Learning from model grids: Tracers of the ionization fraction in the ISM

Emeric Bron1,* Evelyne Roueff1 Maryvonne Gerin1, Jérôme Pety2,1, Pierre Gratier3, Franck Le Petit1, Viviana Guzman4, Jan Orkisz5, Victor de Souza Magalhaes2, Mathilde Gaudel1, Pierre Palud6, Lucas Einig2,7, Sébastien Bardeau2, Maryvonne Gerin1, Pierre Chainais6, Jocelyn Chanussot1, Javier Goicoechea2, Annie Hughes7, Jouki Kainulainen7, David Languignon1, Jacques Le Bourlot1, François Levrier7, Derek Lis7, Harvey Liszt7, Karin Öberg7, Nicolas Peretto7, Antoine Roueff7, Albrecht Sievers1, Pierre-Antoine Thouvenin6, and Pascal Tremblin7

1LERMA, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Universités, France
2IRAM, 300 rue de la Piscine, 38406 Saint Martin d’Hères, France
3Laboratoire d’Astrophysique de Bordeaux, Univ. Bordeaux, CNRS, France
4Instituto de Astrofísica, Pontificia Universidad Católica de Chile, Chile
5Chalmers University of Technology, Department of Space, Earth and Environment, Sweden
6Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRISTAL, 59651 Villeneuve d’Ascq, France
7See https://www.iram.fr/pety/ORION-B/team.html for the other affiliations

Abstract. The ionization fraction in neutral interstellar clouds is a key physical parameter controlling multiple physical and chemical processes, and varying by orders of magnitude from the UV irradiated surface of the cloud to its cosmic-ray dominated central regions. Traditional observational tracers of the ionization fraction, which mostly rely on deuteration ratios of molecules like HCO\(^+\), suffer from the fact that the deuterated molecules are only detected in a tiny fraction of a given Giant Molecular Cloud (GMC). In [1], we propose a machine learning-based, semi-automatic method to search in a large dataset of astrochemical model results for new tracers of the ionization fraction, and propose several new tracers relevant in different ranges of physical conditions.

1 Context

The ionization fraction \(x(e^-)\) is a key physical parameter controlling several aspects of the physics and chemistry of neutral interstellar clouds. Fast ion-neutral reactions are one of the main drivers of interstellar chemistry. The ionization fraction also controls the coupling of the gas to the magnetic field, as well as the charge state of grains and thus the efficiency of the photoelectric heating mechanism on grains. Finally, it can affect our interpretation of the observed excitation of high dipole moment molecules such as HCN, which are very sensitive to collisional excitation by electrons. In a typical Giant Molecular Cloud, \(x(e^-)\) can vary in the neutral gas from \(~ 10^{-4}\) cm\(^{-3}\) at the UV illuminated surface of the cloud, where carbon is fully ionized, to \(10^{-6} – 10^{-9}\) cm\(^{-3}\) in the central shielded regions, where cosmic rays maintain a residual source of ionization.

*e-mail: emeric.bron@obspm.fr

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
Usual tracers of the ionization fraction use the molecular deuteration of species such as HCO\(^+\) [3] or N\(_2\)H\(^+\) [4] that indirectly trace the H\(_2\)D\(^+\)/H\(_2\)\(^+\) ratio, whose enrichment at low temperature is limited by electronic dissociative recombination and which is thus sensitive to the ionization fraction. However deuteration-based tracers of the ionization fraction suffer from several limitations. They are sensitive to the ortho-para ratio of H\(_2\), itself a parameter difficult to estimate observationally. They are also relatively insensitive at very low ionization fractions, yielding estimates with large uncertainties. More importantly, these deuterated species are typically only detected in dense cores, which represent only small fraction of the observable area of a GMC. In the context of the Orion-B IRAM-30m Large Program [7], which has mapped the Orion B GMC over 5 square degrees with an angular resolution of 26\(^\prime\) in the full 71 to 116 GHz band, we thus aim at finding new tracers of the ionization fraction to allow a full scale, unbiased view of the ionization fraction in a GMC, covering all the types of environments constituting a GMC.

2 Method

Massive datasets of model results produced by running large grids of models exploring wide ranges for multiple parameters and predicting tens or hundreds of observable quantities are now becoming common. These datasets contain a wealth of information on how sensitive the different observable quantities are to different unobservable physical quantities, and can thus help in finding new observable tracers of unobservable quantities such as \(x(e^-)\).

The possible relationship between one observable quantity (e.g. an integrated line intensity ratio) and one unobservable quantity of interest (e.g. the ionization fraction) will however be affected by other unobservable parameters (e.g. the gas density, temperature, UV field, etc...). These will act as sources of noise and cause a scatter in the relationship. The predictive power of an observable quantity to predict a given unobservable quantity (here \(x(e^-)\)) can thus be quantified as the magnitude of this scatter around the mean trend of the relationship. We now describe successively the two components needed for this purpose: a grid of models sampling randomly all of the important unobservable parameters over realistic ranges of values for the environment of interest, and a statistical method to find the mean trend of this relationship and to quantify the scatter around this mean trend.

**Table 1.** Ranges of parameters explored in the two model grids.

| Parameter | Translucent medium grid | Cold dense medium grid |
|-----------|--------------------------|------------------------|
| \(m_1\) [cm\(^{-3}\)] | \(3 \times 10^2 \sim 3 \times 10^7\) | \(10^7 \sim 10^8\) |
| \(T_{\text{gas}}\) [K] | 15 \(-\) 100 | 7 \(-\) 20 |
| \(G_0\) | 1 \(-\) 1000 | 1 |
| \(A_V\) | 2 \(-\) 6 | 5 \(-\) 20 |
| \(\zeta\) [cm\(^{-1}\)] | \(10^{-17} \sim 10^{-15}\) | \(10^{-17} \sim 10^{-16}\) |
| OPR\(_{\text{H}_2}\) | 0.1 \(-\) 3 | \(10^{-4} \sim 0.1\) |
| depletion factor | 1 | 1 \(-\) 10 |
| \([S]\) | \(1.86 \times 10^{-8} \sim 1.86 \times 10^{-5}\) | \(1.86 \times 10^{-8} \sim 1.86 \times 10^{-5}\) |

Using the astrochemical model presented in [8], we compute single zone models with fixed density and temperature at stationary state, including a chemical network of 310 species and 8711 chemical reactions. As we expect the best tracers of the ionization fraction \(x(e^-)\) to be different in different types of environment, we run two grids of 5000 models each, exploring ranges of parameter values relevant respectively for translucent medium and for cold dense medium (the ranges of parameters explored are listed in Table 1). We select a list of species based on the frequency range of the Orion-B Large Program : CS, SO, C\(_2\)H,
Multi-line Diagnostics of the Interstellar Medium

We select a set of 5000 models, from several limitations. They are sensitive to the ortho-para ratio of H$_2$, itself a parameter to the ionization fraction. However, deuterated species are not sensitive to the ionization fraction. Temperature is limited by electronic dissociative recombination and thus sensitive to the ionization fraction. Transition probabilities for deuterated species are thus not a very good tracer of the ionization fraction. However, deuterated species are sensitive to other unobservable parameters (e.g. the gas density, temperature, UV field, etc...). These will act as sources of noise and cause a scatter in the relationship. The $R^2$ value thus represents the fraction of the initial variance in the quantity of interest (here $x(e^{-})$) in the dataset that can be predicted from the fitted relationship with the single line ratio under consideration. We fit one Random Forest model and compute the corresponding $R^2$ coefficient for each line ratio (separately for each of the two model grids), and then rank the tracers based on the $R^2$ values. More details on how we trained the Random Forest models can be found in [1].

### 3 Results

The rankings of line ratios based on the $R^2$ value are shown on Fig. 1 (left panel: translucent medium conditions, right panel: cold dense medium conditions). Several line ratios can predict a large fraction of the variations of $x(e^{-})$ in the model grids. In the translucent case, most of the best ratios involve C$_2$H, which is known to be enhanced in UV-illuminated environments (e.g. [5, 6]). In translucent gas, UV photons are still the main source of ionization, and a tracer of the impact of the UV field such as C$_2$H can thus also trace accurately the ionization fraction.
fraction of the gas. In both types of conditions, we find very good tracers of the ionization fraction, that can give accurate estimates of the ionization fraction despite not knowing the density, temperature, UV illumination, ortho-para ratio of H$_2$, etc. For instance, Fig. 2 shows a scatter plot of the ionization fraction versus one of the best tracers in the translucent case: the C$_2$H (1-0)/$^{13}$CO (1-0) line ratio. The scatter induced by the 7 unknown parameters in the model grid remains very limited, as can be expected from the $R^2 > 0.9$ found. An ad hoc analytical fit, more easy to reuse than the Random Forest model, is also provided (black solid line, uncertainties as dashed lines). Complete results can be found in [1].

The method that we propose is very general and can be applied to find good tracers of any physical quantities (e.g., gas density, cosmic ray ionization rate) based on datasets of any types of more complex models (e.g., PDR models, astrochemical models including surface chemistry). A general implementation of this method is publicly available at http://autorank.ism.obspm.fr/.

References

[1] Bron E., Roueff E., Gerin M., Pety J., Gratier P., Le Petit F. et al., A&A 645 A28 (2021)
[2] Breiman, L., Mach. Learn. 45 5 (2001)
[3] Caselli, P., Walmsley, C. M., Terzieva, R., & Herbst, E., ApJ 499 234 (1998)
[4] Caselli, P., Planet. Space Sci. 50 1133 (2002)
[5] Gratier, P., Bron, E., Gerin, M., et al., A&A 599 A100 (2017)
[6] Pety, J., Teysier, D., Fossé, D. et al., A&A 435 885 (2005)
[7] Pety, J., Guzmán, V. V., Orkisz, J. H., et al., A&A 599 A98 (2017)
[8] Roueff, E., Loison, J. C., & Hickson, K. M., A&A 576 A99 (2015)
[9] van der Tak, F. F. S., Black, J. H., Schöier, F. L., Jansen, D. J. et al., A&A 468 627 (2007)