} + \alpha^{rr})n_H^{tot}\tilde{x}^2} + 1} - 1\right), \quad (70)$$

where

$$\tilde{x} = \frac{1}{2}\left(x_C + \frac{k_5 x_O A}{\alpha^{rr} + \tilde{\alpha}^{gr}}\right). \quad (71)$$

2.5 OH$^+$, H$_2$O$^+$, and H$_3$O$^+$

H$_3^+$ and H$_2^+$ molecules provide the main channels of OH$^+$ formation following reactions (23) and (24). Once OH$^+$ is formed it mainly produces H$_2$O$^+$ in reaction (26). Additionally, OH$^+$ can be destructed in recombination with electrons (25) and photodissociated (52). Therefore

$$n_{OH^+} = \frac{k_{23}n_{H_3^+}n_O + k_{24}n_{O^+}n_{H_2}}{k_{26}n_{H_2} + k_{25}n_e + D^{OH^+}}. \quad (72)$$

Apart from reaction (26), there is an additional channel of H$_2$O$^+$ formation in reaction (27). The destruction H$_2$O$^+$ includes production of H$_3$O$^+$ in reaction (31) and three recombination channels with electrons (28), (29), and (30). Hence

$$n_{H_2O^+} = \frac{k_{26}n_{OH^+}n_{H_2} + k_{27}n_{H_3^+}n_O}{k_{31}n_{H_2} + (k_{28} + k_{29} + k_{30})n_e}. \quad (73)$$

The abundance of H$_3$O$^+$ is determined by

$$n_{H_3O^+} = \frac{k_{31}n_{H_2O^+}n_{H_2}}{(k_{32} + k_{33} + k_{34})n_e}. \quad (74)$$

Note that one of the channels of recombination of H$_3$O$^+$ molecules leads to the formation of H$_2$O (34), which can be photodissociated to OH or react with C$^+$ to form HCO$^+$ (39) or its isomer HOC$^+$ (40).

2.6 OH and H$_2$O

Finally, the number density of OH molecules, is determined by the formation in reactions (30), (32), (33), and (35). Additionally, OH is formed by the photodissociation of H$_2$O. This, along with the destruction by reaction with C$^+$ (see reactions 39 and 40), is the dominant destruction channel for H$_2$O. H$_2$O is mainly produced by reaction (34). Therefore, we include the reaction (34) among the formation channels of OH. The destruction of OH is due to

reactions (37), (38), (41), and photodissociation (51). The reaction (37) can dominate at low metallicities.

Using these reactions, we can write that

$$n_{OH} = \frac{(k_{30}n_{H_2O^+} + (k_{32} + k_{33})n_{H_3O^+})n_e + k_{35}n_H n_O}{k_{38}n_{C^+} + k_{41}n_O + k_{37}n_{H^+} + D^{OH}}$$

$$+ \frac{k_{34}n_{H_3O^+}n_e}{k_{38}n_{C^+} + k_{41}n_O + k_{37}n_{H^+} + D^{OH}}\frac{D^{H_2O}}{(k_{39} + k_{40})n_{C^+} + D^{H_2O}}. \quad (75)$$

Finally, since OH and H$_3$O$^+$ provide the main production channel of H$_2$O in the atomic envelope (where OH predominantly formed in the reaction 35) and molecular core, respectively, the H$_2$O abundance can be written using OH and H$_3$O$^+$ abundances as

$$n_{H_2O} = \frac{k_{36}n_{OH}n_H + k_{34}n_{H_3O^+}n_e}{(k_{39} + k_{40})n_{C^+} + D^{H_2O}}. \quad (76)$$

2.7 Asymptotics

Here, we obtain asymptotics for OH abundances in certain cases:

(i) Atomic gas, $f_{H_2} \ll 1$. In this case, the OH is mainly produced in reaction (35) and destructed by photodissociation (51). Hence, x_{OH} is simply

$$x_{OH} = \frac{n_{OH}}{n_H^{tot}} = \frac{k_{35}n_O}{D^{OH}\chi} = 1.9 \times 10^{-10}\left(\frac{n_2 Z}{\chi}\right)d_O\left(\frac{T}{300}\right)^{-0.38}, \quad (77)$$

which gives $x_{OH} \approx 1.2 \times 10^{-10}n_2 Z\chi^{-1}$ at $T = 100$ K and $d_O = 0.59$. It is evident that x_{OH} is proportional to the same combination of physical parameters $n_2 Z\chi^{-1}$ as H$_2$ abundance in the atomic envelope, $n_{H_2} = n_H^{tot}f_{H_2}/2 \approx n_H^{tot}/\alpha$ (see equations 55 and 56). Therefore, the relative OH/H$_2$ abundance will not be sensitive to the number density, metallicity, and UV flux, and hence

$$\frac{n_{OH}}{n_{H_2}} = \frac{\alpha k_{35}n_O}{D^{OH}\chi} = 2.5 \times 10^{-6}d_O\left(\frac{T}{300}\right)^{-0.38}, \quad (78)$$

which for $T = 100$ K and $d_O = 0.59$ is 1.5×10^{-6}.

(ii) Molecular gas, $f_{H_2} \lesssim 1$. In this case, we can write

$$x_{OH} \approx x_O \frac{k_5 n_{H^+}}{D^{OH}}\left(\frac{1 + \frac{k_{30}n_e}{k_{31}n_{H_2}}}{1 + \frac{k_{30}n_e}{k_{31}n_{H_2}}\frac{k_{28} + k_{29} + k_{30}}{k_{30}}}\right)\frac{1}{1 + \frac{k_6 n_H}{k_{24}n_{H_2}}}. \quad (79)$$

For $T = 100$ K, we get

$$x_{OH} \approx 5.6 \times 10^{-3}\left(\frac{n_2 Z}{\chi}\right)d_O\frac{1 + B}{1 + 5B}\frac{2f_{H_2}}{1 + f_{H_2}}f_{H^+}, \quad (80)$$

where $B = \frac{k_{30}n_e}{k_{31}n_{H_2}} \approx 6 \times 10^{-2}\frac{Z}{f_{H_2}}$, if the electrons are determined by the carbon abundance. Since, f_{H^+} is a complex function of the physical parameters (see equation 70), it is hard to obtain a simple form of dependence on the physical parameters. However, in the cold ISM f_{H^+} is typically in the range of $\sim 10^{-3}-10^{-5}$ and it depends sub linearly on the physical parameters. Therefore, as $f_{H_2} \to 1$, we can write

$$x_{OH} \approx 1.1 \times 10^{-2}\left(\frac{n_2 Z}{\chi}\right)d_O f_{H^+}. \quad (81)$$

For intermediate f_{H_2} using equation (55), this gives

$$x_{OH} \approx 7.2 \times 10^{-7}\left(\frac{n_2 Z}{\chi}\right)^2 d_O\frac{f_{H^+}}{S_{H_2}}. \quad (82)$$

Thus, for intermediate f_{H_2}, x_{OH} has a steeper dependence on $n_2 Z\chi^{-1}$.
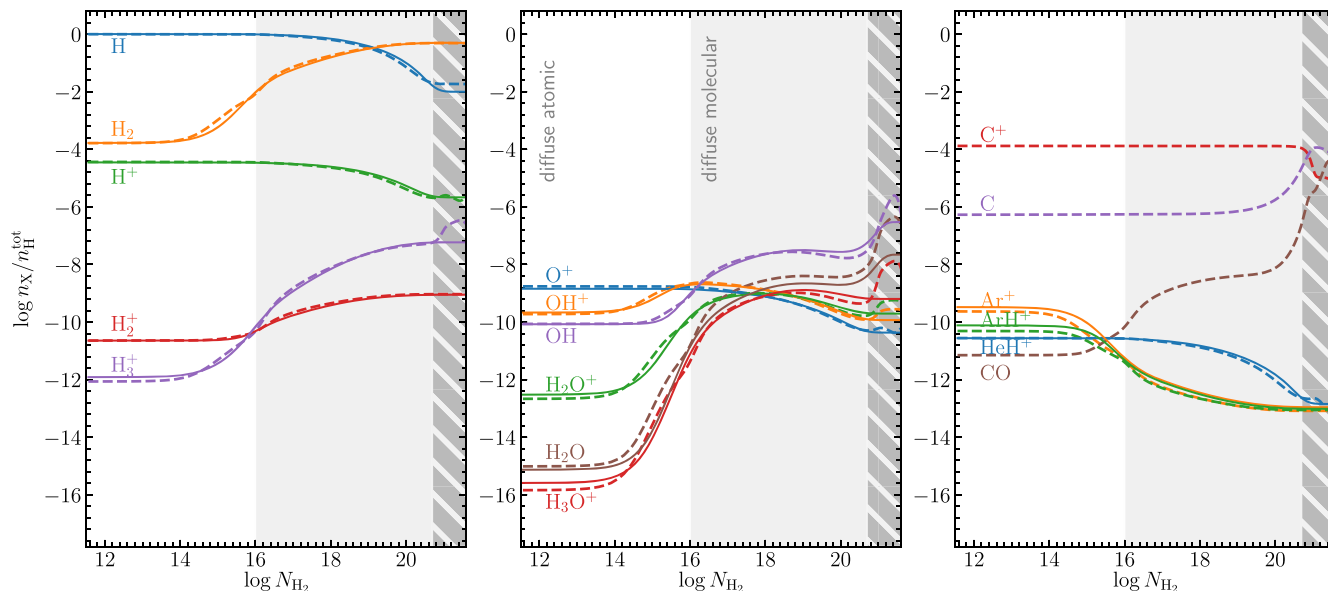
Figure 2. The relative abundances of different species involved in OH production as a function of N_{H_2} column density. The solid lines show the result of our calculation for a model with $Z = 1$, $\chi = 1$, $\zeta = 3$, and $n_H^{tot} = 50 \, cm^{-3}$, while the dashed lines represent MEUDON PDR results. The filled grey region conditionally represents the diffuse molecular part of the cloud. Since we did not take into the the conversion of C^+ convert into C I and CO and the carbon network reactions (involved mainly C and CO), therefore our calculations are not representative for the H_2 column density, $\log N_{H_2} \gtrsim 20.7$, at specified parameters of the model. This region is marked by hatch grey area.

3 RESULTS

In Fig. 2, we plot the abundances of various species estimated using the relationships discussed in Section 2. We consider $n_H^{tot} = 50$, $\chi = 1$, $\zeta = 3$, and $Z = 1$ i.e. the typical physical conditions in the local ISM. For the purpose of discussion of results, we split the cloud into three regions based on standard ISM classification: *(i)* diffuse atomic – where hydrogen is predominantly in atomic form; *(ii)* diffuse molecular – where hydrogen is in molecular form but carbon is ionized; and *(iii)* dense molecular – where hydrogen is in H_2 and carbon is in neutral and molecular forms. We merge the dense molecular region with translucent one. Recall that our calculations in dense molecular region are not appropriate for the estimation of O-bearing molecular abundances (see Section 3.1). Indeed, we did not take into account C and CO chemistry that leads to significant drop in the electron fraction due to C^+/C transition. Additionally, once CO is formed it can cool the gas to temperatures $\sim 10 \, K$, that can affect the reaction rates and number density.

For the diffuse atomic and molecular regions, our calculations reproduce the well-known behaviour of abundances of all the ions, molecules, and radicals under consideration. Briefly, in the atomic envelope, the abundances of most molecular species, such as H_2^+, H_3^+, H_2O^+, and H_3O^+, are relatively low. This is understandable because their production is coupled with the H_2 abundance. The main distinction is OH which can be formed through a neutral–neutral reaction (35), and the abundances of both H and O are relatively high in the atomic region. But, once the hydrogen starts to efficiently convert into the molecular form, due to self-shielding (at $\log N_{H_2} \gtrsim 15$), there is a proportional increase in the abundances of aforementioned molecular species. Deeper into the cloud their abundances slightly reduce. This is due to the drop in the hydrogen ionization fraction, which leads to a drop in O^+ and hence OH^+ abundances. The latter provides a dominant channel for the production of H_2O^+, H_3O^+, and consequently, OH in diffuse molecular clouds.

The profiles of various species presented in Fig. 2 would be vastly different for different physical conditions. For example, the low-metallicity environment would result in much higher hydrogen ionization fraction (see e.g. Balashev & Kosenko 2020) due to lower recombination rate on the dust grains; and also the electron fraction would be less sensitive to C^+ abundance. Additionally, in case of extremely high CRIR, $\zeta \gtrsim 100$, the electron abundance will also be less sensitive to the C^+ abundance, and H_2 can be significantly dissociated by cosmic rays in the self-shielded regions of the cloud.

3.1 Comparison with MEUDON PDR code

Here, we compare our analytical calculations with the results from MEUDON PDR code (Le Petit et al. 2006). For MEUDON PDR modelling, we consider of slab gas irradiated from one side by the radiation field (with the Mathis spectrum) with constant density and temperature $T = 100 \, K$. In general, we find that for the diffuse parts of the medium, the results from our calculations for different species involved in OH production and OH itself are in a good agreement with MEUDON PDR results.

Fig. 2 provides a comparison of our calculations and MEUDON PDR results for a model with $Z = 1$, $\chi = 1$, $\zeta = 3$, and $n = 50 \, cm^{-3}$. For diffuse atomic and diffuse molecular regions of the cloud, the major differences between our model and MEUDON PDR results are due to: *(i)* slightly different reaction rates for the O–H charge exchange reactions. This results in enhanced O^+ abundance in MEUDON PDR and hence an excess of O-bearing molecules. *(ii)* different reaction rates for ionization of Ar and He by cosmic rays. This results in enhanced Ar^+ and ArH^+ abundances in our calculation, while insignificantly changes the results of O-bearing species. *(iii)* different treatment of H_2/H I-transition. This results in differences in the regions with $\log N_{H_2} \sim 14 - 16$ for H_2, and hence, for all the other molecules and radicals. *(iv)* The difference in the
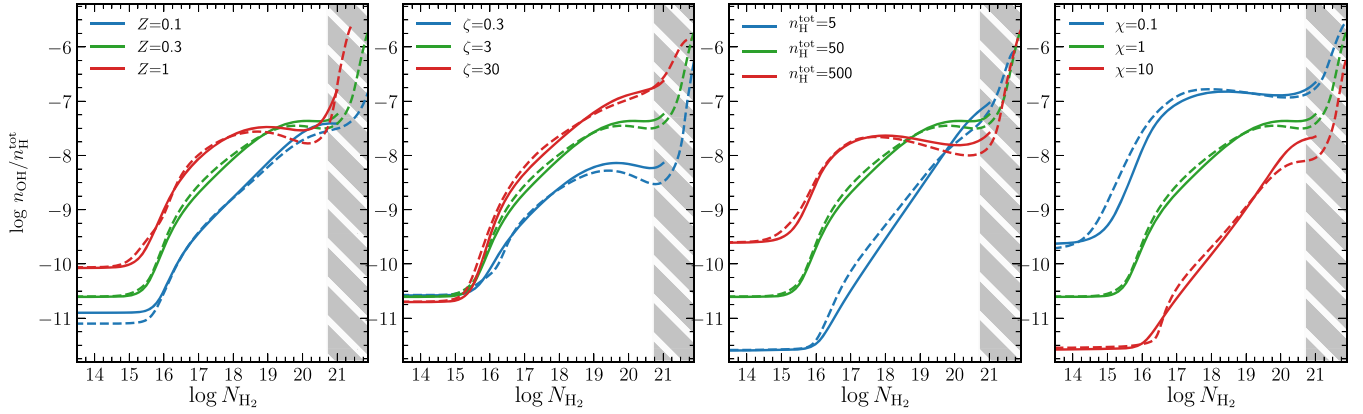
Figure 3. Dependence of $x_{OH} = n_{OH}/n_H^{tot}$ as a function of N_{H_2} on various physical parameters of the ISM. The green curves show the base model with $Z = 0.3$, $\chi = 1$, $\zeta = 3$, and $n_H^{tot} = 50$. In each panel from left- to right-hand, the blue and red lines represent the profiles obtained by varying one of the physical parameters, i.e. metallicity, CRIR, UV field strength, and number density (see legend in each panel), respectively. The solid and dashed lines correspond to our and MEUDON PDR code calculation, respectively. The grey hatch portion schematically represents the translucent and dense molecular gas region, where our calculations are no longer applicable (the exact position of this region depends on the combination of the physical parameters and hence should be calculated individually for each line).

H_3^+, OH, H_2O, and H_3O^+ photodissociation rates. Out of these, the most important is the difference for OH, which is directly reflected in the difference in OH abundance. We verified that once we use MEUDON PDR rates for the aforementioned reactions, there is almost one-to-one accordance between the results from our calculations and MEUDON PDR code, with typical differences being <30 per cents. (v) The different description of the depletion of O and C. In MEUDON PDR calculations for simplicity, we used fixed values, $d_C = 0.47$ and $d_O = 0.59$, while in our calculations, we used metallicity-dependent d_C and d_O (see equation 3). This results in the gradual increase of calculated O-bearing molecules abundances for some regions of the models with $Z < 0.2$ (see e.g. the left-hand panel in Fig. 3).

As previously mentioned, our model is not appropriate for the translucent parts of the cloud. But MEUDON PDR profiles reproduce well the expected behaviour of molecular species discussed in previous section. However, here we are concerned only with the abundance of OH molecule in diffuse atomic and molecular regions. Therefore, the results from our calculations are adequate for the purpose of this paper.

3.2 Dependence on the physical conditions

In Figs 3 and 4, we present profiles of x_{OH} as a function of N_{H_2} and N_{HI}, respectively. The profiles of N_{OH} versus N_{H_2} are presented in Fig. 5. Specifically, in each case we first obtained a base model which corresponds to $Z = 0.3$, $\chi = 1$, $\zeta = 3$, and $n_H^{tot} = 50$, and then independently varied metallicity, UV flux, CRIR, and number density to cover a range of physical conditions in the medium. For completeness, we also plot the profiles obtained using the MEUDON PDR code. As expected, for diffuse medium we find a good agreement between our and MEUDON PDR code calculations. Since the latter is able to correctly estimate OH abundances in translucent and molecular ISM, for completeness, we have retained calculations corresponding to these regions in the figures.

As obtained in Section 2.7, the OH abundance in the diffuse ISM directly depends on the combination $n_2 Z \chi^{-1}$. Such a dependence is clearly seen in Figs 3 and 4. In the diffuse atomic ISM i.e. the regions with low H_2 and low H I column densities, this dependence is very stiff and OH abundance is well reproduced by the equation (77).

In the diffuse molecular region, this dependence gets altered as the ionization and molecular fractions, with latter also determining the x_{OH}, are the functions of physical conditions n_H^{tot}, χ, ζ, and Z. Note that the dependence of x_{OH} on the cosmic ray ionization rate is not so strong. This is because the ionization fraction f_{H^+} depends sublinearly on ζ. However, in diffuse molecular ISM x_{OH} starts to have a linear dependence on ζ. Another interesting feature corresponds to deep into the cloud near the $C^+/C/CO$ transition layer which demarcates the transition from the diffuse to dense gas phase. Here, for a wide range of physical conditions, x_{OH} tends to approach a constant asymptotic value which is reasonably reproduced by approximation given in equation (81).

In Fig. 5, we plot how N_{OH}^{obs} as a function of N_{H2} depends on the variation the physical conditions. We show OH column density as $N_{OH}^{obs} = 2 \times N_{OH}(N_{H_2}/2)$. Since our calculations are based on one-side radiation field models, this procedure simulates the slab medium exposed by radiation field and CRIR at both the sides. As discussed in subsequent sections, while comparing models with observations, we are interested only on column densities of $\log N_{OH} \gtrsim 12$, which corresponds to $\log N_{H_2} \gtrsim 18$ (see Fig. 5). It is interesting to note that at these column densities, the relative N_{OH}/N_{H_2} abundance is most sensitive to UV flux and CRIR (at $\zeta < 3$), but has very little sensitivity to metallicity. Additionally we note, that at the current observational sensitivity limits $N_{OH} \gtrsim 10^{13}$ cm^{-2}, the OH molecules probe the medium with H_2 column densities $N_{H_2} \gtrsim 10^{20}$ cm^{-2}, that corresponding to the diffuse molecular ISM, where H I/H_2 transition is already complete. From Fig. 5, one can see that the absorption systems with N_{OH} in ranges $10^{13}-10^{14}$ cm^{-2} correspond to the molecular gas before onset of CO, i.e. the 'CO-dark' molecular gas.

3.3 Systematics and limitations

As already mentioned in previous sections, our calculations are applicable only to the diffuse atomic and molecular phases of ISM, i.e. well before the onset of $C^+/C/CO$ transition. Here, we discuss other limitations and possible systematics that may impact the obtained profiles of the abundances of O-bearing molecules which one should be aware of.
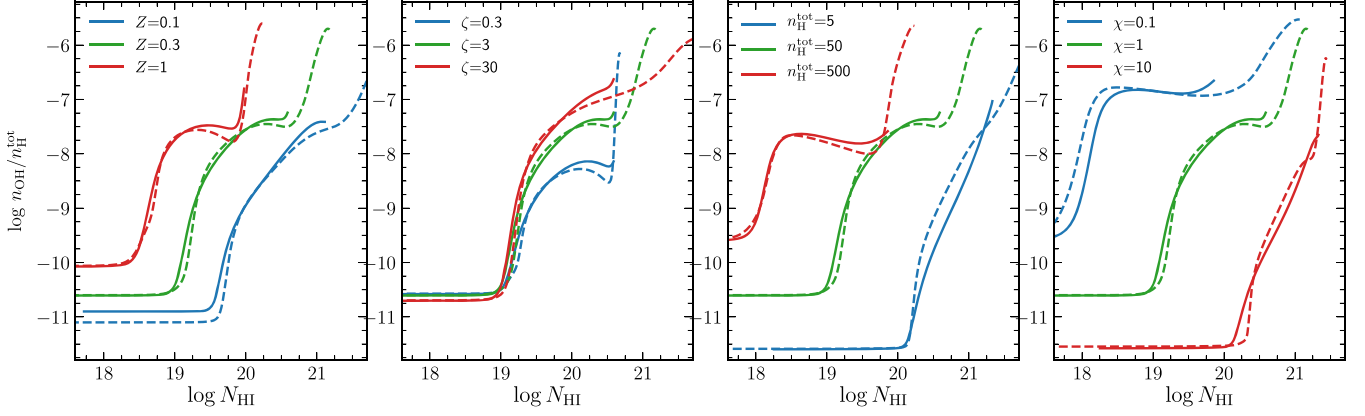
Figure 4. Dependence of $x_{OH} = n_{OH}/n_{H}^{tot}$ as a function of N_{HI} on various physical parameters of the ISM. The other details are the same as in Fig. 3.
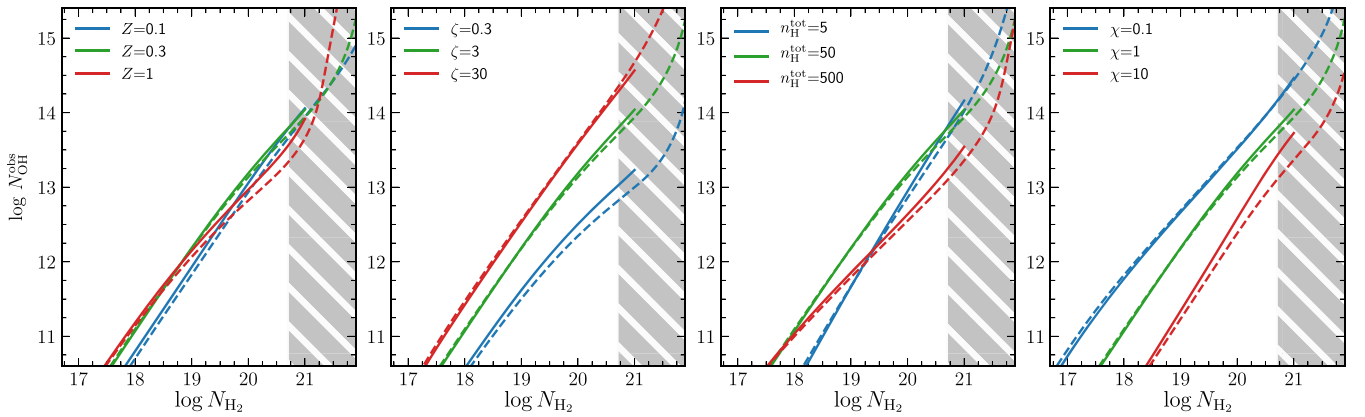
Figure 5. Dependence of N_{OH} as a function of N_{H_2} on various physical parameters of the ISM. The other details are the same as in Fig. 3.

3.3.1 Temperature gradient

The O–H charge exchange reactions are quite sensitive to the temperature of the gas. This is especially important if the gas gets heated to temperatures of $\sim 1000\,K$, for example, in the H_2 dissociated shell of an H_2 cloud. The situation corresponds to the case of a medium with low number densities ($n \lesssim$ few $\times\,10\,cm^{-3}$) and exposed to high UV fluxes ($\chi \gtrsim 10$). In this case, in comparison to the isothermal case with $T \approx 100\,K$, the abundance of O^+ is significantly enhanced. At such high temperatures $\sim 1000\,K$, the H_2 formation rate is also expected to be enhanced (see e.g. Le Bourlot et al. 2012), resulting in the enhancement of the O-bearing molecules production in the atomic envelope of cloud. Both these factors result in the significant increase ($\gtrsim 1$ dex) for $\log N_{OH}$. However, we note that such regions still have little effect on the total OH column density which are accessible through current observations ($\log N_{OH} \gtrsim 12$). Additionally, these warm region are also expected to have high excitation temperatures, so will not contribute significantly to the total OH and H I.

3.3.2 H_2 formation rate

The H_2 formation rate, R^{H_2}, is not well-known in the ISM. The original measurements using *Copernicus* satellite provided the first estimates of $R^{H_2} \approx 3 \times 10^{-17}\,cm^3\,s^{-1}$ (Jura 1975). The latter measurements updated this to slightly higher values $R^{H_2} \approx 4 \times 10^{-17}\,cm^3\,s^{-1}$ (Gry et al. 2002). We use $R^{H_2} \approx 8.0 \times 10^{-17}\,cm^3\,s^{-1}$

provided by the default calculations in MEUDON PDR code with default dust size distribution parameters (power law with slope of -3.5 in the ranges of $3 \times 10^{-3} - 3 \times 10^{-1}\,\mu m$) and mass to dust ratio of 6.3×10^{-3}, polycyclic aromatic hydrocarbons fraction of 4.6×10^{-2}.

3.3.3 H_2 self-shielding function

Since H_2 self-shielding function determine the H_2 abundance profile in the cloud, it has a direct impact on the OH abundance. In our calculations, we used Draine & Bertoldi (1996) approximation, which is found to be in good agreement with the MEUDON PDR code calculations which use the approximation given by Federman, Glassgold & Kwan (1979). However, the self-shielding function depends on the H_2 level populations and radiative transfer in H_2 lines. Therefore, it is a complex function of the cloud structure. In the first approximation, the function significantly depends on the turbulent Doppler parameter in the medium. In our calculations, we used $b = 2\,km\,s^{-1}$, which is the typical value observed in the cold diffuse ISM. We found that varying the Doppler parameter in ranges $1-10\,km\,s^{-1}$ resulted in insignificant changes of N_{OH} at any N_{H_2} in the observable ranges of $\log N_{OH} \gtrsim 12$. Note that this column density range already corresponds to the shielded medium.

3.3.4 OH photodissociation rate

The unshielded photodissociation rates used in our calculations are provided in Table 1. This implicitly assumes that the thickness of
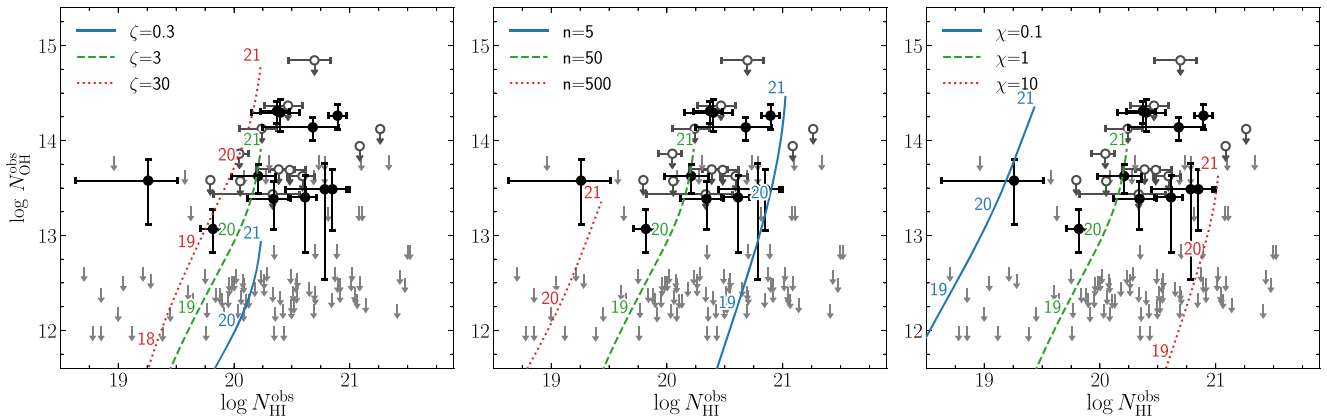
Figure 6. Comparison between relative N_{OH} and N_{HI} abundances from our calculations and observations of the Milky Way (Li et al. 2018). The circles (filled and open) and grey arrows represent OH absorption detections and non-detections, respectively, for sightlines with confirmed H I absorption detections. The error bars represent 1σ uncertainties. The green dashed curve in each panel represents the calculated $N_{OH}(N_{HI})$ profile for the base model with $Z = 1$, $\chi = 1$, $\zeta = 3$ and $n_H^{tot} = 50 \, cm^{-3}$. The red dotted and blue solid curves show models obtained by varying a parameter with respect to the base model. The parameter being varied i.e. cosmic ray ionization rate, UV field strength, or number density is indicated in the legend of panels from left- to right-hand, respectively. The numbers along the curves show corresponding $\log N_{H_2}$.

the medium is small; and the gas on the opposite side of the slab is not shielded. This is reasonable, since we are interested in the diffuse ISM, i.e. gas with small amounts of dust, which otherwise could be very effective in shielding. However, for some molecules, such as OH, the shielding by H_2 lines can be important, since the photodissociation cross-sections of OH and H_2 are intersected. Therefore, locally this shielding will depend on the H_2 column densities in the gas. Using the calculations by Heays et al. (2017), one can estimate that when $\log N_{H_2}$ is equal to 18 and 20 then the shielding of OH by H_2 is ~ 0.95 and 0.68, respectively. Since, the abundance of OH scales directly with UV flux, we can roughly estimate that taking into account this shielding will result in a $\lesssim 30$ per cent increase of $\log N_{OH}$ for the highest H_2 column densities corresponding to the diffuse ISM.

3.3.5 Multilayer structure

The observations of cold ISM always correspond to the measurement of quantities that are integrated over the LOS. Therefore, it is nearly impossible to distinguish the individual spatially separated clouds with the same LOS velocities. The fact that the observed column densities at a given velocity can be a result of the concatenation of the several individual clouds, is refer to as the multilayer structure of the observed ISM. Indeed many studies provide evidences of the mutlilayer structure of the diffuse ISM (e.g. Bialy et al. 2017). In this case, if the layers have the similar thickness in terms of H_2, then the discrepancy in the calculations will be maximum. Otherwise, the thicker component will likely dominate the observed column densities.

3.3.6 Time-dependent chemistry

All calculations in our study were done with the assumption of the steady-state chemistry, i.e. when all the reaction rates are balanced. However, some reaction rates may have very long time-scales in the diffuse ISM. The most important among these is the time-scale for the formation of H_2 on the dust grains, which can be estimated as $t_{H_2}^{form} \sim 1/(R^{H_2} n_H^{tot}) \sim 10 n_2 \, Myr$. It is similar or even less than

the dynamical time-scales for some ISM clouds. A possible impact of the absence of steady-state is that H_2 abundance and hence OH abundance would be less than compared to the steady-state scenario. The overall impact of this on calculations is hard to properly quantify. The deviations from the steady-state are most likely induced by hydrodynamical motions in the medium. Hence, the time-dependent magneto-hydrodynamical and radiative transfer calculations with very high spatial resolution are required to properly resolve H I–H_2 transition and take into account ISM chemistry. Such modelling is hardly accessible with the state of art numerical codes.

4 DISCUSSION

In this section, we compare the results from our calculations with observational data from our Galaxy and extragalactic OH absorbers. We also derive the incidence rate of OH absorbers, and discuss prospects with upcoming blind extragalactic H I and OH absorption line surveys with a focus on MALS.

4.1 Measurements in the Milky Way

In Fig. 6, we compare the relative N_{OH} and N_{HI} abundances obtained using our calculations with the observations of the Milky Way. The observational data are taken from the recent work by Li et al. (2018), who presented data for H I and OH absorption from the Galaxy towards 44 extragalactic radio continuum sources. As can be seen from fig. 8 of Li et al. (2018), the errors on H I and OH measurements in some cases are quite large. For our analysis, we have excluded *(i)* one data point which has no uncertainties on N_{OH}, and *(ii)* treat all the N_{OH} measurements with errors extending down the y-axis in fig. 8 of Li et al. (2018) as upper limits. The latter are plotted as open circles in Fig. 6.

The model profiles of $N_{OH}(N_{HI})$ from our calculations are shown as curves in Fig. 6. The OH and H I column densities were calculated as $N_x^{obs} = 2 \times N_x(N_{H_2}/2)$ to simulate the situation where the slab medium is exposed by radiation field and CRIR at both sides. In each panel, the base model with $Z = 1$, $\zeta = 3$, $n_{H_2}^{tot} = 50$, and $\chi = 1$ is shown as the dashed curve. Note that the relative N_{OH}/N_{HI} abundance is mostly sensitive to UV field and much less sensitive

to CRIR. It is evident that the base model reproduces most of the observational data points reasonably well. However, some observed sightlines have a lower or higher relative N_{OH}/N_{HI} abundance. The deviation of these points can be easily described assuming that the physical conditions have a dispersion within 1 dex of the base model. The majority of measurements with N_{OH} upper limits agree well with the model calculations. The exceptions are LOSs with relatively high H I but low OH column densities i.e. the bottom right-hand corner in Fig. 6. Principally, these points can be explained by the concatenation of several low-N_{HI} clouds along the LOS. Since N_{OH} has a steep dependence on N_{HI} in region of log N_{HI} 20−21, a few components with high N_{HI} can even have relatively small N_{OH}.

4.2 Measurements at $z > 0$

There are only four detections of OH absorption in intervening galaxies at $z > 0$. Three of them are found towards strongly lensed radio sources: PKS 1830–211 ($z = 0.886$; Chengalur et al. 1999; Combes et al. 2021), B 0218+357 ($z = 0.685$; Chengalur & Kanekar 2003), and J 0134–0931 ($z = 0.765$; Kanekar et al. 2005). These sightlines have OH column densities, log $N_{OH} \gtrsim 15$. In the Galaxy, such column densities are associated with the dense molecular ISM. The absorber towards PKS 1830–211 is particularly special. More than 50 molecular species tracing dense gas corresponding to this have been detected (Wiklind & Combes 1998; Muller et al. 2011, 2014). Our prescription is not applicable to dense gas phases associated with these.

Here, we focus on the remaining OH detection, which is at $z \approx 0.05$ towards Q 0248+430 (Gupta et al. 2018b). This sightline has log $N_{OH} \sim 14$ estimated using the excitation temperature, $T_{ex} = 3.5$ K, which is comparable to the average Galactic measurements in the diffuse ISM. But the H I column density, log $N_{HI} \sim 19.7$ estimated using the spin temperature, $T_s = 70$ K, is ∼0.5 dex lower in comparison. A moderately high spin temperature (see for example Gupta et al. 2018a) will reconcile the measurement with the base model and average Galactic measurements. Nevertheless, if that is not the case, then without any tension the measurement can be explained through a model with either relatively enhanced number densities $n_H^{tot} \gtrsim 500$ or CRIR $\zeta > 100$ or a suppressed UV field $\chi < 0.2$, or some other moderate combinations of these. In our opinion, the latter possibilities are plausible because the absorbing gas in this case is associated with the tidal tail from a galaxy at an impact parameter of ∼15 kpc.

Clearly, large samples of extragalactic H I and OH measurements of diffuse ISM are required to disentangle various possibilities and make reasonable comparisons with the Galactic measurements.

4.3 OH column density distribution function

In this section, we derive the expected OH column density distribution function which can be used to estimate the incidence rate of OH absorbers for blind absorption line surveys such as MALS. While investigating cosmological evolution of an absorber population, it is convenient to define the column density distribution function as (see for example Lanzetta et al. 1991)

$$\mathcal{F}(N, X) = \frac{d\mathcal{N}}{dNdX}, \quad (83)$$

where N represents column density, X represents absorption distance (see e.g. Noterdaeme et al. 2009, for definition), and $d\mathcal{N}$ is the number of absorbers between N and $N + dN$, and X and $X + dX$. The use of dimensionless absorption distance, X, instead of the redshifts

is convenient because it gives the value of distribution in comoving frame which guarantees that the objects with the same absorption cross-section will have the same column density function at all redshifts. Consequently, any cosmological evolution in the absorber population properties can be directly constrained.

As previously shown, at the column densities relevant for observations (log $N_{OH} \gtrsim 12$), the OH molecule traces the gas with an associated H_2 column density of log $N(H_2) > 18$. At these column densities, we can use the measured H_2 column density distribution to determine the shape of the OH column density distribution function. Specifically, for OH, we define

$$\mathcal{F}_{OH}(N_{OH}, X) = \iiint \mathcal{F}_{H_2}(N_{H_2}, X) \frac{dN_{H_2}}{dN_{OH}} f_{UV}(\chi) f_n(n)$$
$$\times f_{cr}(\zeta) f_Z(Z) d\chi dn d\zeta dZ, \quad (84)$$

where $N_{H_2} \equiv N_{H_2}(N_{OH}, \chi, n, \zeta, Z)$. The functions $f_{UV}(\chi)$, $f_n(n)$, $f_{cr}(\zeta)$, and $f_Z(Z)$ are the distributions expected for the UV field, number density, CRIR, and metallicity, respectively. The integral needs to be performed over the each of these parameters. We can also use

$$\frac{dN_{H_2}}{dN_{OH}} = \frac{n_{H_2}}{n_{OH}}, \quad (85)$$

which ease the calculations, since n_{H_2}/n_{OH} can be defined using equations (75) and (55).

It is reasonable to assume for $N(H_2)$ column density distribution $\mathcal{F}_{H_2}$ the shape of Γ-distribution[1] with the cut at the low column density region. This choice is motivated observationally. First, at the high end of H_2 column density distribution, it can be well extrapolated from the measurements of the CO (Zwaan & Prochaska 2006). Secondly, the direct measurements of the H_2 column density column density distribution using the composite spectrum of the DLAs (Balashev & Noterdaeme 2018), not only provides the constraints on the slope of the lower end of $\mathcal{F}_{H_2}$, $\beta = -1.29 \pm 0.12$, but also shows that this slope is in a good agreement with the measurements based on the CO. This value of the slope β was recently confirmed in the statistical model for QSO-DLA absorbers (Krogager & Noterdaeme 2020). Additionally, the slope based on composite spectra is in the agreement with the estimated H_2 column density distribution based on the detection in high-resolution spectra (Noterdaeme et al. 2008)[2] Hence for $\mathcal{F}_{H_2}$, we use

$$\mathcal{F}_{H_2}(N, X) = C(z)N^\beta e^{-N/N_{H_2}^{up}} \Theta \left(N - N_{H_2}^{low} \right), \quad (86)$$

where $C(z)$ is a scaling constant describing possible redshift evolution of column density distribution (i.e. we assume that the shape did not change). Θ is a Heaviside function, which take into account the abrupt drop in the number of the H_2 absorption systems at column densities lower. This drop of $\mathcal{F}_{H_2}$ is connected with self-shielding of H_2 in the medium and we set log $N_{H_2}^{low} = 18$ based both on observations and sophisticated numerical simulations (e.g. Bellomi et al. 2020). However, we note that the OH column density distribution at log $N_{OH} > 13$ (corresponding to reasonable detection limit) will be little sensitive to the exact value of $N_{H_2}^{low}$ around chosen value, since at log $N_{H_2} \sim 18$ expected OH column densities, log N_{OH}

---

[1]This function was introduced for H I gas by (Pei & Fall 1995; Storrie-Lombardi, McMahon & Irwin 1996). The motivation to use this function for H_2 gas is based on the comparison of high-z and local measurements (e.g. Balashev & Noterdaeme 2018).

[2]However, current high resolution statistic may inherit observational bias, see discussion in Balashev & Noterdaeme 2018.

$\lesssim 12$ for the wide range of the physical conditions (see Fig. 5). Additionally, the calculations are valid only for the OH column densities $\log N(\mathrm{OH}) \lesssim 14$, since at higher OH column densities, OH is predominantly associated with dense CO-bearing medium. At this regime (corresponding to $\log N(\mathrm{H}_2) \gtrsim 21$), the OH column density at given N_{H_2} is significantly enhanced in comparison to extrapolation of our calculations (see Fig. 5). Therefore, our calculation does not sensitive to the exact choice of cutoff column density $N_{\mathrm{H}_2}^{\mathrm{up}}$ in equation (86) unless $\log N_{\mathrm{H}_2}^{\mathrm{up}} \gtrsim 21$, which seems to be reasonable based on the local measurements (Zwaan & Prochaska 2006), which found $\log N_{\mathrm{H}_2}^{\mathrm{up}} \approx 23.1$. We will keep later value in the following. Also one can note that typical cross-sections of the gas at the higher end of the H$_2$ (and OH) column density distribution is very low; and hence it will be technically difficult to probe this regime in observations.

For any redshift, we can match the normalization constants of H$_2$ and OH column density distributions

$$\int_{10^{18}}^{\infty} \mathcal{F}_{\mathrm{H}_2} \mathrm{d}N = \frac{\mathrm{d}\mathcal{N}_{\mathrm{H}_2}}{\mathrm{d}X}(z) \approx \frac{\mathrm{d}\mathcal{N}_{\mathrm{OH}}}{\mathrm{d}X}(z) = \int_{\approx 10^{11}}^{\infty} \mathcal{F}_{\mathrm{OH}} \mathrm{d}N. \quad (87)$$

This equality is motivated by our modelling since we see that any sightline with $N_{\mathrm{H}_2} > 18$ will bear OH molecules with $N_{\mathrm{OH}} \gtrsim 11$ (see Fig. 5). At this lower end, one can expect that the increase of the OH column density, comes from the CNM, where hydrogen is mostly in atomic form. However, the abundance of the OH is very low in such regions (see equation 77) and therefore, to get $\log N_{\mathrm{OH}} \approx 11$, one should have $\log N_{\mathrm{HI}} \sim 21$ in the CNM. For such HI column densities, it is expected that H$_2$/H I transition should already occur.[3] Additionally, this lower end of OH column density distribution is very challenging to assess at the sensitivity limits of the current facilities. From Balashev & Noterdaeme (2018) normalization constant for H$_2$ column density distribution function, $\mathrm{d}\mathcal{N}_{\mathrm{H}_2}/\mathrm{d}X$, at $z \sim 3$ is equal $\approx 10^{-2.4}$. Since there is no agreement on the evolution of $C(z)$ between the different modelling and observations, for the illustrations purposes in the following, we will show derived quantities of OH column density distribution function and incidence rate scaled to the fraction of the diffuse molecular gas at $z \sim 3$ measured directly using QSO sightlines.

To estimate OH column density distribution using equation (84), we assume that $f_{\mathrm{UV}}(\chi), f_n(n), f_{\mathrm{cr}}(\zeta), f_Z(Z)$ are

$$f_p \sim 10^{\mu_p + \sigma_p \mathcal{N}(0,1)}, \quad (88)$$

where $\mathcal{N}(0, 1)$ is a normal distribution with zero mean and dispersion equals one. The means and dispersion chosen to be ($\mu_\chi = 0, \sigma_\chi = 0.5$), ($\mu_n = 2, \sigma_n = 0.5$), ($\mu_\zeta = 0.5, \sigma_\zeta = 0.5$), and ($\mu_Z = -0.5, \sigma_Z = 0.5$), which correspond to the expected reasonable ranges. To calculate the OH column density distribution, we randomly sampled parameters from these distributions. Then, for each realization, we also sampled the slope and normalization of the H$_2$ column density distribution from the measurements using the Sloan Digitial Sky Survey (SDSS) composite spectrum (Balashev & Noterdaeme 2018) taking into account its anticorrelation. Using the dependence of the $N_{\mathrm{H}_2}(N_{\mathrm{OH}})$, this allow us to calculate the $\mathcal{F}_{\mathrm{OH}}$ for individual realization. The estimated median and dispersion of $\mathcal{F}_{\mathrm{OH}}$ is shown in the Fig. 7. The OH column density distributions in the column density
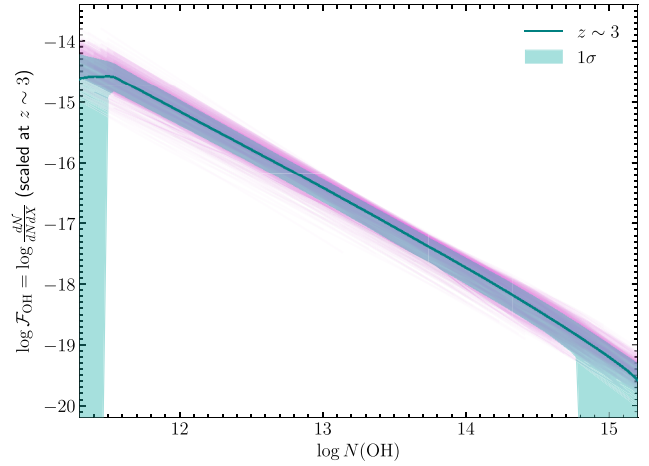
Figure 7. The OH column density distribution function, $\mathcal{F}_{\mathrm{OH}}$. The blue line show the median $\mathcal{F}_{\mathrm{OH}}$, while the blue region represents 0.68 confidence region, based on the distribution of the individual $\mathcal{F}_{\mathrm{OH}}$ (shown by red lines), calculated for a range of physical conditions, with distributions determined by equation (88).

ranges $\log N_{\mathrm{OH}} = [12, 15]$ have a power-law dependence $\sim N^\alpha$ with $\alpha \sim [-1.2, -1.3]$. This is similar to slope of H$_2$ column density distribution $\beta = -1.29$, and can be easily justified, since there is the $N_{\mathrm{H}_2}/N_{\mathrm{OH}}$ factor in the equation (86) has a almost constant at considered column densities and for the reasonable ranges of the physical parameters and hence $\alpha \approx \beta$.

4.4 The incidence rate and future prospects

Using the derived OH column density distribution, we estimate incidence rate i.e. the number of absorption systems per unit absorption distance with $N > N_{\mathrm{OH}}$ as a function of N_{OH}. This quantity can be easily calculated as a survival function of the column density distribution function provided as

$$\frac{\mathrm{d}\mathcal{N}_{\mathrm{OH}}(N > N_{\mathrm{OH}})}{\mathrm{d}X} = \int_{N_{\mathrm{OH}}}^{\infty} \mathcal{F}_{\mathrm{OH}}(N, X) \mathrm{d}N. \quad (89)$$

The estimated incidence rate of OH at $z \sim 3$ is shown in Fig. 8. Based on this, one can see that at $z \sim 3$ to detect a single OH absorption at the sensitivity limits corresponding to $\log N_{\mathrm{OH}}$ of 12, 13, and 14, one needs to probe the total absorption distance paths, ΔX, of 440_{-90}^{+260}, 1000_{-210}^{+850}, and 3100_{-100}^{+3900}, respectively.

Observationally, for an optically thin cloud, under the LTE conditions, the integrated optical depth of the strongest OH 18-cm line is related to N_{OH} through

$$N_{\mathrm{OH}} = 2.24 \times 10^{14} \frac{T_{\mathrm{ex}}}{f_c^{\mathrm{OH}}} \int \tau_{1667}(v) \, \mathrm{d}v \, \mathrm{cm}^{-2}, \quad (90)$$

where T_{ex} is the excitation temperature in Kelvin, $\tau_{1667}(v)$ is the optical depth of the 1667 MHz line at velocity v, and f_c^{OH} is the covering factor (e.g. Liszt & Lucas 1996). The milliarcsecond-scale spectroscopy of OH absorbers using the Very Long Baseline Interferometry (VLBI) is needed to probe the parsec-scale structures in the diffuse molecular gas and constrain f_c^{OH} (Srianand et al. 2013; Gupta et al. 2018a). For the purpose of calculations presented in this paper, we assume that $f_c^{\mathrm{OH}} = 1$ i.e. the absorbing gas completely covers the background radio continuum. Also, we adopt $T_{\mathrm{ex}} =$

---

[3]This consideration do not take into account neither time-dependence of H$_2$ formation in the CNM nor a possibility that large HI column density can be due to concatenation of the many layers of the CNM, each having lower HI column density to have H$_2$/HI transition.
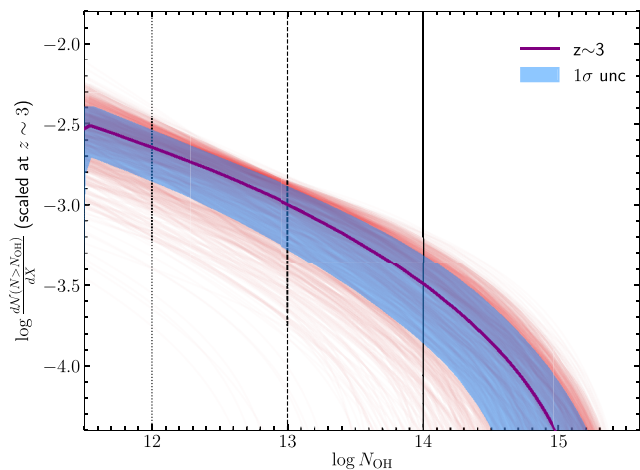
Figure 8. The incidence rate of absorption systems with $N > N_{OH}$ as a function of N_{OH}. The violet line indicates the averaged incidence rate, while the blue region represents 0.68 confidence region, based on the distribution (shown by red lines), calculated for a range of expected physical conditions (see text).

3.5 K which is the peak of the lognormal function fitted to the T_{ex} distribution of OH absorbers observed in the Galaxy (Li et al. 2018). Note that if T_{ex} is coupled to the cosmic microwave background (CMB), then it will be higher at cosmologically significant redshifts.

Recently, Gupta et al. (2020) used the uGMRT to carry out the first blind search of OH 18-cm main lines at $0.14 < z < 0.67$. For the modest total redshift path-length of 24.4 corresponding to $N_{OH} > 2.4 \times 10^{14}(T_{ex}/3.5)(1.0/f_c^{OH})$ cm^{-2}, they estimate the absorption per unit comoving path-length to be, $l_{OH}(z \sim 0.4) < 0.08$. Due to the small absorption path-length, not only the statistical power is limited but also the survey has probed only distant outskirts of star forming galaxies, mostly at impact parameters larger than 30 kpc. Clearly, as motivated by this and the Fig. 8 much larger surveys are needed.

The next major improvement in the field of extragalctic OH absorption line search may come from MALS. The survey is using MeerKAT's *L*- and *UHF*-bands covering 900–1670 and 580–1015 MHz to search for OH 18-cm main lines at $0 < z < 0.85$ and $0.64 < z < 1.87$, respectively. The *L*-band observations of the survey are well underway. For the *UHF*-band, the science verification observations are in progress. Each MALS pointing is centred at a radio source brighter than 200 mJy at 1 GHz. The absorption lines will be searched towards this central as well as numerous off-axis radio sources as faint as 5 mJy with in the telescope's field of view (FWHM $\sim 1.5°$ at 1 GHz; Jonas & MeerKAT Team 2016). Since the central radio source will be the brightest in the field of view, to maximize the low column density H I 21-cm absorption path of the survey, these are being selected to be typically at $z > 0.6$ and 1.4 for *L*- and *UHF*-bands, respectively.

For MALS, a total of ~ 1100 pointings, equally split among both the band, are expected to be observed. The target spectral rms is 0.5 mJy beam^{-1} per 5 km s^{-1}. For the central bright sources, this implies a 5σ integrated optical depth sensitivity, $\mathcal{T}(\equiv \int \tau dv) \leq 0.066$ km s^{-1}. This corresponds to a sensitivity to detect $N_{OH} \geq 5 \times 10^{13} (1.0/f_c^{OH})(T_{ex}/3.5)$ cm^{-2} towards all the sightlines. Considering the redshift distribution of MALS targets, the total absorption path lengths achieved towards the central bright radio sources for *L*-

and *UHF*-band components of the survey are, $\Delta X \approx 600$ and 1200, respectively. Based on our model predictions at $z \sim 3$, we get an incidence rate of $\approx 4.7^{+2.0}_{-2.5} \times 10^{-4}$ for log $N_{OH} > 13.7$. This roughly corresponds to $\Delta X \approx 2100^{+2300}_{-600}$ needed to obtain a single detection. If this along with the 1σ uncertainties presented in Fig. 8 is applied directly to the MALS search at $0 < z < 1.87$, only a handful of OH absorption detections are expected from the survey (see also Gupta et al. 2018b, for constraints based on Galactic measurement). But this extrapolation is highly uncertain on two accounts: *(i)* the T_{ex} may be varying as a function of z, and *(ii)* the normalization at $z \sim 3$ i.e. $C(z)$ may be varying too.

Here, it is worth elaborating on the following two caveats related to $C(z)$. Recall that the $C(z)$ presented in Fig. 8 has been estimated using the H$_2$ column density distribution at $z \sim 3$ measured using the UV absorption lines in the SDSS spectroscopic data base. First, in SDSS the spectroscopic targets have been selected on the basis of optical colours. This makes the survey biased against dust-bearing sightlines. Since the presence of dust and cold gas (H$_2$) are interlinked, the measurements of H$_2$ column density distribution function based on SDSS should be treated as a lower limit. On the other hand, the off-axis radio sources in MALS which have not been included in the above estimate because they would generally be sensitive to $N_{OH} > 14$, will have no optical or infrared colour selection function applied. Even the central MALS targets have been primarily selected on the basis of *WISE* infrared colours, and also include optically faint ($r > 21$ mag) targets which are generally excluded from optical spectroscopic surveys. Therefore, MALS is much less biased against dust. Secondly, caveat is related to the redshift evolution of $C(z)$ itself. Indeed, the simulations and emission line studies predict that the fraction of the molecular gas, Ω_{H_2} (which is proportional to $C(z)$ in equation 86), may show significant evolution, with Ω_{H_2} several times higher than at $z = 0$ and $z \sim 3$ (see Péroux & Howk 2020 and references therein). These caveats may actually push $C(z)$ in a direction which is favourable for the upcoming lower redshift radio absorption lines surveys such as MALS. Indeed, Muzahid, Srianand & Charlton (2015) find the detection rate of H$_2$ absorbers in DLAs and sub-DLAs at $z \sim 0$ to be a factor of 2 higher than at $1.8 < z < 4.2$ (Noterdaeme et al. 2008).

Eventually, the mid-frequency component of the upcoming first phase of the SKA, SKA1 (https://www.skatelescope.org/), will have the frequency coverage (350–1760 MHz) to search for OH main lines at $0 < z < 3.8$, and trace the cosmological evolution of OH abundance in diffuse ISM. The target column density sensitivity for a SKA1 survey ought to be $N_{OH} \sim 12$; and a total absorption line path in excess of 50 000. Based on $C(z \sim 3)$, this will imply a prospect of detecting at least ~ 100 OH absorbers and a realistic chance of constraining the column density distribution of diffuse OH. The required integrated 5σ optical depth sensitivity, $\mathcal{T} = 0.0013$ km s^{-1} is demanding. Pragmatically, such a survey will proceed commensally with other planned surveys (see for example table 1 of Morganti, Sadler & Curran 2015).

It is important to bear in mind that currently the physical modelling and observations based on molecular emission lines can only constrain the evolution of the dense molecular gas. The direct correspondence of the diffuse and dense molecular gas in terms of the column density distribution is not at all constrained. In fact, the normalization $C(z)$ will depend on physical conditions i.e. Z, n, ζ, and χ prevailing in the diffuse ISM at a certain redshift. Therefore, it may not scale with Ω_{H_2} as simply stated above. Thus, the upcoming OH absorption line searches with the SKA pathfinders and precursors, and eventually SKA1, will be a powerful tool to constrain the physical conditions in the diffuse ISM and be highly complementary

to ongoing molecular emission line surveys in millimeter regime that mostly probe the dense molecular gas.

5 CONCLUSIONS

In this paper, using a formalism based on the analytical description of HI/H_2 transition and a simplified network of major chemical reactions, we presented a semi-analytical prescription to estimate the abundances of O-bearing species in the diffuse ISM. We focused on OH molecule and our prescription is applicable only to diffuse ISM i.e. before C^+/C/CO transition, which corresponds to $N_{OH} \lesssim 14$. We investigated the dependence of relative OH/HI and OH/H_2 abundances on the variations of the physical conditions, i.e. the metallicity (Z), number density (n), cosmic ray ionization rate (ζ), and strength of UV field (χ) in the ISM. For the reasonable ranges of the physical conditions in diffuse ISM (within ± 1 dex of typical values), we showed that the abundances obtained using our simple prescription are in agreement with the calculation with the MEUDON PDR code, which utilizes the full reaction network to predict abundances of various species. We confirmed that as was found previously, OH is strongly enhanced in the presence of H_2. We also confirmed that to the first order the OH abundance depends on the combination of nZ/χ, while the scaling factor significantly increases as H_2 molecular fraction approaches unity, resulting in typical n_{OH}/n_H^{tot} values of $10^{-8}-10^{-7}$ in the H_2-dominated region of the cloud. This indicates that at current observational sensitivity limits of $N_{OH} \gtrsim 10^{13}$ cm^{-2}, the OH molecules probe the medium with H_2 column densities $N_{H_2} \gtrsim 10^{20}$ cm^{-2}, corresponding to the diffuse molecular ISM, where HI/H_2 transition is already complete. Additionally, we show that the absorption systems with N_{OH} in ranges $10^{13}-10^{14}$ cm^{-2} probe molecular gas before onset of CO, i.e. the 'CO-dark' molecular gas.

We found that the Galactic measurements of OH main lines from Li et al. (2018) can be reproduced by models with $n \sim 50$ cm^{-3}, $\chi \sim 1$ (Mathis field), and $\zeta \sim 3 \times 10^{-17}$ s^{-1}, with a variation of about 1 dex allowed around these values. The only available measurement of OH at $z > 0$ in the diffuse regime (log $N_{OH} \lesssim 14$) towards Q 0248+430 from Gupta et al. (2018b) indicates similar physical conditions.

Utilizing the observed H_2 column density distribution function at $z \sim 3$ (Balashev & Noterdaeme 2018), we derived the expected OH column density distribution function and incidence rate of OH absorbers. Both derived quantities include a scaling constant $C(z)$, which describes the possible redshift evolution of the cross-section of OH-bearing diffuse molecular gas. We apply measured $C(z \sim 3)$ directly to MALS (Gupta et al. 2016), the large survey project at the MeerKAT telescope, which will cover $0 < z < 1.87$ for the OH main lines. Considering only the central brightest source in the telescope's field of view, we estimate an OH incidence rate $dN/dX \approx 4.7^{+2.0}_{-2.5} \times 10^{-4}$, which indicates that MALS may only detect only a handful of OH absorbers from diffuse ISM. We discuss the caveats related to the extrapolation of $z \sim 3$ distribution function to the lower redshift. We suggest that the constraints on the OH based on this should be treated as a lower limit. The upcoming large radio absorption line surveys with SKA precursors and pathfinders will have the capacity to constrain $C(z)$ and inform the design of next generation surveys.

An ambitious SKA1 survey sensitive to detect $N_{OH} \sim 12$ and a total absorption line path in excess of 50 000, will have the prospect of detecting at least $\sim$100 OH absorbers and a realistic chance of constraining the column density distribution of diffuse OH. This will be highly complementary to the molecular emission line surveys and physical models of ISM which are mostly focused on the dense molecular gas.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Allison J. R. et al., 2020, MNRAS, 494, 3627
Balashev S. A., Kosenko D. N., 2020, MNRAS, 492, L45
Balashev S. A., Noterdaeme P., 2018, MNRAS, 478, L7
Balashev S. A. et al., 2017, MNRAS, 470, 2890
Balashev S. A. et al., 2019, MNRAS, 490, 2668
Barriault L., Joncas G., Lockman F. J., Martin P. G., 2010, MNRAS, 407, 2645
Bellomi E., Godard B., Hennebelle P., Valdivia V., Pineau des Forêts G., Lesaffre P., Pérault M., 2020, A&A, 643, A36
Bettens R. P. A., Hansen T. A., Collins M. A., 1999, J. Chem. Phys., 111, 6322
Bialy S., Sternberg A., 2015, MNRAS, 450, 4424
Bialy S., Sternberg A., 2016, ApJ, 822, 83
Bialy S., Sternberg A., 2019, ApJ, 881, 160
Bialy S., Bihr S., Beuther H., Henning T., Sternberg A., 2017, ApJ, 835, 126
Black J. H., Dalgarno A., 1973, ApJ, 184, L101
Bovino S., Tacconi M., Gianturco F. A., Galli D., 2011, A&A, 529, A140
Boyce P. J., Cohen R. J., 1994, A&AS, 107, 563
Brogan C. L. et al., 2013, ApJ, 771, 91
Caswell J. L., 1999, MNRAS, 308, 683
Chengalur J. N., Kanekar N., 2003, Phys. Rev. Lett., 91, 241302
Chengalur J. N., de Bruyn A. G., Narasimha D., 1999, A&A, 343, L79
Combes F., Gupta N., Jozsa G. I. G., Momjian E., 2019, A&A, 623, A133
Combes F. et al., 2021, A&A, 648, 12
Cotten D. L., Magnani L., Wennerstrom E. A., Douglas K. A., Onello J. S., 2012, AJ, 144, 163
Dame T. M., Hartmann D., Thaddeus P., 2001, ApJ, 547, 792
Darling J., Giovanelli R., 2002, AJ, 124, 100
Dawson J. R. et al., 2014, MNRAS, 439, 1596
Donate E., White J., Magnani L., 2019, MNRAS, 486, 4414
Draine B. T., 2011, Physics of the Interstellar and Intergalactic Medium, Princeton Univ. Press, Princeton, NJ
Draine B. T., Bertoldi F., 1996, ApJ, 468, 269
Engels D., Bunzel F., 2015, A&A, 582, A68
Epée Epée M. D., Mezei J. Z., Motapon O., Pop N., Schneider I. F., 2016, MNRAS, 455, 276
Federman S. R., Glassgold A. E., Kwan J., 1979, ApJ, 227, 466
Fernandez M. X., Momjian E., Salter C. J., Ghosh T., 2010, AJ, 139, 2066
Freundlich J. et al., 2019, A&A, 622, A105
Glassgold A. E., Langer W. D., 1974, ApJ, 193, 73
Grenier I. A., Casandjian J.-M., Terrier R., 2005, Science, 307, 1292
Grossmann V., Heithausen A., Meyerdierks H., Mebold U., 1990, A&A, 240, 400
Gry C., Boulanger F., Nehmé C., Pineau des Forêts G., Habart E., Falgarone E., 2002, A&A, 391, 675
Gupta N. et al., 2016, PoS, 14, 16
Gupta N. et al., 2018a, MNRAS, 476, 2432
Gupta N., Momjian E., Srianand R., Petitjean P., Noterdaeme P., Gyanchandani D., Sharma R., Kulkarni S., 2018b, ApJ, 860, L22

Gupta N. et al., 2020, ApJ, 907, 24

Habing H. J., 1968, Bull. Astron. Inst. Neth., 19, 421

Heays A. N., Bosman A. D., van Dishoeck E. F., 2017, A&A, 602, A105

Heiles C., Troland T. H., 2003a, ApJS, 145, 329

Heiles C., Troland T. H., 2003b, ApJ, 586, 1067

Herbst E., Klemperer W., 1973, ApJ, 185, 505

Hollenbach D., Kaufman M. J., Neufeld D., Wolfire M., Goicoechea J. R., 2012, ApJ, 754, 105

Jenkins E. B., 2013, ApJ, 764, 25

Jonas J., MeerKAT Team, 2016, PoS, 252, 1

Jones J. D. C., Birkinshaw K., Twiddy N. D., 1981, Chem. Phys. Lett., 77, 484

Jura M., 1975, ApJ, 197, 575

Kanekar N. et al., 2005, Phys. Rev. Lett., 95, 261301

Karpas Z., Anicich V., Huntress W. T., 1979, J. Chem. Phys., 70, 2877

Krogager J.-K., Noterdaeme P., 2020, A&A, 644, 6

Lanzetta K. M., Wolfe A. M., Turnshek D. A., Lu L., McMahon R. G., Hazard C., 1991, ApJS, 77, 1

Le Bourlot J., Le Petit F., Pinto C., Roueff E., Roy F., 2012, A&A, 541, A76

Le Petit F., Roueff E., Le Bourlot J., 2002, A&A, 390, 369

Le Petit F., Nehmé C., Le Bourlot J., Roueff E., 2006, ApJS, 164, 506

Leroy A. K. et al., 2009, AJ, 137, 4670

Li D. et al., 2018, ApJS, 235, 1

Liszt H., Lucas R., 1996, A&A, 314, 917

Liszt H., Lucas R., 1999, in Carilli C. L., Radford S. J. E., Menten K. M., Langston G. I., eds, ASP Conf. Ser. Vol. 156, Highly Redshifted Radio Lines. Astron. Soc. Pac, San Francisco, p. 188

Madau P., Dickinson M., 2014, ARA&A, 52, 415

Martinez Oscar J., Betts N. B., Villano S. M., Eyet N., Snow T. P., Bierbaum V. M., 2008, ApJ, 686, 1486

Mathis J. S., Mezger P. G., Panagia N., 1983, A&A, 500, 259

McBride J., Alatalo K., Nyland K., 2015, MNRAS, 447, 392

McCall B. J. et al., 2004, Phys. Rev. A, 70, 052716

McElroy D., Walsh C., Markwick A. J., Cordiner M. A., Smith K., Millar T. J., 2013, A&A, 550, A36

Mitchell J. B. A., 1990, Phys. Rep., 186, 215

Mitchell J. B. A. et al., 2005, J. Phys. B At. Mol. Phys., 38, L175

Morganti R., Sadler E. M., Curran S., 2015, PoS, 134, 18

Muller S. et al., 2011, A&A, 535, A103

Muller S. et al., 2014, A&A, 566, A112

Muzahid S., Srianand R., Charlton J., 2015, MNRAS, 448, 2840

Neufeld D. A., Wolfire M. G., 2017, ApJ, 845, 163

Noterdaeme P., Petitjean P., Ledoux C., Srianand R., Ivanchik A., 2008, A&A, 491, 397

Noterdaeme P., Petitjean P., Ledoux C., Srianand R., 2009, A&A, 505, 1087

Novotný O. et al., 2010, J. Phys. Chem. A, 114, 4870

Pei Y. C., Fall S. M., 1995, ApJ, 454, 69

Péroux C., Howk J. C., 2020, ARA&A, 58, 363

Pineda J. L., Langer W. D., Velusamy T., Goldsmith P. F., 2013, A&A, 554, A103

Prasad S. S., Huntress W. T. J., 1980, ApJS, 43, 1

Rebrion C., Rowe B. R., Marquette J. B., 1989, J. Chem. Phys., 91, 6142

Rosen S. et al., 2000, Faraday Discuss., 115, 295

Schilke P. et al., 2014, A&A, 566, A29

Smith D., Adams N. G., Miller T. M., 1978, J. Chem. Phys., 69, 308

Spirko J. A., Zirbel J. J., Hickman A. P., 2003, J. Phys. B At. Mol. Phys., 36, 1645

Srianand R., Gupta N., Rahmani H., Momjian E., Petitjean P., Noterdaeme P., 2013, MNRAS, 428, 2198

Stancil P. C., Lepp S., Dalgarno A., 1998, ApJ, 509, 1

Stancil P. C., Schultz D. R., Kimura M., Gu J. P., Hirsch G., Buenker R. J., 1999, A&AS, 140, 225

Sternberg A., Le Petit F., Roueff E., Le Bourlot J., 2014, ApJ, 790, 10

Storrie-Lombardi L. J., McMahon R. G., Irwin M. J., 1996, MNRAS, 283, L79

Tacconi L. J. et al., 2018, ApJ, 853, 179

te Lintel Hekkert P., Chapman J. M., 1996, A&AS, 119, 459

te Lintel Hekkert P., Versteege-Hensel H. A., Habing H. J., Wiertz M., 1989, A&AS, 78, 399

Theard L. P., Huntress W. T., 1974, J. Chem. Phys., 60, 2840

Villinger H., Futrell J. H., Howorka F., Duric N., Lindinger W., 1982, J. Chem. Phys., 76, 3529

Walsh A. J. et al., 2016, MNRAS, 455, 3494

Weingartner J. C., Draine B. T., 2001, ApJ, 563, 842

Weinreb S., Barrett A. H., Meeks M. L., Henry J. C., 1963, Nature, 200, 829

Wiklind T., Combes F., 1998, ApJ, 500, 129

Zheng Z., Li D., Sadler E. M., Allison J. R., Tang N., 2020, MNRAS, 499, 3085

Zwaan M. A., Prochaska J. X., 2006, ApJ, 643, 675

This paper has been typeset from a TeX/LaTeX file prepared by the author.