



**HAL**  
open science

# Semiclassical Estimates of Pressure-induced Line Widths for Infrared Absorption in Hot (Exo)planetary Atmospheres

Jeanna Buldyreva, Kathleen Stehlin, Sergei Yurchenko, Elizabeth Guest,  
Jonathan Tennyson

► **To cite this version:**

Jeanna Buldyreva, Kathleen Stehlin, Sergei Yurchenko, Elizabeth Guest, Jonathan Tennyson. Semiclassical Estimates of Pressure-induced Line Widths for Infrared Absorption in Hot (Exo)planetary Atmospheres. *The Astrophysical Journal Supplement Series*, 2025, 276 (1), pp.23. 10.3847/1538-4365/ad9b19. hal-04904362

**HAL Id: hal-04904362**

**<https://hal.science/hal-04904362v1>**

Submitted on 22 Jan 2025

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

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



Distributed under a Creative Commons Attribution 4.0 International License



# Semiclassical Estimates of Pressure-induced Line Widths for Infrared Absorption in Hot (Exo)planetary Atmospheres

Jeanna Buldyreva<sup>1</sup>, Kathleen Stehlin<sup>1</sup>, Sergei N. Yurchenko<sup>2</sup> , Elizabeth R. Guest<sup>2</sup>, and Jonathan Tennyson<sup>2</sup> <sup>1</sup> Institut UTINAM, UMR 6213 CNRS, Université de Franche-Comté, 16 Route de Gray, Besançon cedex, 25030, France; [jeanna.buldyreva@univ-fcomte.fr](mailto:jeanna.buldyreva@univ-fcomte.fr)<sup>2</sup> Department of Physics and Astronomy, University College London, Gower Street, London, WC1E 6BT, UK; [s.yurchenko@ucl.ac.uk](mailto:s.yurchenko@ucl.ac.uk)

Received 2024 July 06; revised 2024 November 26; accepted 2024 December 03; published 2025 January 9

## Abstract

Because of elevated temperatures and high fluxes of stellar radiation irreproducible in laboratory conditions, molecules and molecular ions found or expected in exoplanetary atmospheres are generally poorly characterized from the viewpoint of their spectroscopic line-shape parameters; in many cases, there are no data at all. Advanced theoretical approaches (classical, semiclassical, and quantum mechanical), without mentioning their high computational cost, are also impracticable due to the lack of potential energy surfaces. To fill this gap of crucially missing line-broadening parameters, we provide estimated values issued from a simple rotationally independent semiclassical expression. Only the index related to the leading long-range interaction term, molecular masses and kinetic diameters, as well as temperature are used as input parameters. A wide range of absorbers and perturbation by He, Ar, H<sub>2</sub>, N<sub>2</sub>, O<sub>2</sub>, CO, NO, CO<sub>2</sub>, H<sub>2</sub>O, CH<sub>4</sub>, and NH<sub>3</sub> as well as self-perturbation are considered. The explicit temperature dependence  $T^{-0.5}$  allows calculations to be limited to the single reference temperature of 296 K; for other temperatures a simple scaling can be used. The full set of line-broadening coefficients obtained with various possible values of kinetic diameters is provided by the new Collisional Line-broadening Parameters database, which is specifically designed for this purpose. “Midvalue” (or more recent) kinetic diameters are retained to create one-value line-broadening data required to populate the ExoMol database. A way to generate rotationally dependent line widths is proposed.

*Unified Astronomy Thesaurus concepts:* [Molecular data \(2259\)](#); [Molecular spectroscopy \(2095\)](#); [Collisional broadening \(2083\)](#); [Exoplanet atmospheres \(487\)](#); [Stellar atmospheric opacity \(1585\)](#)

## 1. Introduction

In the three decades since the first extrasolar planet was discovered (A. Wolszczan & D. Frail 1992) and the first detection of an exoplanet orbiting a standard star was reported (M. Mayor & D. Queloz 1995), we have learnt that such planets are ubiquitous, with nearly every star supporting a planetary system. Most planets detected so far are quite unlike those in our own solar system, and astronomers have taken the first steps in characterizing these exoplanetary atmospheres through spectroscopy. The need to perform spectroscopic studies of exoplanet atmospheres is driving the next-generation space missions, including the European Space Agency’s (ESA’s) Ariel mission, which will observe about 1000 exoplanets ranging from Jupiter and Neptune size down to super-Earth size at visible and infrared (IR) wavelengths, and the current program of the James Webb Space Telescope (NASA/ESA), which is intended to provide the first spaceborne glimpse of mid-IR exoplanetary spectra at a resolving power that ranges from low to moderate (from around 100 to a few thousand). However, complete interpretation and exploitation of observation results require extended “reference” spectroscopic parameters for various molecules whose presence is already confirmed or expected in exoplanetary atmospheres.

Among these parameters, those related to pressure broadening of spectral lines are known to have important consequences for radiative transport (see, e.g., G. Tinetti et al. 2012), and

pressure-broadening data were identified as the most important current data needs in the recent NASA white paper by J. J. Fortney et al. (2019). Recent studies devoted to radiative-transfer modeling of exoplanetary atmospheres (see, e.g., C. Hedges & N. Madhusudhan 2016; K. L. Chubb et al. 2021; E. Gharib-Nezhad & M. R. Line 2019) have assessed the impact of pressure broadening. The importance and status of the laboratory broadening data for exoplanetary missions were recently discussed in a review by K. L. Chubb et al. (2024). Whereas line broadening has been well studied for the conditions found in the terrestrial atmosphere (data collected in spectroscopic databases such as HITRAN; I. E. Gordon et al. 2022), elevated temperatures and high fluxes of stellar radiation (characteristic for exoplanets but irreproducible in laboratory conditions) result in a dearth of data for most species, both radiators and broadeners, found in atmospheres of exoplanets.

The specific “exoplanet-oriented” ExoMol database, created to provide all the spectroscopic data required for modeling spectra of atmospheres of exoplanets and cool stars (as well as for other atmospheric, astrophysical, laboratory, or industrial applications; J. Tennyson & S. N. Yurchenko 2012), currently hosts line lists giving access to line positions and intensities for 91 molecules and 224 isotopologues (J. Tennyson et al. 2024). Some of the lists are huge, containing billions of lines; additional molecular data such as opacities (K. L. Chubb et al. 2021), absorption cross sections (C. Hill et al. 2013), photoabsorption and photodissociation cross sections (M. Pezzella et al. 2021), lifetimes (J. Tennyson et al. 2016), cooling functions, specific heats (R. Wang et al. 2023), and partition functions are available. However, this huge amount of data is of limited usefulness without information on line broadening. This becomes especially pressing for studies of secondary



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

atmospheres (L. Anisman et al. 2022), where data will be necessary for a variety of broadening species (such as CO<sub>2</sub>, H<sub>2</sub>O, N<sub>2</sub>, and CO, etc.) and the advent of high-resolution techniques for exoplanetary observations. Despite some progress achieved using the novel method based on machine learning (ML; E. R. Guest et al. 2024) in the provision of air-broadening data for molecules in the ExoMol database, these line-broadening data are incomplete for He and H<sub>2</sub> or largely nonexistent for any other important broadeners.

The present paper aims to address this issue and to start a systematic production of the line-broadening parameters with a final goal of filling the ExoMol database. As advanced theoretical methods require intermolecular potential energy surfaces currently unavailable for exomolecules, as a first “guess-value” step, we employ a simple but powerful semiclassical approach suggested recently (J. Buldyreva et al. 2022) for producing rotationally independent pressure-broadened line widths for (vib)rotational transitions applicable for a large variety of collision partners. Since the current database structure of ExoMol is not capable of coping with the full complexity as well as the magnitude of the line-shape data, we intend first to develop a prototype of a line-shape database for exoplanetary atmospheric studies that will underpin the next great advance in our understanding of exoplanets. The prototype and the underlying data structure will be tested and then connected to the ExoMol website and other databases. As a possible way to getting rotational line-width dependencies, the ML approach is examined through comparisons of its air-broadening results with the constant-value semiclassical estimates.

## 2. Pressure-broadening and Related Data in ExoMol

Molecules in an exoplanet atmosphere, besides their thermal motion, which gives a Gaussian profile due to the Doppler effect, are subject to constant collisions with other species resulting in pressure broadening that is represented by a Lorentz profile. Under the assumption of statistical independence of these two broadening mechanisms, the resulting shape of a spectroscopic line is given by a convolution of the corresponding profiles (Voigt profile). Whereas the temperature-dependent Doppler broadening coming from the thermal agitation can be easily evaluated from the mass of the active molecule and the line-position frequency, pressure broadening is due to collisions and relies on the specifics of the particular intermolecular interactions; the line-broadening parameters are therefore much more difficult to evaluate.

Until recently, the majority of available line-broadening data came from the HITRAN database (I. E. Gordon et al. 2022): the best represented are the terrestrial atmospheric molecules with the air being the major broadener (self-broadening parameters are given too). Although HITRAN now provides some pressure-broadening parameters for collisions with He, H<sub>2</sub>, and CO<sub>2</sub> (J. S. Wilzewski et al. 2016; Y. Tan et al. 2022), important for spectroscopy of hydrogen-dominated atmospheres, the majority of the diatomics, especially transient and with high boiling temperatures, have minimal or no broadening data in ExoMol (nor in any other spectroscopic database, nor in the literature). Very recently, E. R. Guest et al. (2024) used state-of-the-art ML techniques to produce air-broadening data for a large number of molecules in ExoMol using HIRTRAN as the source of the training data. “Beyond-air” broadeners still have not been systematically addressed.

As billions of transitions are often required for the high-temperature studies of exoplanetary spectroscopy (S. N. Yurchenko et al. 2014), a special state-oriented format is adopted for the ExoMol line lists, ensuring a more compact data storage with respect to the line-by-line format commonly used in HITRAN (I. E. Gordon et al. 2022) or GEISA (T. Delahaye et al. 2021). Namely, the ExoMol line-list format is based on two files: a “States file” containing the descriptions of the molecular states (energies, statistical weights, and quantum numbers, etc.) and a “Transition file” containing the Einstein A coefficients (J. Tennyson et al. 2013). Within the ExoMol data structure, the presentation of line-broadening data in ExoMol differs from that used by HITRAN and GEISA.

The line-broadening values in the ExoMol database are presented in a form of the ExoMol Diet (E. J. Barton et al. 2017). Namely, pressure-broadening line widths  $\gamma_{\text{ref}}$  (half-width at half-maximum, in cm<sup>-1</sup>, at the reference temperature of  $T_{\text{ref}} = 296$  K and the reference pressure of  $P_{\text{ref}} = 1$  atm) and the corresponding temperature-dependence exponents  $n$  are provided. The line width  $\gamma(T, P)$  at a given temperature  $T$  and a given pressure  $P$  can be retrieved then by the commonly used power law as given by

$$\gamma(T, P) = \gamma_{\text{ref}} \frac{P}{P_{\text{ref}}} \left( \frac{296}{T} \right)^n. \quad (1)$$

Unlike the HITRAN database structure, where line-broadening parameters  $\gamma_{\text{ref}}$  and  $n$  are provided in the line-by-line manner for each transition individually, the ExoMol Diet organizes  $\gamma_{\text{ref}}$  and  $n$  based on the state quantum numbers. For a given collision pair “active molecule” + “perturber” (e.g., H<sub>2</sub>O and H<sub>2</sub>),  $\gamma_{\text{ref}}$  and  $n$  are collected in a separate ASCII file (e.g., 16H2-O\_H2.broad) with a set of quantum numbers (e.g., the rotational angular momentum quantum number  $J$ , its projection  $K_a$  and  $K_c$  on the molecular axes, vibrational quantum numbers  $v$ , and the rotational index  $m_J$ ; see J. Tennyson et al. 2024). In principle, there are no restrictions on the choice of quantum numbers to be employed for this purpose. However, the usual preference is using the single quantum number  $J$  of the lower state (commonly denoted in the spectroscopic literature by  $J''$ ) representing the most important dependence of the pressure-broadening parameters. This case is labeled as “a0.” It is also common in the line-broadening literature to use the rotational index  $m_J$  that is defined as  $-J, J, J + 1$  (for  $J$  lower and the P, Q, and R branches, respectively) as an alternative for  $J$ . This case of a single lower-state rotational dependence is represented by the label “m0.”

Such  $J$ -based cases are also most straightforward to use in connection with the ExoMol line-list format, since  $J$  is a compulsory quantum number descriptor of the States file, while other (especially nonrigorous) quantum numbers are not always well defined or even provided. This is because the nonrigorous quantum numbers associated with theoretical line lists (i.e., most of ExoMol) in most cases can only be defined approximately, if at all. Moreover, their definition becomes increasingly ambiguous with excitations; see, e.g., the discussion and associated review in E. K. Conway et al. (2021). This includes the rotational quantum numbers  $K$ ,  $K_a$ , and  $K_c$ , etc., which are responsible for the so-called “branch” dependence of line-broadening parameters. While the “a0” classification ignores completely the “branch” dependence, the “m0” case can be in principle used to recover this important information. In any case, the “branch” dependence, pronounced typically at

**Table 1**  
Extract from a Diet File for H<sub>2</sub>O Broadened by H<sub>2</sub>

Case Label	$\gamma_{\text{ref}}$ (cm <sup>-1</sup> )	$n$	$J$
a0	0.0916	0.790	0
a0	0.0880	0.684	1
a0	0.0824	0.602	2
a0	0.0759	0.542	3
a0	0.0691	0.495	4
a0	0.0635	0.458	5
a0	0.0590	0.426	6
a0	0.0558	0.400	7
a0	0.0526	0.373	8
a0	0.0496	0.350	9
a0	0.0463	0.327	10
a0	0.0433	0.310	11
a0	0.0405	0.291	12
a0	0.0382	0.278	13

small  $J$  values, remains minor (see, as one of numerous examples, experimental and theoretical line widths for pure CO<sub>2</sub> bands; L. Daneshvar et al. 2014) and can be completely neglected for practical purposes of providing rotationally independent estimates. By default, the  $\gamma_{\text{ref}}$ ,  $n$  entry corresponding to the highest given  $J = J_{\text{max}}$  should be used for all  $J > J_{\text{max}}$ . The bare minimum of a Diet file is a single entry for  $J = 0$  with no quantum number dependence. Thus, using  $J = 0$  in the ExoMol Diet as the reference entry will set the same value of  $\gamma_{\text{ref}}$  to all  $J > 0$  transitions. As an example of a Diet file for H<sub>2</sub>O broadened by H<sub>2</sub>, an extract from 16H2-O\_H2.broad is given in Table 1. Finally, we note that in ExoMol it is generally assumed that minor isotopologues that have the same symmetry as the parent isotopologue share the same line-broadening parameters.

### 3. Line-width Estimates for Exomolecules

As mentioned in the previous section, the weak coverage of line broadening by various collision partners for molecules important for exoplanetary studies reflects the extreme difficulty of both experimental and theoretical approaches to generating these data. Accurate laboratory measurements are only realistic for a limited number of lines, making an experimental description of the line lists with tens of billions of lines a hopeless task, especially for radicals and molecular ions. An accurate theoretical description from the first principles is currently also limited to a small number of simple molecules; see, e.g., P. Wcisło et al. (2021). Therefore, the only resort is to use approximate methods such as ML (E. R. Guest et al. 2024), requiring significant quantities of training data, or cost-effective semiclassical procedures, such as proposed previously by us (J. Buldyreva et al. 2022). Note that these two approaches are not completely independent as the cited ML study used the parameters employed in our approximate model—see Equation (2) below—as training features.

In a recent study (J. Buldyreva et al. 2022), we explored a simple semiclassical estimate, initially suggested by C. Tsao & B. Curnutte (1962), for a production of pressure-broadened line widths and justified its use by comparison with available measurements. Therefore, for “exotic” active molecules with unknown intermolecular interaction potential energy surfaces and, as a consequence, impossibility of advanced theoretical calculations, this simple formula is expected to provide the best

**Table 2**  
Collision Partners Considered in this Work to Produce (Rotationally Independent) Line-broadening Coefficients  $\tilde{\gamma}$  Using the Semiclassical Method by J. Buldyreva et al. (2022)

Active Molecules				Perturbers	
AlCl	CP	MgH	PO <sub>2</sub>	TiO	self
AlH	CrH	MgO	PS	VO	Ar
AlO	CS	NaCl	ScH	ZnS	CH <sub>4</sub>
AsH <sub>3</sub>	FeO	NaF	SH	...	CO
BeH	KCl	NaH	SiC	...	CO <sub>2</sub>
C <sub>3</sub>	KF	NaOH	SiH	...	H <sub>2</sub>
CaF	KOH	NH	SiH <sub>2</sub>	...	H <sub>2</sub> O
CaH	LiCl	NS	SiH <sub>3</sub>	...	He
CaO	LiF	PF <sub>3</sub>	SiO	...	N <sub>2</sub>
CaOH	LiH	PH	SiO <sub>2</sub>	...	NH <sub>3</sub>
CH <sub>2</sub>	LiH <sup>+</sup>	PN	SiS	...	NO
CH <sub>3</sub>	MgF	PO	TiH	...	O <sub>2</sub>

guess values. In this approach, a rotationally independent line-broadening coefficient  $\tilde{\gamma}$  (in cm<sup>-1</sup> atm<sup>-1</sup>),

$$\tilde{\gamma} = 1.7796 \times 10^{-5} \frac{m}{m-2} T^{-0.5} m^{* -0.5} d^2, \quad (2)$$

is calculated from the index  $m$  corresponding to the leading long-range intermolecular interaction ( $m = 2(q - 1)$ , where  $q$  is the power of the inverse intermolecular distance dependence), the temperature  $T$  considered (in kelvin), the reduced molecular mass  $m^* = m_a m_p / (m_a + m_p)$  ( $m_a$  and  $m_p$  being the masses of the active and perturbing molecules in dalton), and the molecular pair’s kinetic diameter  $d = (d_a + d_b)/2$  (in picometers). When applied to NO and OH colliding with noble gases and nonpolar molecules, good consistency with available measurements was obtained over the temperature range of 200–3000 K (J. Buldyreva et al. 2022). The temperature dependence of this expression is consistent with Equation (1) ( $n = 0.5$ ), which lets us define a single parameter  $\gamma_{\text{ref}}$  for the reference temperature of 296 K and the reference pressure of 1 atm.

It should be noted that it is not uncommon to use a generic value for the temperature exponent  $n$  as in our Equation (2) when no laboratory data on the temperature dependence of the line broadening are available. For example, C. M. Sharp & A. Burrows (2007) adopted the value of  $n = 0.4$  for their molecular opacities for brown dwarfs and giant planet atmospheres. Similar to us, E. Gharib-Nezhad et al. (2021) relied on the classical collision theory by P. W. Anderson (1949) and applied the temperature exponent of  $n = 0.5$  when computing their EXOPLINES opacities. The exponent of  $n = 0.5$  was also one of the default values in the production of the ExomolOP opacities (K. L. Chubb et al. 2021). We recognize that the actual temperature dependence can significantly differ from such simplistic models, but the value of  $n = 0.5$  seems to be a reasonable choice when no data is available.

To start producing systematic data for “exotic” molecular species, we selected optically active molecules with confirmed or expected presence in exoplanetary atmospheres but without line-broadening data available (Table 2); the broadeners were chosen because of their important abundances. The main difficulty of applying Equation (2) to the “exotic” molecular pairs is the lack of the kinetic diameters  $d$  in the literature; even the Lennard-Jones parameters  $\sigma_{LJ}$  that could replace  $d$  were not



**Table 3**  
Line-broadening Coefficients  $\tilde{\gamma}$  (in  $\text{cm}^{-1} \text{atm}^{-1}$ ) at the Reference Temperature of 296 K for the Active Molecule  $\text{SiO}_2$  ( $m_a = 60.09$  Da)

Perturber	$m$	$m_p$	$d_p$	$\tilde{\gamma}(443.53)$	$\tilde{\gamma}(414.04)$	$\tilde{\gamma}(436.76)^*$	$\tilde{\gamma}(419.39)$
$\text{SiO}_2$	6	60.09	$=d_a$	0.0557	0.0485	0.0540	0.0498
He	10	4.00	258	0.0821	0.0754	0.0806	0.0766
Ar	10	39.95	333	0.0398	0.0368	0.0391	0.0374
$\text{O}_2$	6	32.00	346	0.0529	0.0490	0.0520	0.0497
$\text{H}_2$	6	2.02	292	0.1501	0.1383	0.1474	0.1404
$\text{N}_2$	6	28.02	368	0.0584	0.0543	0.0575	0.0550
CO	4	28.01	376	0.0795	0.0739	0.0782	0.0749
$\text{CO}_2$	6	44.01	330	0.0460	0.0426	0.0452	0.0432
NO	4	30.01	317	0.0669	0.0618	0.0657	0.0627
$\text{H}_2\text{O}$	4	18.02	265	0.0697	0.0640	0.0684	0.0651
$\text{NH}_3$	4	17.04	260	0.0703	0.0645	0.0689	0.0655
$\text{CH}_4$	10	16.05	380	0.0616	0.0573	0.0606	0.0580

**Note.** The  $d_a$  values (in picometers) coming from four polarizability values found in the literature and used for calculations are indicated as arguments in the column headings. The values in the column marked by an asterisk are the closest to the averaged values and are retained for the use in ExoMol.

found. Therefore, for nearly all active molecules listed in Table 2, the required  $d_a$  values were estimated using a semiempirical formula (B. I. Loukhovitski & A. S. Sharipov 2021) from the molecular polarizabilities.

Another point to be mentioned is the choice of the leading-interaction parameter  $m$ . As our data are intended to be used in IR absorption spectra (i.e., generally for vibration-rotation transitions), we assumed that IR active vibrational modes are concerned and that the leading intermolecular interaction corresponds to the active molecule's geometry in this mode. Namely, among the active molecules in ExoMol with completely unknown line-broadening parameters, the polyatomics  $\text{C}_3$  and  $\text{SiO}_2$  (having (near) zero dipole moments in their equilibrium configuration) were considered as vibrating in an IR active mode, creating an instantaneous dipole moment interacting with the leading perturber's multipole. We also disregarded homonuclear diatomics such as  $\text{C}_2$ . Broadeners were assumed to be in their equilibrium geometry. In such a way, the possible  $m$  values could be limited to 4, 6, and 10 (dipole-dipole, dipole-quadrupole, and dipole-induced dipole interactions, respectively).

Table 3 gives an example of line-broadening coefficients produced for  $\text{SiO}_2$ . Because of various polarizability values found in the literature for the active molecule and different resulting  $d_a$  values, four  $\tilde{\gamma}$  estimates were obtained. The full set of possible  $\tilde{\gamma}$  values for each molecular pair and related references are accessible from the prototype COLLISIONAL LINE broadening parameters (COLLINE) database<sup>3</sup> described in the next section. To choose just one estimate to be included in ExoMol, we calculated the average  $\langle d_a \rangle$  value and, from the set of available  $d_a$  values, took that closest to  $\langle d_a \rangle$ ; in Table 3 the column with the selected line-broadening coefficients is marked by an asterisk. (For the active molecules with two possible  $d_a$  values, the one coming from the more recent polarizability value was selected.) Details on the ExoMol format update for line-broadening data are discussed below.

Figure 1 shows the spread of the line-broadening coefficients  $\tilde{\gamma}$  calculated in the present work for the active species included in ExoMol and perturbers of Table 2. The diameter and the colors of the circles represent the magnitude of the  $\tilde{\gamma}$  values.

An illustration of the temperature dependence of  $\text{SiO}$  line widths (at the reference pressure of 1 atm) calculated in the

present work is given in Figure 2 for several perturbers as well as their impact on the spectrum of  $\text{SiO}$  ( $T = 3000$  K and  $P = 1$  atm). It is seen that the collisional partner strongly influences the broadening and will have a significant impact on the collisional cross-section calculations.

## 4. Providing New Data

### 4.1. COLLINE Database

The limited amount of information provided by the current ExoMol format a0 initiated reflection on a more complete source of line-broadening data for “exotic” collision pairs. As different values of the kinetic diameter (Lennard-Jones parameter  $\sigma_{\text{LJ}}$  or polarizability) for the same active molecule were available from different sources and led to multiple possible values of the pressure-broadening coefficients, a specific database containing all possible estimates was created (COLLINE).<sup>3</sup> It references the input parameters for the bulk of active and perturbing species (so that the user can track where the data come from) and for each active molecule provides tables of possible pressure-broadening coefficients at 296 K. Among 51 active molecules listed in Table 2, 45 are “ExoMol molecules,” i.e., species included into the ExoMol database<sup>4</sup> with line lists provided; work on a further six molecules ( $\text{PO}_2$ ,  $\text{SiC}$ ,  $\text{FeO}$ ,  $\text{SiH}_3$ ,  $\text{CH}_2$ , and  $\text{ZnS}$ ) is currently in progress.

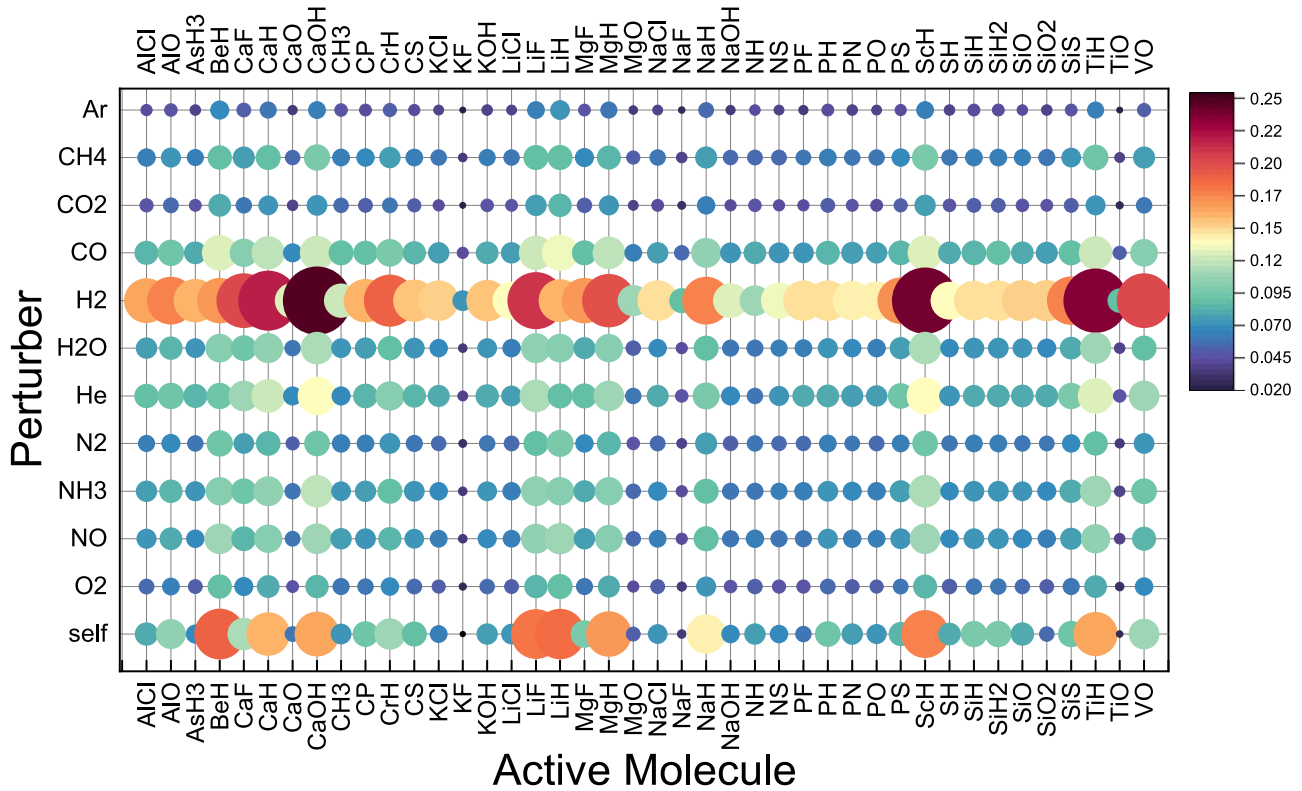
As follows from Equation (2), line-broadening estimates for a desired temperature  $T$  can be obtained by a simple scaling of the reference temperature value  $\tilde{\gamma}(296)$ :

$$\tilde{\gamma}(T) = \tilde{\gamma}(296) \left( \frac{296}{T} \right)^{0.5}. \quad (3)$$

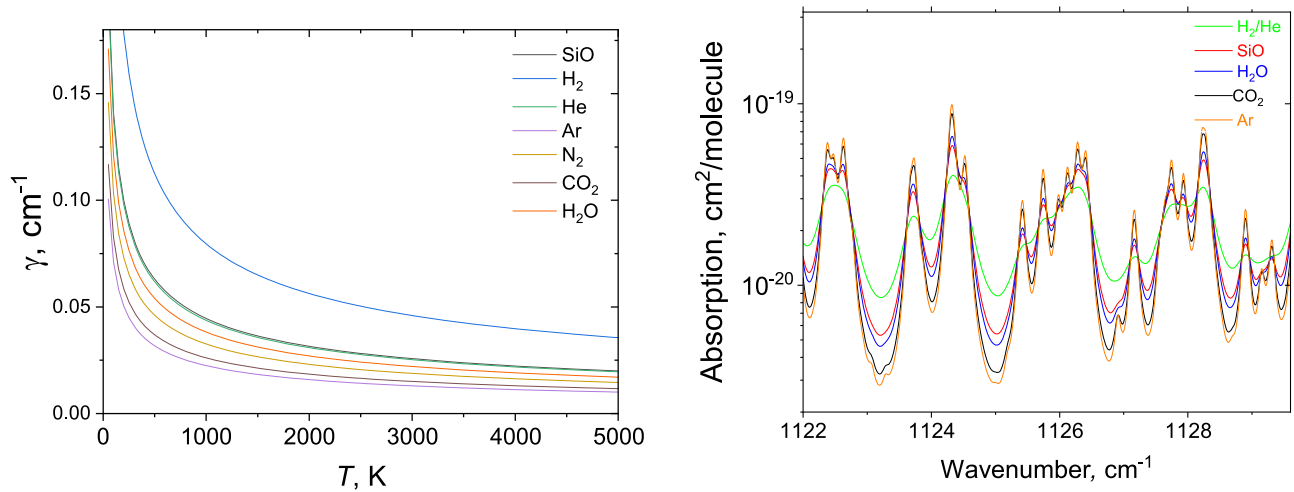
This equation is expected to work well for high  $T$ -values since the semiclassical description works better with temperature increase. The low-temperature limit  $T_{\text{min}}$  (lowest temperature for which the trajectories remain open) is different for each molecular pair and can be roughly evaluated as the depth of the isotropic intermolecular interaction potential expressed in kelvin. If the temperature (kinetic energy of the relative molecular motion) is above  $T_{\text{min}}$ , the molecules remain separated after collision (giving an open trajectory); conversely low-energy shocks can lead to a quasi-complex formation

<sup>3</sup> <https://colline.u-bourgogne.fr/>

<sup>4</sup> <https://www.exomol.com>



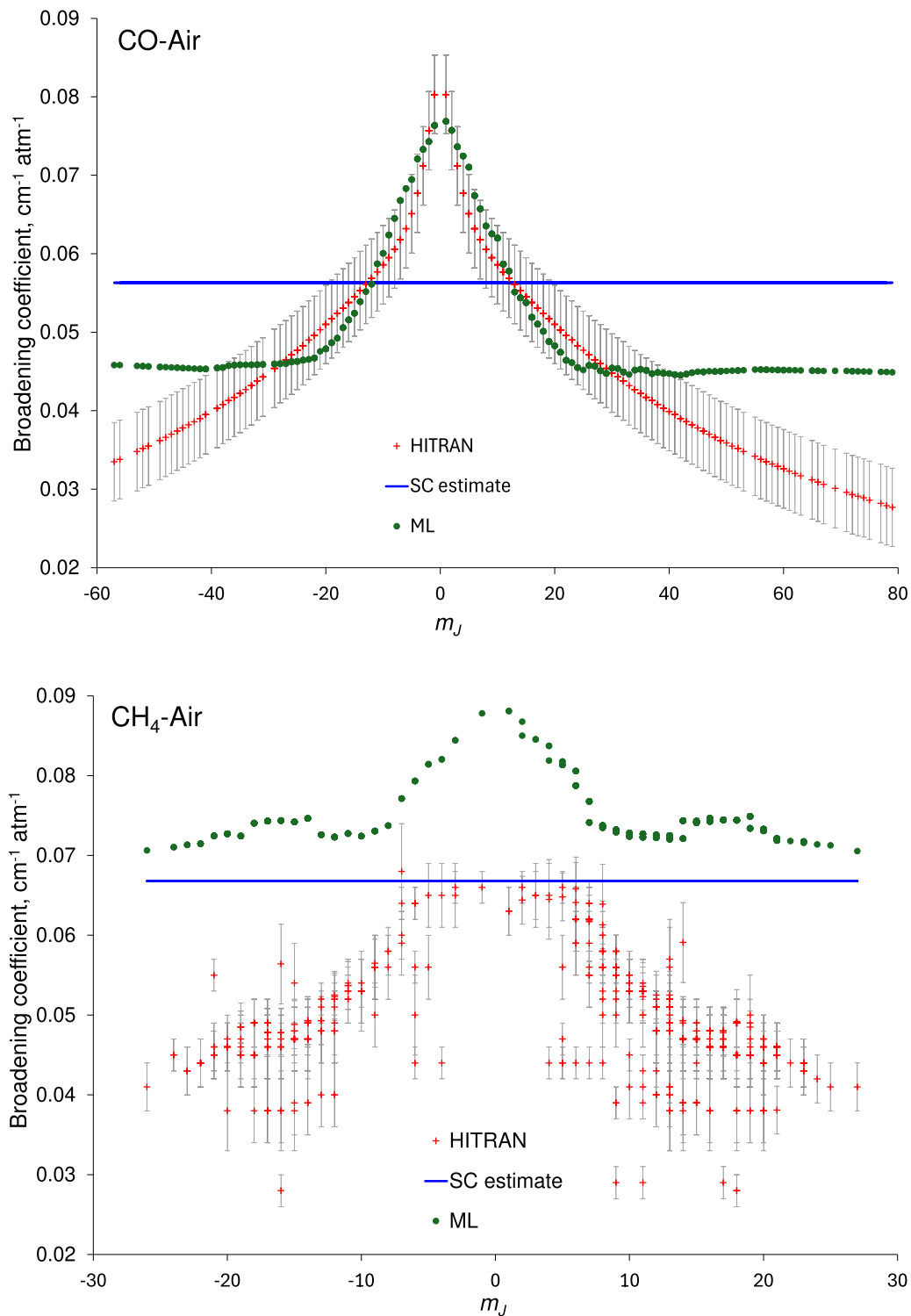
**Figure 1.** The spread of the line-broadening coefficients  $\tilde{\gamma}$  ( $\text{cm}^{-1} \text{atm}^{-1}$ ) for the “exotic” molecules included in ExoMol and collision partners considered in this work. The sizes of the circles and the color heatmap represent the magnitude of  $\tilde{\gamma}$ : the sizes are proportional to its value, while the heatmap is specified in the figure.



**Figure 2.** Left display: temperature dependence of the SiO line broadening for some perturbers considered in the present work. Right display: absorption spectrum of SiO computed for  $T = 3000$  K and  $P = 1$  atm.

(orbiting trajectory) in which case the semiclassical formulae do not work. The intermolecular potentials are however generally not well known for the active species considered (kinetic diameters in most cases were obtained from the molecular polarizabilities and not from the Lennard-Jones  $\sigma_{LJ}$  parameters coupled to the isotropic-potential depth  $\varepsilon_{LJ}$ ). Therefore, a separate study has been conducted to estimate low-temperature limits  $T_{\min}$  for the temperatures  $T$  to be used in Equation (3). Typically, X–He ( $X = \text{active molecules of Table 2}$ ) potential energy surfaces were sought in the literature with the aim of deducing the isotropic component depth; then,

using the isotropic-potential depth for He–He and the combination rule  $\varepsilon_{X-\text{He}} = (\varepsilon_{X-\text{X}}\varepsilon_{\text{He-He}})^{1/2}$ , the X–X isotropic-potential depth was obtained. The isotropic-potential depths and the related references for the active and perturbing molecules are given on the COLLINE web page. The  $T_{\min}$  values are provided in the last column of the data files for each active species referenced in COLLINE. For the majority of the collision partners considered,  $T_{\min}$  values lie far below the reference temperature of 296 K, so that the line-broadening estimates provided for 296 K are physically meaningful and can be scaled to any desired temperature  $T > T_{\min}$ . However,



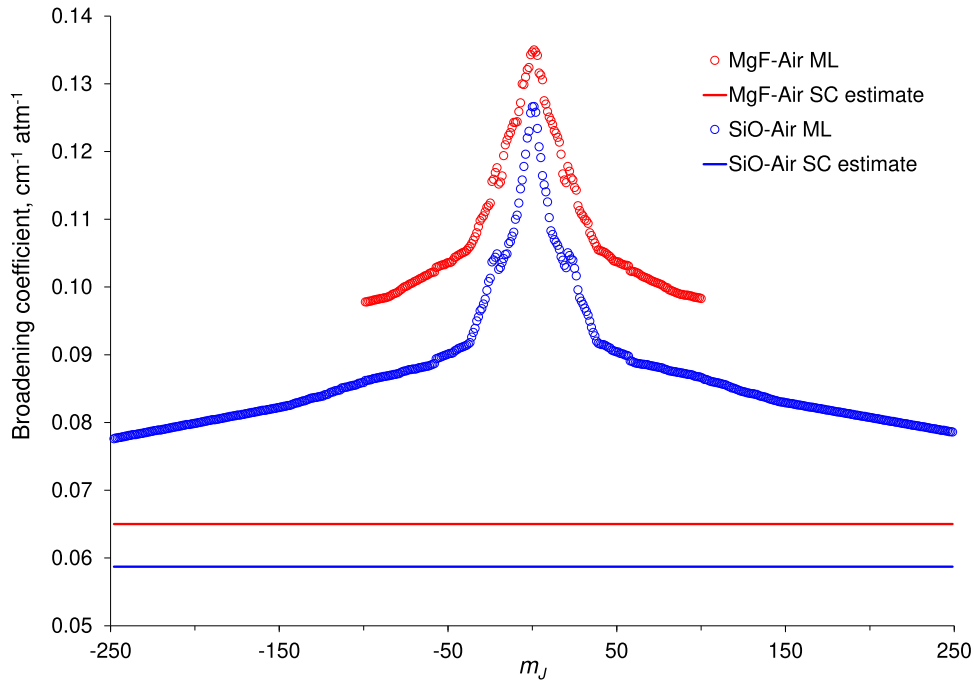
**Figure 3.** Comparison of air-broadening coefficients  $\gamma_{\text{ref}}$  obtained by the ML approach with HITRAN values (I. E. Gordon et al. 2022) and semiclassical (SC) estimates  $\tilde{\gamma}$  of the present work, shown for CO (upper panel) and CH<sub>4</sub> (lower panel).

for some collision partners (KCl with KCl, CO<sub>2</sub>, H<sub>2</sub>O, NH<sub>3</sub>, and self-perturbed LiF and MgO), the information found in the literature on the isotropic interaction potentials suggests  $T_{\text{min}}$  values higher than 296 K (see files for these active species), and the line-broadening coefficients provided are valid only for  $T > T_{\text{min}} > 296$  K. It should be also noted that the use of the broadening coefficients at  $T < T_{\text{min}}$  will lead to an underestimation of the real broadening as contributions from

molecular quasi-complexes are not accounted for. However, species such as KCl, LiF, and MgO usually only occur in the gas phase at temperatures well above 296 K.

#### 4.2. ExoMol Update

Exploiting semiclassically estimated pressure-broadening line widths obtained for “exotic” molecular pairs, the corresponding new ExoMol Diet files (E. J. Barton et al. 2017) were created for



**Figure 4.** Line-broadening coefficients  $\gamma_{\text{ref}}$  obtained by the ML approach (E. R. Guest et al. 2024) and the semiclassical (SC) estimates  $\tilde{\gamma}$  of the present work for MgF and SiO broadened by air.

528 “absorber”–“perturber” systems representing 44 molecules available in the ExoMol database and the 12 perturbers considered here; see Figure 1. Because of the rotational dependence ignored, these new data files contain a single entry (no dependence on quantum numbers), and the last column value  $J''$  is arbitrarily set to zero. These values are also included in the ExoMol definition .def files (see J. Tennyson et al. 2024) for all line lists corresponding to a given molecule.

### 5. Prospective of Getting $J$ Dependence for Line-broadening Data

The sheer volume of line-broadening data required by the ExoMol project, and other projects concerned with providing molecular line lists for exoplanetary environments, has led us to explore other means of generating the data. ML provides one possible methodology. Recently E. R. Guest et al. (2024) used ML models to fit to air-broadening data from the HITRAN database (I. E. Gordon et al. 2022). They used an ensemble learning model to develop a cheap method for the large-scale production of pressure-broadening parameters  $\gamma_{\text{ref}}$ , which proved to be reasonably accurate for active molecules not included in the training set: 69% of the predicted widths were found to lie within the uncertainties given by the HITRAN database. A particular feature of this model was that in nearly all cases the shape of the rotational dependence of the line-broadening parameter was well reproduced and, importantly for applications for hot molecules, provided stable values of  $\gamma$  at high  $J$ . Figure 3 provides an illustration of ML-generated broadening coefficients  $\gamma_{\text{ref}}$  of CO and CH<sub>4</sub> (molecules of importance to the exoplanetary community) broadened by air. In the case of CO, ML values reproduce HITRAN’s data within their uncertainty up to  $|m_J| \approx 40$ . For CH<sub>4</sub>, ML results produce a good  $J$  dependence but show a systematic offset; in this case using a semiclassical estimate for scaling may give more accurate results. Concerning “exotic” active species, in Figure 4, we show results for  $\gamma_{\text{ref}}$  of MgF and SiO broadened

by air from the ML calculations (E. R. Guest et al. 2024) compared to the corresponding semiclassical estimates  $\tilde{\gamma}$ . For MgF as well as for SiO, the ML results lie systematically higher than the semiclassical values, but the higher value of the broadening coefficient for MgF lines is obtained with both theoretical methods. We cannot, however, declare if any of these theoretical data sets are better (no measurements or other independent computations are available).

The plan will be to use the model of E. R. Guest et al. (2024) or a generalization of it to consider other broadening species to provide the  $J$  dependence of the line-broadening coefficients. We note that for simplicity E.R. Guest et al. (2024) neglected the dependence of  $\gamma_{\text{ref}}$  on the vibrational state of the active molecule.

### 6. Conclusion

A lack of adequate procedures to measure or compute broadening parameters for the billions of molecular transitions capable of absorbing radiation in exoplanetary and other hot atmospheres sets severe limits on the quality of the atmospheric modeling and retrievals. To overcome this difficulty and ensure data completeness, we have produced rotationally independent semiclassical estimates of line-broadening coefficients for 51 active molecules and 12 main perturbers detected or expected in the atmospheres of exoplanets. The bulk of new data has been placed in the prototype database COLLINe for testing by the exoplanetary community. One value for each molecular pair has been selected and included in the ExoMol database.

These reference temperature (296 K) values can be easily scaled to higher temperatures. It should be kept in mind that minimal-temperature limits exist and are different for each molecular pair. Below these minimal temperatures the semiclassical formula becomes invalid because of orbiting trajectories associated with molecular complexes formation. For the large majority of molecular pairs considered, however, the minimal temperature lies below the reference temperature of 296 K so that the provided reference temperature broadening coefficients are



meaningful and can be safely used for scaling to  $T > T_{\text{ref}} > T_{\text{min}}$ . If the reference temperature line-broadening coefficient is scaled to a temperature lower than  $T_{\text{min}}$ , line broadening will be underestimated since contributions from orbiting collisions are neglected in our semiclassical theoretical approach.

As a next step, we plan to extend our calculations of semiclassical estimates to the parent isotopologues of the molecular species considered in the ExoMol database. To go further and produce line-broadening coefficients with pronounced dependencies on the rotational quantum numbers, we plan to build on our tested ML approach (E. R. Guest et al. 2024), which will be used in conjunction with our semiclassical estimates to generate pressure-broadening data for billions of transitions listed in ExoMol. Other theoretical methods such as advanced semiclassical calculations for molecular pairs with available interaction potential energy surfaces—see A. Sokolov et al. (2025)—will be also employed to provide rotationally dependent high-accuracy line-broadening data where required.


### Acknowledgments

This work was supported by ERC Advanced Investigator Project 883830 (ExoMolHD) and by the STFC grant ST/Y001508/1; E.R.G. is part supported by the UCL Centre for Doctoral Training in Data Intensive Science, funded by the STFC training grant reference ST/P006736/1. S.Y. thanks the Franche-Comte Region for the financial support of his stay at Université de Franche-Comté.

### Data Availability

The bulk of calculated semiclassical estimates for line-broadening coefficients is accessible from the COLLINE database (<https://colline.u-bourgogne.fr>). Selected one-value entries are also in the appropriate .broad files in the ExoMol database (<https://www.exomol.com>).

### ORCID iDs

Sergei N. Yurchenko  <https://orcid.org/0000-0001-9286-9501>  
Jonathan Tennyson  <https://orcid.org/0000-0002-4994-5238>

### References

- Anderson, P. W. 1949, *PhRv*, **76**, 647
- Anisman, L., Chubb, K. L., Changeat, Q., et al. 2022, *JQSRT*, **283**, 108146
- Barton, E. J., Hill, C., Czurylo, M., et al. 2017, *JQSRT*, **203**, 490
- Buldyreva, J., Yurchenko, S. N., & Tennyson, J. 2022, *RASTI*, **1**, 43
- Chubb, K. L., Rocchetto, M., Yurchenko, S. N., et al. 2021, *A&A*, **646**, A21
- Chubb, K. L., Robert, S., Sousa-Silva, C., et al. 2024, *RASTI*, **3**, 636
- Conway, E. K., Gordon, I. E., Hargreaves, R. J., et al. 2021, *JQSRT*, **270**, 107716
- Daneshvar, L., Foldes, T., Buldyreva, J., & Vander Auwera, J. 2014, *JQSRT*, **149**, 258
- Delahaye, T., Armante, R., Scott, N., et al. 2021, *JMoSp*, **380**, 111510
- Fortney, J. J., Robinson, T. D., Domagal-Goldman, S., et al. 2019, arXiv:1905.07064
- Gharib-Nezhad, E., Iyer, A. R., Line, M. R., et al. 2021, *ApJS*, **254**, 34
- Gharib-Nezhad, E., & Line, M. R. 2019, *ApJ*, **872**, 27
- Gordon, I. E., Rothman, L. S., Hargreaves, R. J., et al. 2022, *JQSRT*, **277**, 107949
- Guest, E. R., Tennyson, J., & Yurchenko, S. N. 2024, *JMoSp*, **401**, 111901
- Hedges, C., & Madhusudhan, N. 2016, *MNRAS*, **458**, 1427
- Hill, C., Yurchenko, S. N., & Tennyson, J. 2013, *Icar*, **226**, 1673
- Loukhovitski, B. I., & Sharipov, A. S. 2021, *JPCA*, **125**, 5117
- Mayor, M., & Queloz, D. 1995, *Natur*, **378**, 355
- Pezzella, M., Yurchenko, S. N., & Tennyson, J. 2021, *PCCP*, **23**, 16390
- Sharp, C. M., & Burrows, A. 2007, *ApJS*, **168**, 140
- Sokolov, A., Yurchenko, S. N., Tennyson, J., Gamache, R. R., & Vispoel, B. 2025, *JQSRT*, **330**, 109225
- Tan, Y., Skinner, F. M., Samuels, S., et al. 2022, *ApJS*, **262**, 40
- Tennyson, J., Hill, C., & Yurchenko, S. N. 2013, in AIP Conf. Proc. 1545, 8th Int. Conf. on Atomic and Molecular Data and Their Applications: ICAMDATA-2012 (Melville, NY: AIP), 186
- Tennyson, J., Hulme, K., Naim, O. K., & Yurchenko, S. N. 2016, *JPhB*, **49**, 044002
- Tennyson, J., & Yurchenko, S. N. 2012, *MNRAS*, **425**, 21
- Tennyson, J., Yurchenko, S. N., Zhang, J., et al. 2024, *JQSRT*, **326**, 109083
- Tinetti, G., Tennyson, J., Griffith, C., & Waldmann, I. 2012, *RSPTA*, **370**, 2749
- Tsao, C., & Curnutte, B. 1962, *JQSRT*, **2**, 41
- Wang, R., Balciunaite, U., Chen, J., et al. 2023, *JQSRT*, **306**, 108617
- Wcislo, P., Thibault, F., Stolarczyk, N., et al. 2021, *JQSRT*, **260**, 107477
- Wilzewski, J. S., Gordon, I., Kochanov, R. V., Hill, C., & Rothman, L. S. 2016, *JQSRT*, **168**, 193
- Wolszczan, A., & Frail, D. 1992, *Natur*, **355**, 145
- Yurchenko, S. N., Tennyson, J., Bailey, J., Hollis, M. D. J., & Tinetti, G. 2014, *PNAS*, **111**, 9379