Effect of pressure broadening on molecular absorption cross sections in exoplanetary atmospheres

Christina Hedges⋆ and Nikku Madhusudhan
Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

ABSTRACT
Spectroscopic observations of exoplanets are leading to unprecedented constraints on their atmospheric compositions. However, molecular abundances derived from spectra are degenerate with the absorption cross-sections which form critical input data in atmospheric models. Therefore, it is important to quantify the uncertainties in molecular cross-sections to reliably estimate the uncertainties in derived molecular abundances. We investigate and quantify the effects of various factors involved in line broadening in exoplanetary atmospheres – the profile evaluation width, pressure versus thermal broadening, broadening agent, spectral resolution and completeness of broadening parameters – on molecular absorption cross-sections. We use H₂O as a case study as it has the most complete absorption line data. For low-resolution spectra (R ≲ 100) for representative temperatures and pressures (T ~ 500–3000 K, P ≲ 1 atm) of H₂-rich exoplanetary atmospheres, we find the median difference in cross-sections (δ) introduced by various aspects of pressure broadening to be ≲ 1 per cent. For medium resolutions (R ≲ 5000), including those attainable with James Webb Space Telescope, we find that δ can be up to 40 per cent. For high resolutions (R ~ 10⁵), δ can be ≳ 100 per cent, reaching ≳ 1000 per cent for low temperatures (T ≲ 500 K) and high pressures (P ≳ 1 atm). The effect is higher still for self-broadening. We generate a homogeneous data base of absorption cross-sections of molecules of relevance to exoplanetary atmospheres for which high-temperature line lists are available, particularly H₂O, CO, CH₄, CO₂, HCN, and NH₃.

Key words: methods: laboratory – planets and satellites: atmospheres – planetary systems.

1 INTRODUCTION
In recent years, it has become possible to observe high-precision atmospheric spectra of a variety of exoplanets detected via transits, direct imaging and Doppler surveys (see e.g. review by Madhusudhan et al. 2014b). For example, high-precision observations with the Hubble Space Telescope (HST) Wide Field Camera 3 (WFC3) in the near-infrared (1.1–1.7 μm) have led to unambiguous detections of H₂O in several hot Jupiter atmospheres (Deming et al. 2013; Kreidberg et al. 2014; McCullough et al. 2014; Madhusudhan et al. 2014a). Despite the modest resolution (R ~ 10–100; depending on spectral binning), the high photometric precisions of HST WFC3 spectra have allowed unprecedented constraints on the H₂O abundances in these atmospheres (e.g. Kreidberg et al. 2014; Madhusudhan et al. 2014a). On the other hand, it has also become possible to detect molecules such as H₂O and CO in atmospheres of hot Jupiters orbiting bright stars using very high resolution (R = 10⁵) infrared Doppler spectroscopy with large ground-based facilities (Snellen et al. 2010; Brogi et al. 2012). Complementary to short-period exoplanets, high-resolution and high-precision spectra have also been reported for young giant planets on large orbital separations discovered via direct imaging from ground leading to detections of several key molecules including H₂O, CO and CH₄ (Konopacky et al. 2013). While these observational advancements are already leading to detailed constraints on the chemical compositions and physical processes in exoplanetary atmospheres, the field will be further revolutionized with upcoming facilities including James Webb Space Telescope (JWST; R ~ 1000–3000) in space and several large ground-based facilities such as the European Extremely Large Telescope (E-ELT) (R ~ 10⁵).

Central to the interpretation of such observations, however, is the accuracy of the fundamental inputs to atmospheric models. The opacity contributing to a spectrum due to a given chemical species is proportional to the product of its absorption cross-section and the molar abundance of that species. Therefore, the chemical abundances derived from atmospheric spectra, using standard retrieval

* E-mail: clh93@ast.cam.ac.uk

© 2016 The Authors
Published by Oxford University Press on behalf of the Royal Astronomical Society
Molecular absorption cross-sections are generated from transition line lists by incorporating the appropriate line broadening and binning to the required resolution. Several factors can cause line broadening in exoplanetary atmospheres, such as thermal Doppler broadening and pressure broadening. However, while detailed high-temperature line lists are becoming available for several molecules, there is still a lack of detailed line broadening parameters that are required for generating accurate cross-sections. Freedman et al. (2008, 2014) report opacity calculations using existing high-temperature line lists and highlight the lack of detailed pressure broadening parameters. In particular, current atmospheric observations are most sensitive for giant exoplanets with H$_2$-rich atmospheres, and hence molecular opacities in models need to incorporate pressure broadening due to H$_2$. However, while latest data bases such as HITRAN (Rothman et al. 2013) provided data for a large compendium of molecules, the data were typically available for terrestrial applications of temperatures $\lesssim$300 K. Since exoplanetary atmospheres that are observable with current instruments span much higher temperatures ($\sim$600–3000 K), it has become necessary to generate high-temperature line lists for many molecules. Therefore, several new high-temperature molecular line lists have been reported in recent years. For example, the recently revised HITTEMP data base (Rothman et al. 2010) provides a compilation of theoretical and experimental high-temperature line lists and line broadening parameters for important molecules such as H$_2$O, CO and CO$_2$. More recently, and more extensively, the ExoMol data base (Tennyson & Yurchenko 2012) has reported high-temperature theoretical line lists for numerous molecules of relevance to exoplanetary atmospheres, including the largest line lists for CH$_4$ and NH$_3$ known to date (e.g. Yurchenko, Barber & Tennyson 2011; Yurchenko & Tennyson 2014).

Molecular absorption cross-sections are generated from transition line lists by incorporating the appropriate line broadening and binning to the required resolution. Several factors can cause line broadening in exoplanetary atmospheres, such as thermal Doppler broadening and pressure broadening. However, while detailed high-temperature line lists are becoming available for several molecules, there is still a lack of detailed line broadening parameters that are required for generating accurate cross-sections. Freedman et al. (2008, 2014) report opacity calculations using existing high-temperature line lists and highlight the lack of detailed pressure broadening parameters. In particular, current atmospheric observations are most sensitive for giant exoplanets with H$_2$-rich atmospheres, and hence molecular opacities in models need to incorporate pressure broadening due to H$_2$. However, while latest data bases such as HITRAN provide broadening data for self- and air-broadening, H$_2$-broadening data are currently unavailable for most molecules. Thus, high-temperature molecular cross-sections for several molecules rely on insufficient H$_2$ pressure broadening data (Freedman et al. 2008, 2014; Sharp & Burrows 2007). On the other hand, purely theoretical line lists such as ExoMol are computed in zero-pressure conditions due to which pressure broadening data are typically unavailable and hence not included while computing cross-sections (e.g. Hill, Yurchenko & Tennyson 2013). Finally, even where broadening parameters are available, deriving cross-sections from line lists involves a series of numerical considerations which can influence the cross-sections.

In the present work, we investigate the dependence of molecular cross-sections on the various factors involved in computing them from molecular line lists including both thermal and pressure broadening. We discuss each aspect of the construction of the cross-sections and where errors can be introduced by comparing different data sources, temperatures, pressures, broadening agents, evaluation widths and resolutions. Meaningful comparisons can be difficult to make as cross-sections span many orders of magnitude and contain many peaks and troughs in features across wavelengths. As discussed in Section 6.1, we find the median percentage difference to be a useful measure and representative of the change due to pressure broadening as a whole. To assess this, we also generate a bank of cross-sections with only thermal broadening included for comparison with those including pressure broadening. Using this method, we create a data base of molecular cross-sections from a range of available line list sources. The cross-sections generated span a wide range of pressures ($10^{-4}$–100 atm) and temperatures (300–3500 K) that are relevant for exoplanetary atmospheres and sub-stellar objects.

We investigate the cross-sections over a wide range in spectral resolution ($R = 10^5$–$10^7$) which reflects current and future instruments for atmospheric characterization of exoplanets as discussed above. There are a wide array of future instruments that we can prepare for. The JWST is scheduled for launch in 2018 and will host several instruments in the infrared range. The Near-Infrared Spectrograph (NIRSpec) and Mid-Infrared Instrument (MIRI) spectrographs will hold significance for the characterization of exoplanet atmospheres. NIRSpec will encompass the 0.6–5 μm range using three overlapping bands. MIRI will be a particularly broad-band instrument stretching from 5–20 μm in wavelength. NIRSpec will have a spectral resolution ranging up to $R = 1400$–3600 μm for its highest resolution grating and MIRI will achieve $R \approx 3000$. This is a drastic increase in our current capabilities with HST which warrants an investigation into our current modelling inputs. In the distant future, we also anticipate the E-ELT to have $R = 100$ 000 and be highly capable of atmospheric characterization for exoplanets. Based on these instruments, we present our tests of cross-sections at a range of resolutions of $R = 100$, 1000, 3000, 10 000 and 100 000. This gives a representative picture of what we can achieve now and in the future.

### 2 LINE LIST SOURCES

There are several line list repositories available where lists of transitions for molecules can be obtained. Those most relevant to atmospheric characterization and that are publicly available are given in Appendix A1 including the number of transitions each list contains. These data bases are given in a variety of formats but each contain the key parameters (e.g. Einstein coefficients, degeneracies, energy levels, etc.) and a method for obtaining the line positions (e.g. in wavenumber, $\nu$) and the line intensities $S(T)$ for a given temperature ($T$). These line data sources cover a wide spectral range, typically spanning the visible to mid-IR (e.g. $\sim$0.5 to $\sim$30 μm), although this can vary between the different line lists. Line transitions are uniformly spread in frequency space leading to line lists being given in wavenumber rather than wavelength. Sources also range in their completeness with some containing fewer transitions than others. Lack of completeness leads to less reliable cross-section data for two reasons; gaps in the wavenumber coverage cause some features to be missed from the cross-section and lines of lower intensity which can contribute significantly to the cross-section are not represented. As such accurate cross-sections require the most complete lists of molecular transitions.

One of the largest and most well established of all the repositories for molecular line lists is the HITRAN data base, which has been updated every few years (Rothman et al. 1998, 2013). The
HITRAN data base has been mainly used for terrestrial applications and predominantly includes molecules of importance for the Earth’s atmosphere, with temperatures below ~300 K. Because of this lower temperature approach, these data are less applicable for the most observable exoplanetary atmospheres. For current atmospheric observations of exoplanets, we require spectroscopic data covering a higher temperature range, e.g. for applications to highly irradiated planets and young directly imaged planets which have a wide range of temperatures going up to 3000 K. This need has been met by the newer HITRAN data base, also known as HITRAN2010 to distinguish from an earlier version, which contains fewer molecules but many more transitions for each (Rothman et al. 2010). HITRAN currently covers OH, NO, CO, CO₂ and H₂O which are particularly useful for hot Jupiter atmospheres, accurate up to temperatures of 4000 K. More recently, the ExoMol data base has begun addressing the deficit of data available for molecules of astrophysical importance at high temperatures (Tennyson & Yurchenko 2012).

3 LINE BROADENING

The diversity of physical conditions in exoplanetary atmospheres can lead to different types and degree of line broadening. Exoplanetary atmospheres span a wide range of temperatures (~400–3000 K) and dynamical parameters (e.g. wind speeds, and orbital and spin rotation rates) ranging from tidally locked close-in planets to young giant planets on wide orbital separations. The two prominent sources of line broadening in planetary atmospheres are thermal (Doppler) broadening and pressure (collisional) broadening (Chamberlain 1978; Mihalas, Auer & Mihalas 1978; Seager & Deming 2010). Thermal Doppler broadening is caused by the line-of-sight thermal velocity distribution of molecules at a given temperature in the planetary atmosphere. Pressure broadening is induced by collisions between chemical species with the collision frequency being a strong function of pressure.

Other sources of broadening can be prevalent depending on the planetary properties and observing geometry. In principle, natural broadening due to the intrinsic uncertainty in energy levels is always present but is expected to be negligible compared to other broadening mechanisms discussed above. Further broadening and shifting of spectral lines can be caused by the planetary rotation and strong winds in the planetary atmosphere, especially for close-in hot Jupiters observed in transmission spectra (Spiegel, Haiman & Gaudi 2007; Miller-Ricci Kempton & Rauscher 2012; Showman et al. 2013). Finally, rotational broadening due to the spin of the planet can also be significant, especially for exoplanets that are not tidally locked such as those on wide orbital separations (Snellen et al. 2014). These sources of broadening can be important on a case-by-case basis.

3.1 Broadening profiles

Opacities used in models of exoplanetary atmospheres typically include thermal and pressure broadening of the spectral lines wherever the broadening parameters are available. While thermal broadening is straightforward to include (e.g. Hill et al. 2013), the parameters for pressure broadening are sparse as relevant approximations need to be made for each molecule considered (Freedman, Marley & Lodders 2008; Freedman et al. 2014), as discussed in Section 3.3. In the present work, we consider both thermal and pressure broadening and investigate the effect of the assumed pressure broadening parameters on the resulting absorption cross-sections.

Under the assumption of a Maxwell–Boltzmann thermal velocity distribution, the Doppler broadening takes the form of a Gaussian profile. On the other hand, pressure broadening is represented by a Lorentzian profile. The Doppler and Lorentzian broadening profiles are given, in wavenumbers in cm⁻¹, as

\[ f_D(v - v_0) = \frac{1}{\gamma_D \sqrt{\pi}} \exp\left(-\frac{(v - v_0)^2}{\gamma_D^2}\right) \]  

\[ f_L(v - v_0) = \frac{1}{\pi} \frac{\gamma_L}{(v - v_0)^2 + \gamma_L^2} \]

where \(v_0\) is the centroid in wavenumbers, \(\gamma_D\) is the Doppler width and \(\gamma_L\) is the Lorentzian pressure broadening half-width at half-maximum (HWHM) both in units of cm⁻¹. These are given by (see e.g. Rothman et al. 1998; Hill et al. 2013)

\[ \gamma_D = \sqrt{\frac{2k_BT}{m} v_0} \]  

\[ \gamma_L = \left(\frac{T_{ref}}{T}\right)^p \sum_b \gamma_{L,b} P_b, \]

where \(P\) is pressure in atm, \(T\) is the temperature in kelvin, \(T_{ref}\) is the reference temperature (usually 296 K), \(p_b\) is the partial pressure of the broadener, \(n\) is a temperature scaling factor and \(\gamma_{L,b}\) indicates the Lorentzian HWHM due to a specific broadening molecule in units of cm⁻¹ atm⁻¹. Here \(k_B\) is the Boltzmann constant, \(m\) is the mass of the molecule in grams and \(c\) is the speed of light in cm s⁻¹. Here, \(\sum\) signifies the sum over all the broadening parameters for each broadening medium.

Pressure broadening is typically harder to evaluate than thermal broadening for multiple reasons. First, as alluded to above, the line-by-line pressure broadening parameters, \(n(v)\) and \(\gamma(v)\), are typically unavailable for most molecules under the conditions encountered in exoplanetary atmospheres, e.g. high temperatures up to ~3000 K and varied atmospheric compositions such as H₂-rich gas giant atmospheres (Freedman et al. 2008, 2014). We discuss this further in Section 3.3. Secondly, the Lorentzian profile contributes a higher percentage to its extensive wings which can result in a significant amount of the intensity being moved into the wings of the line profile. In the case of extreme broadening, this can significantly increase the impact of high-intensity transitions far from the line centre and influence cross-sections from neighbouring low-intensity transitions. Therefore, the profile needs to be treated particularly carefully by sampling the Voigt well to approximate the profile shape with low uncertainties. This ensures appropriate normalization so that no intensity is lost from binning the profile.

Fig. 1 shows the spread of both the Lorentzian HWHM \(\gamma_{L,b}\) and the temperature scaling parameter \(n\) with the intensity for CO, one of the simplest molecules in the data base. We can see discrete levels in each parameter due to the discrete nature of the \(J\) quantum number that the pressure broadening values are generated from in complex Robert–Bonamy (CRB) calculations Robert & Bonamy (1979); Antony et al. (2006). The spread is quite narrow in this parameter space leading us to believe that a mean value of each parameter could reasonably approximate the results of a detailed line-by-line treatment of pressure broadening.

It is useful to understand where each of the two broadening mechanisms contributes most significantly in the pressure–temperature (\(P-T\)) space relevant for exoplanetary atmospheres. However, as the two profiles are disparate with a Gaussian profile containing more information in the core compared to the extended Lorentzian
wings, it is difficult to compare the two with a common metric. Nevertheless, Fig. 2 shows a comparison between the Doppler HWHM ($\gamma_G \ln 2$) and the Lorentzian HWHM over the $P-T$ space of interest to give an approximation of where each profile contributes most significantly. As expected, at low pressures thermal (Gaussian) broadening provides a significant contribution to the final profile core, whereas at high pressures pressure (Lorentzian) broadening is stronger. Closer to the boundary between these two regimes, both broadening mechanisms are likely to contribute significantly to core of the profile. Furthermore, due to the extended wings of the Lorentzian, it is generally advisable to consider both broadening contributions even when the Lorentzian HWHM is narrow in comparison to the Gaussian. This is done using the Voigt profile discussed below.

As discussed in Ngo et al. (2012), it has been shown through comparisons with experimental data on pressure broadening that there is some deviation in reality from the Voigt profile. This is due to the change in velocity of the broadener particle by the collision with the broadening agent which affects the profile shape, the width and the shift of the line. Here we have used only the standard Voigt profile and have not investigated further in terms of profile shape. It would be possible to change to a more sophisticated profile and regenerate molecular cross-sections if it were found to be an important factor. These deviations due to the velocity of the particles undergoing collisions are of the order of a few per cent, and we do not expect a more physically accurate Voigt profile shape to impact our results.

### 3.2 Evaluating the Voigt profile

The joint contributions due to thermal and pressure broadening are modelled using a Voigt profile which is a convolution of the Gaussian and Lorentzian profiles, given as

$$f\nu(v - v_0) = \int_{-\infty}^{\infty} f\nu(v' - v_0) f\nu(v - v') dv'.$$

The characteristic width of the Voigt function is investigated by Olivero & Longboothum (1977). Using coefficients from this work, we define the Voigt width $\gamma_V$ with the approximation

$$\gamma_V \approx 0.5346\gamma_G + \sqrt{0.1686\gamma_G^2 + \gamma_L^2}.$$

This Voigt width is used in later sections to approximate the width of the combination of the two profile types.

The issue of how best to evaluate the Voigt profile is a well-known problem. The profile must be calculated accurately and quickly for a wide range of Lorentzian and Gaussian profiles, corresponding to the wide range of temperature and pressure values, for potentially millions of lines of a given molecule. The two parameters used for generating the profile are

$$\alpha = \frac{v - v_0}{\gamma_G}, \quad \alpha' = \frac{\gamma_L}{\gamma_G}.$$

where $\alpha$ is the distance from the profile centroid and $\alpha'$ is the ratio of the Lorentzian to Gaussian widths (Zaghloul 2007). To be able to calculate the Voigt function accurately, $\alpha$ must be evaluated over many orders of magnitude to encompass the relevant temperature and pressure region.

The Voigt function is given as

$$H(a, u) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{ae^{-t^2}}{(u - t)^2 + a^2} dt,$$

From formula 7.4.13 in Abramowitz & Stegun (1964)

$$\int_{-\infty}^{\infty} \frac{ye^{-t^2}}{(x - t)^2 + y^2} dt = \pi\Re\mathcal{w}(x + iy),$$

where

$$w(z) = e^{-z^2} \text{erfc}(-iz),$$

where $\text{erfc}$ is the complimentary error function

$$\text{erfc}(z) = 1 - \text{erf}(z).$$
Comparison of purely thermally broadened H$_2$O cross-sections (blue) with cross-sections including both thermal and pressure broadening using a Voigt profile (red) at the native line spacing of 0.01 cm$^{-1}$. A combination of the two broadening types brings extensive wings from the Lorentzian component which brings up the level of the continuum.

where erf is the error function

$$
erf(z) = \frac{2}{\sqrt{\pi}} \int_{0}^{z} e^{-t^2} dt.
$$

From this we can see that

$$
f_i(v, \gamma_L, \gamma_a) = Re(w(\alpha + i\beta)).
$$

The function $w(z)$ is known as the Faddeeva function and here is calculated using the Faddeeva package (Johnson, Cervellino & Wutteke 2012).

Several numerical methods have been proposed to compute the Voigt profile accurately and efficiently over different regions of parameter space (Schreier 2011). We implement the Voigt profile using a method based on the complex error function or Faddeeva function (Johnson et al. 2012). This is a fast and accurate method with the relevant libraries publicly available. The Faddeeva package includes Algorithm 916 from Zaghloul & Ali (2011) which is known to provide accurate results. We find that this package gives converged profiles over our required parameter space with fast computational speeds of less than 2 ms per profile. We use a sampling rate of 6 points per Voigt width, which is much finer than our required final resolution. We evaluate the Voigt profile to 500 Voigt widths around the centroid to accurately capture the information in the Lorentzian wings, as the derived cross-sections are critically dependent on this evaluation width (discussed in more detail in Section 5.2).

3.3 Availability of broadening parameters

Computing line broadening, as discussed above, requires broadening parameters for each line for both the thermal and pressure broadening components. For thermal broadening of a spectral line of a given molecule at a given temperature, the Doppler width ($\gamma_L$) is easily calculated from equation (3). Several recent works have reported such thermal broadened cross-sections for several molecules, e.g. Barber et al. (2006) and Hill et al. (2013). Fig. 3 shows a comparison between cross-sections for H$_2$O generated with only thermal broadening and those generated with a full Voigt profile including both thermal and pressure broadening. It is clear that the Voigt profile has a significant effect on the low-intensity lines, and increases the overall continuum of the molecular cross-sections, especially where pressure is high.

Despite their critical importance, pressure broadening parameters are not yet readily available for all molecules of relevance to exoplanetary atmospheres (Freedman et al. 2008, 2014). As shown in equation (4), the parameters required for computing pressure broadening are the Lorentzian HWHM ($\gamma_L$), the pressure broadening coefficient ($\gamma_a$), and the pressure broadening parameter ($n$) for each line in the line list. These parameters are hard to determine. Theoretical calculations of pressure broadening parameters are particularly time-consuming and not covered for a wide range of molecules or broadening agents, particularly at high temperatures. Such methods have been explored by Gamache et al. (1997) and Gamache, Laraia & Lamouroux (2011) and used to generate the HITRAN database, e.g. CRB calculations are used as discussed in Gamache, Lynch & Neshyba (1998) and Gamache et al. (2012) where values are also verified experimentally. Molecular line lists in the HITRAN database do contain pressure broadening parameters for self-broadening and air-broadening, but are typically relevant only to low-temperature atmospheres ($< 300$ K). For high-temperature exoplanetary atmospheres, particularly of H$_2$-rich atmospheres that are most observable, pressure broadening data are still scarce. Typically, state-of-the-art ab initio line lists such as ExoMol are computed under zero-pressure conditions due to which the pressure broadening line parameters are not available. On the other hand, experimental data are also scarce for the conditions relevant for exoplanetary atmospheres.

In recent years, significant efforts have been dedicated towards generating pressure broadening parameters for exoplanetary applications, particularly with a focus on broadening molecules such as H$_2$ and other molecules of interest outside of Earth applications. Useful parameters for important molecules can be found in works such as Li et al. (2015) and Faure et al. (2013) where pressure broadening coefficients are made available. In this work, we use only air- and self-broadening provided by HITEMP and HITRAN and the PS 1997 list for water from Partridge & Schwenke (1997), which contains H$_2$-broadening (Freedman, personal communication). We focus on the H$_2$O case as it is the most well-studied molecule currently with a large variety of line lists and three different broadening molecules available to test. Further, H$_2$O is one of the best measured molecules in exoplanet atmospheres to date. We anticipate expanding this work to cover other molecules such as CO with further broadening agents as it is another molecule of interest, particularly in very high resolution Doppler spectroscopy of hot exoplanetary atmospheres (e.g. Snellen et al. 2013).

However, despite these sources, there is a lack of both experimental and theoretical high-temperature data on H$_2$-broadening parameters for many molecules of interest for exoplanetary atmospheres, e.g. CH$_4$, CO$_2$, NH$_3$, CH$_3$H$_2$, HCN, TiO, CO, etc. (Madhusudhan 2012). Currently, most molecules that are available with pressure broadening information have only self-broadening and air-broadening parameters. This is beginning to be addressed and more H$_2$-broadening parameters are becoming available (Wilzewski et al. 2016).

There has been a great deal of work in making these parameters accurate and accessible which is made difficult by the sheer volume of lines for which broadening parameters are needed. Data bases such as HITEMP, HITRAN and GEISA (Gestion et Etude des Informations) Spectroscopiques Atmosphériques give pressure broadening values for each transition for each molecule. Where values are not available, we are forced to turn to what is available in the literature from experiments and from other line lists. Where

Figure 3. Comparison of purely thermally broadened H$_2$O cross-sections (blue) with cross-sections including both thermal and pressure broadening using a Voigt profile (red) at the native line spacing of 0.01 cm$^{-1}$. A combination of the two broadening types brings extensive wings from the Lorentzian component which brings up the level of the continuum.
Table 1. Pressure–temperature grid for cross-sections. Values between these points could be interpolated by the user to allow a finer grid. These points have been chosen as they represent observable temperatures and pressures of currently known exoplanets.

| T (K)  | 300   | 400   | 500   | 600   | 700   | 800   | 900   |
|--------|-------|-------|-------|-------|-------|-------|-------|
| P (atm)| 0.0001| 0.001 | 0.01  | 0.1   | 1     | 10    | 100   |

Detailed line-by-line pressure broadening values are not available for a molecule, we use the mean of pressure broadening values available from other sources. While this approach is not ideal, it is currently the only option in some cases. For example in this work, the YT10to10 line list for CH₄ is highly complete and a useful asset to modelling of exoplanetary atmospheres; however, pressure broadening values have not been calculated. An alternative HITRAN list has pressure broadening components but only a small fraction of the molecular transitions are covered and the list is only appropriate for room temperature applications. We will discuss the consequences of this method of taking the average broadening in Section 6.3 and the impact using a mean broadening value has on the final cross-section. Table 3 also shows which molecules have detailed broadening available and which only have mean broadening. For each case where detailed broadening parameters are available, a mean case has also been investigated in order to make a comparison.

4 GENERATING CROSS-SECTIONS

The combination of molecular line parameters and broadening profiles discussed above allows us to compute the molecular absorption cross-sections which form inputs to exoplanetary atmospheric models. We compute molecular cross-sections over a wide range of pressures (P) and temperatures (T) relevant for exoplanetary atmospheres. Our grid in P–T space is shown in Table 1. For each point in P–T, each cross-section will usually be calculated multiple times using different broadening parameter or source. If a cross-section was required between these points, interpolation could be used on the cross-sections between the nearest pressure and temperature values as outlined in Hill et al. (2013). In this section, we describe the procedure we use to compute cross-sections.

4.1 Line intensities and partition functions

Generation of cross-sections requires the intensity of each transition to be accurately calculated. Most line list data bases give Einstein coefficients for each transition with degeneracies and energies for each state. These can be converted into line intensities as (Rothman et al. 2013)

\[
S_i,j(T) = \frac{A_{i,j}}{8\pi c v_{i,j}^2 Q(T_{ref})} g_i e^{-\hbar c E_i/k_B T_{ref}} \frac{Q(T)}{Q(T_{ref})} g_j e^{-\hbar c E_j/k_B T_{ref}} (1 - e^{-\hbar c v_{i,j}/k_B T}),
\]

(15)

where \( A_{i,j} \) is the Einstein coefficient for spontaneous emission for the transition between states \( i \) and \( j \), \( g_i \) is the upper state degeneracy, \( E_i \) is the lower state energy in cm\(^{-1}\), \( v_{i,j} \) is the transition frequency between \( i \) and \( j \), also in cm\(^{-1}\) and finally \( h \) is the Planck constant. Here, \( Q(T_{ref}) \) is the partition function at the required reference temperature. When an intensity is given at a reference temperature, usually 296 K, it can be converted to an intensity at any temperature using

\[
S_i,j(T) = \frac{Q(T_{ref})}{Q(T)} S_i,j(T_{ref}) \exp(-\hbar c E_i/k_B T) \left[1 - \exp(-\hbar c v_{i,j}/k_B T)\right],
\]

(16)

This gives the intensity of a transition in units of cm\(^{-1}\)/molecule cm\(^{-2}\).

As evident from equations (15) and (16), the partition function scales the line intensities. The partition function gives a measure of how many of the molecules of a gas are in the ground state compared with all other states. This ratio of state populations increases with temperature as it becomes more likely to find particles in higher energy states. This is intrinsically linked to the energy of each transition which makes the partition function unique for each molecule. The partition function is given by

\[
Q(T) = \sum_j g_j e^{-E_j/k_B T},
\]

(17)

where \( g_j \) is the lower state degeneracy and \( E_j \) is the lower state energy as described in equation (15). However, to calculate a partition function this way, the spectral information of the molecule must be complete. Missing transitions result in the partition function which will not be representative, especially at high temperatures. Information on how partition functions are calculated for ExoMol can be found in Tennyson & Yurchenko (2012). Errors in the partition function may cause discrepancies that are small when considering individual cross-sections at particular temperatures but may impact significantly when considering full atmospheric models. Models of atmospheres involve temperature–pressure profiles with respect to altitude and so require cross-sections at many different temperatures to be included in computing the emergent spectrum. This will cause any errors in the cross-section from the partition function to become compounded by many layers of an inhomogeneous atmosphere.

In the present work, partition functions have either been adopted from existing data bases or computed using the TIPS code (Fischer et al. 2003). The ExoMol data base provides partition functions for use with each molecular line list. The HITRAN data base uses the TIPS code as discussed in Fischer et al. (2003) to create partition functions between 70 and 3000 K. There can be some discrepancies between partition functions from different sources, particularly at high temperatures. Some are only available in a small range of temperatures. Where this is the case we use extrapolation to find the values at higher temperatures. In our case, this only affected our highest temperature point of \( T = 3500 \) K.

Using a line list at a temperature greater than that recommended by the source is possible by using the correct partition function. However, while at low temperatures many transitions have low enough intensity to be discounted without effecting the final result, as temperature increases these low-intensity lines begin to contribute more noticeably. Because of this, any low transitions that are missed (either by an intensity cut-off that is too low or from being omitted in the source) can cause errors in the cross-section if molecular line lists are used above their recommended temperatures. These temperatures are given in the last column of Table 3. Here we see with the exception of NH₃ and CH₄, that the ExoMol and HITRAN line lists provide coverage up to and beyond the temperatures used in this work. The HITRAN sources, used primarily for earth applications, are only valid at room temperature. Due to their small number of transitions (shown in Table A1), HITRAN line lists are unreliable at high temperatures even when used with the
correct partition function. Nevertheless, these have been included in this work for two reasons; first to investigate the dependence of completeness of line list sources on the accuracy of cross-sections and, secondly, to include useful molecules that are not covered by any other source.

4.2 Cross-sections from line intensities

Cross-sections are derived from line intensities by first broadening with the appropriate profile. This is followed by binning the resultant cross-sections to the desired spectral resolution. This is a general approach followed by several recent studies (e.g. Hill et al. 2013), albeit with minor differences in implementation. Here we discuss our implementation. The differences from other works are discussed in Section 5. For a given temperature (T) and pressure (P), computing the cross-sections from line intensities involves three steps as follows.

First, the Voigt profile (f_v) is computed at a high resolution in order to accurately evaluate each individual line profile as described in Section 3.2. The spacing of this fine grid, here referred to as the ‘sub-grid’, is given as
\[ \Delta \nu = \gamma_v(v = 500, T, P, m) / 6, \]
where \( \gamma_v(v = 500, T, P, m) \) is the HWHM of a Voigt profile at \( v = 500 \) cm\(^{-1}\). We find this prescription, which samples each Voigt width with 6 points, to provide the requisite resolution and accuracy with optimal computational speed, as discussed in more detail in Section 5. This very fine sampling gives an accurate representation the Voigt function across the whole wavenumber space. This sub-grid spacing is a function of \( T, P \) and molecular mass making it specific to the molecule concerned.

Secondly, for a given spectral line, the cross-section is computed at each point on the sub-grid described above. The cross-section \( \sigma(v) \) of a transition between states \( i \) and \( j \) at a certain pressure (P) and temperature (T) is given by
\[ \sigma_{i,j,P,T}(v) = \frac{f_v(v)}{\int_{\nu_{ij}}^{\nu_{ij} + \Delta \nu_{ij}} f_v(v) \, dv} \approx \frac{S_{i,j,P,T} f_v(v)}{\int_{\nu_{ij}}^{\nu_{ij} + \Delta \nu_{ij}} f_v(v) \, dv} \]
\[ (19) \]
in units of cm\(^{-1}\)/molecule, where \( S_{i,j,P,T} \) is the line intensity and \( f_v(v) \) is the Voigt function with broadening parameters corresponding to the line at the given \( P \) and \( T \). \( \nu_{ij} \) is the wavenumber of the line centre and \( \Delta \nu_{ij} \) is the extent of the profile to which the line wings are evaluated. For a given line, we use a \( \Delta \nu_{ij} \) value of 500 Voigt widths (250 around the line centroid), which we find to be optimal, as discussed in Section 5. The integral is evaluated up to this cut-off and normalizes the profile. Evaluating up to this cut-off effectively folds in the intensity from the missing wings that are not evaluated back into the profile, ensuring no intensity is missed.

Finally, the high-resolution cross-section profile computed for each line as described above is binned to a final cross-section grid with a coarser spacing for saving on storage space. The final cross-section grid spacing is still high resolution at 10\(^{-2}\) cm\(^{-1}\) which corresponds to a resolution of \( R = 100,000 \) or greater depending on wavelength. For instance, at a wavelength of 1\( \mu \)m, this spacing corresponds to a spectral resolution \( (R = \lambda / \Delta \lambda = \nu / \Delta \nu) \) of 10\(^6\). When a lower resolution cross-section grid is desired, this high-resolution grid can be binned down further in frequency space or wavelength space as required. For example, for a given resolution \( R \), the grid points in wavelength space can be determined following \( R = \lambda / \Delta \lambda \) which will give a non-linear grid in \( \lambda \). The mid-points between adjacent bins are selected and all values within these bounds are averaged giving the binned down contribution at each wavelength on the grid.

In this work as discussed in Rothman et al. (2013) and other works, we apply a cut in intensity to only evaluate the significant lines. This provides a reduction in computation time with minimal effect on accuracy as very low intensity lines contribute little to the final cross-section even at high resolution. We apply a cut-off at 10\(^{-30}\) cm\(^{-1}\)/(molecule cm\(^{-2}\)) in intensity across all line lists apart from the BYTe and YT10to10 line lists for NH\(_3\) and CH\(_4\), respectively. Due to their large size, the cut-off was increased to 10\(^{-26}\) cm\(^{-1}\)/(molecule cm\(^{-2}\)). We find that as both these molecules have many complex transitions, our results are unaffected by omitting these low lines across all resolutions covered. Unlike the HITRAN data base, we do not scale our cross-sections by any relative abundances of isotopes based on their terrestrial measurements. This gives each cross-section with the molecule at 100 per cent abundance.

The complete list of molecules used in this work including all sources for data is given in Table 3. Many different line list sources were chosen, particularly for water, for comparison to investigate how completeness affects the resulting cross-section.

5 OPTIMAL RESOLUTION AND CUT-OFF OF BROADENING PROFILE

It is important to consider an appropriately high resolution and extent of the line profile in order to accurately sample the contribution from the profile. Both these properties also influence the computational cost. Therefore, it is desirable to adopt optimal values for each of the properties which facilitate both a high enough accuracy and a reasonable computation time. Several recent studies have adopted different prescriptions for these parameters in the particular context of molecular cross-sections for exoplanetary applications (e.g. Hill et al. 2013; Grimm & Heng 2015). In this section, we systematically investigate the effect of both the profile grid resolution and the extent of the profile wings on the cross-sections in an attempt to determine optimal values for these parameters.

5.1 Effect of profile grid resolution

In the present work, we use a grid resolution that is adaptive with the equivalent width of the Voigt profile for a given line profile as given in equation (6). This approach allows for optimal computational time while ensuring high accuracy of the cross-sections. In this formulation, the grid in frequency space on which we evaluate the Voigt profile, here referred to as the ‘sub-grid’, is defined by equation (18). In this work, we use a minimum sampling of 6 points per Voigt width which was found to be sufficiently accurate based on the investigations carried out below. Evaluating this spacing at \( v = 500 \) cm\(^{-1}\) gives a conservative estimate of the width of a Voigt profile where the Gaussian component is narrowest (see equation 3). This also corresponds to a wavelength of 20\( \mu \)m which is the longest wavelength of interest in the present work and the upper wavelength limit of the infrared observations of exoplanetary atmospheres for instruments such as MIRI on JWST. The necessary grid spacing can become very wide at high pressures where the equivalent width of the Voigt profile becomes large. Therefore, we place an upper limit of 0.01 cm\(^{-1}\) on the grid spacing. When evaluating Voigt profiles on this grid, the range up to which the profile is calculated
Table 2. Table from Hill et al. (2013) giving the staggered spacing of the grid for used for line mapping. Here we present an alternative adaptive grid spacing which is given in equation (18).

| Wavenumber range (cm$^{-1}$) | Grid spacing $\Delta v$ (cm$^{-1}$) |
|------------------------------|-----------------------------------|
| 10–100                      | $10^{-5}$                         |
| 100–1000                    | $10^{-4}$                         |
| 1000–10 000                 | $10^{-3}$                         |
| 10 000–30 000               | $10^{-2}$                         |

Figure 4. Comparison of grid spacing in our adaptive grid with that in the staggered grid from table 2 of Hill et al. (2013) for different pressures. Top: the coloured lines show the Voigt profile widths at different pressures and a representative temperature of 1000 K. The staggered grid spacing is shown as a dashed line for reference, demonstrating that the grid spacing is well below the Voigt profile width for pressures down to $10^{-4}$ atm. Bottom: the coloured lines show the adaptive grid spacing we use. For high pressures ($\geq 1$ atm), the minimum spacing is fixed at $10^{-2}$ cm$^{-1}$ as profiles become very broad. While our grid spacing is coarser than the staggered grid spacing for high pressures, the resultant effect on the cross-sections is small, as shown in Fig. 5. Note that a single value of broadening width and temperature case. Here we take one profile per $T$ and $P$ point.

Figure 5. Comparison of adaptive grid with staggered grid from Table 2. Voigt profiles are mapped to a fine grid from equation (18). The integrated area under the profile for the adaptive case is compared with that for the staggered case to produce this result. Only pressures of $\sim 100$ atm are affected significantly by the resolution of the adaptive grid beyond a few per cent. Below 1 atm, the difference is $\leq 0.2$ per cent or less.

The spacing used in Hill et al. (2013) concerning only thermal broadening where their profile is Gaussian and much narrower requiring finer grid spacing to normalize. However, using our coarser, adaptive grid does not cause significant errors in the final cross-sections. Fig. 4 shows a comparison between the spacing of our adaptive grid and the staggered grid for a representative temperature of 1000 K. It can be seen that both grids are much finer than the Voigt widths for each pressure case (upper panel). While a fine grid spacing gives highly accurate profiles, using the staggered grid can lead unnecessarily to high resolution, especially in the limit of high pressures where the profiles become inherently very broad. This can be computationally expensive particularly for high pressures. The grid we propose in equation (18), referred to here as an adaptive grid as its spacing changes with pressure, uses fewer points at higher pressures to overcome this problem while preserving the accuracy.

Ensuring that the profile is evaluated accurately is imperative to achieve an accurate normalization and neither under- nor overrepresent the line intensity contribution at each grid point. For normalizing the profiles, we compute the area under the curve using a simple trapezium rule. When a profile is only sparsely evaluated, this approach will tend to a greater area estimation, producing a lower intensity contribution from each line profile after normalization. This results in a small percentage of ‘missing’ intensity. To test the validity of our approach, we analyse this amount of missing intensity in a single profile when calculated using the adaptive grid compared with the finer, staggered grid from Hill et al. (2013). This is done for each $T$ and $P$ point.

In order to test our adaptive grid and how well it approximates the Voigt profile, it is compared with the staggered grid using a wide cut-off of $\Delta v_{c} = 100$ cm$^{-1}$ wavenumbers for each pressure and temperature case. Here we take one profile per $T$ and $P$ point only and map to each grid. The profiles are generated in the same way as in Section 3.2. Any difference in the profiles will affect the integrated area, which is used for normalization. We consider the staggered grid to be high enough resolution in all cases that it will produce an accurate area estimation. The comparisons have been conducted over a range of wavelengths but here we select a representative wavelength of $\lambda = 2$ $\mu$m for illustration. Fig. 5 shows the results of this comparison. The difference in the final output of these grids is very small with $\leq 0.2$ per cent of intensity missing at $P < 1$ atm at all temperatures. At $P = 10$ atm, a maximum of 2 per cent of the original intensity is missed at low temperatures.
Several cut-off values have been investigated to establish the optimal balance between accuracy and computational time. As discussed above, we use multiples of Voigt widths to establish a cut-off that adapts to the specific profile. When using a short cut-off, lower intensity lines are underestimated by many orders of magnitude due to the lack of additional intensity from the wings of high-intensity neighbours. This leads to the continuum being poorly approximated by short cut-offs. When the cut-off is increased, we maintain the same normalized area within the profile. Because of this, the profile height becomes slightly shorter as the cut-off becomes wider. Due to this effect, we find that at low cut-off values the cores of very strong lines are slightly overestimated (by approximately 0.2 per cent) at the native spacing of the grid. The effect of this slight overestimation will reduce drastically with resolution. Fig. 6 shows these overestimates and underestimates at the highest resolution.

The underestimation in the profile wings is 10 per cent for H$_2$O at the native spacing. However, this underestimation is confined to the lowest intensity transitions with high-intensity neighbours which are by their nature confined within high-intensity features. As such, an underestimation in these points is of less consequence to most applications.

Due to the profiles being very broad at high pressures where the Voigt profile width is beyond 0.01 cm$^{-1}$, the spacing becomes fixed to 0.01 cm$^{-1}$ wavenumbers. This also fixes our evaluation width to 60 cm$^{-1}$ wavenumbers around the centroid. For high pressures ($P > 10$ atm), this is an underestimate; however, such pressures are less useful for atmospheric modelling in exoplanets.

### 5.3 Comparison of evaluation widths

The wings of the profile will have an effect on any neighbouring lines. If wings are not evaluated out to a large enough separation, the continuum for neighbouring lines will be underestimated. This is particularly important in the case of line lists and cross-sections as intensities span many orders of magnitude and so line wings from high-intensity transitions can be comparable to the peaks of low-intensity neighbours. However, the cut-off also affects the evaluation of the profile. A cut-off that is too close to the centroid will provide poor normalization and will miss some of the intensity contribution.

The cut-off value $\Delta v_c$ is fixed at 500 Voigt widths around the centroid. This is designed so that the evaluation width adapts to both the Gaussian and Lorentzian profiles which is particularly important at extremes of pressure or temperature. This gives 250 Voigt widths around the centroid in each direction. $\Delta v_c$ is increased to 1000 Voigt widths around the centroid at pressures of 1 atm and above. To establish the difference between this method and others in the field, we undertake the following comparisons. First, we take a single Voigt profile on our fine, adaptive grid given by equation (18) and a small Voigt width $\Delta v_c$ of 500 Lorentzian widths as taken from Grimm & Heng (2015). The Voigt profile is calculated in the same way as given in Section 3.2. This is then compared with a single Voigt profile on the same grid with a cut-off at 500 Voigt widths, our value of $\Delta v_c$. The profiles are left unnormalized. The missing area from the Lorentzian cut-off when compared with the method we present is then evaluated as a percentage of the total profile area.

The percentage of missing intensity from this comparison is shown in Fig. 7. It can be seen from the Lorentzian and Voigt width equations that the Voigt width should always cover the same or more wavenumbers and so the Lorentzian cut-off has a smaller integrated area than our approach.

The Lorentzian width evaluation fails at low pressures as the thermal Gaussian component becomes much stronger at lower

---

**Figure 6.** Comparison of cross-sections obtained at the highest resolution using profile evaluation width ($\Delta v_c)$ of 500 Voigt widths versus 10 000 Voigt widths for H$_2$O at 500 K and 0.1 atm. Overestimated and underestimated points are shown in red and blue, respectively. We see that there are some points overestimated by more than 0.1 per cent; however, no points are underestimated by more than 1 per cent. The underestimations average 10 per cent at this high resolution.

The largest differences found are $\sim$10 per cent for pressures of 100 atm, but given that such high pressures are not directly observable for exoplanetary atmospheres, the corresponding differences are less of a concern. We find that in such cases the cut-off chosen $\Delta v_c$ is too narrow for the extreme case of $P = 100$ atm. For the $10^{-4}$–10 atm pressure range, we find the adaptive grid to be very accurate, particularly after binning to the final output grid with a spacing of 0.01 cm$^{-1}$ which corresponds to a spectral resolution of $R = 10^5$–$10^6$ at $\lambda$ <10 $\mu$m.

### 5.2 Effect of profile evaluation width

In order to accurately account for the contribution of the line wings, the broadening profile must be evaluated over a wide enough range centred on the line centre, referred to here as a cut-off value ($\Delta v_c$) as discussed in Section 4.2. The cut-off value determines the extent of evaluation of the profile as well as its normalization as described in equation (19). The choice of $\Delta v_c$ has already been noted in the field as an important factor in computing cross-sections (Sharp & Burrows 2007; Grimm & Heng 2015). A common approach is to apply a cut-off in wavenumber, with values ranging between 10 and 100 cm$^{-1}$, especially for high pressures as discussed by Sharp & Burrows (2007). Another approach is to take a number of Lorentzian widths from the centroid such as in Grimm & Heng (2015); however, this does not take into account the full width of the profile after its convolution with the Gaussian component. We employ a cut-off in multiples of Voigt widths, given by equation (18) and implemented in equation (19). The $\Delta v_c$ in this approach adapts with both the Lorentzian and Gaussian components ensuring that the wings of the profile are accounted for in an adaptive manner depending on the broadening conditions.

Based on our investigation, we find 500 Voigt widths to be sufficient for current applications including JWST-like resolutions and Very Large Telescope (VLT) applications with small uncertainties when using the standard Voigt profile.

---

**Figure 6.** Comparison of cross-sections obtained at the highest resolution using profile evaluation width ($\Delta v_c$) of 500 Voigt widths versus 10 000 Voigt widths for H$_2$O at 500 K and 0.1 atm. Overestimated and underestimated points are shown in red and blue, respectively. We see that there are some points overestimated by more than 0.1 per cent; however, no points are underestimated by more than 1 per cent. The underestimations average 10 per cent at this high resolution.

The largest differences found are $\sim$10 per cent for pressures of 100 atm, but given that such high pressures are not directly observable for exoplanetary atmospheres, the corresponding differences are less of a concern. We find that in such cases the cut-off chosen $\Delta v_c$ is too narrow for the extreme case of $P = 100$ atm. For the $10^{-4}$–10 atm pressure range, we find the adaptive grid to be very accurate, particularly after binning to the final output grid with a spacing of 0.01 cm$^{-1}$ which corresponds to a spectral resolution of $R = 10^5$–$10^6$ at $\lambda$ <10 $\mu$m.
pressures. These low pressures are crucial to understanding the upper atmospheres and are the most likely to affect observations. Any intensity missed in the initial evaluation of the profile then leads to inaccuracies to each individual profile which can then be further compounded by binning several such transitions to the final output grid. An underestimation of the true area under the profile results in an incorrect normalization and an overestimation of the cross-section value for each transition. We see in Fig. 7 that this overestimation can be 50 per cent leading to a loss of information at these crucial low pressures.

![Percentage of Missing Intensity Compared to Voigt Width Analysis](image)

**Figure 7.** Comparison of the Voigt width cut-off from equation (18) with \( \Delta v_c \) set at 500 Lorentzian widths. For pressures of 0.001 atm and below, we see that missing intensity can be as high as \( \approx 50 \) per cent leading to a loss of information at these crucial low pressures.

Other approaches have included a simple cut-off at a given wavenumber from the centroid of the line which would be very similar to the tests performed in Section 5.3. We find that changes in this cut-off cause small underestimations in the lower intensity transitions across wavelength. However, we find our metric, discussed in Section 6.1, to be quite general even accounting for this effect. At most resolutions, many transitions are binned together mitigating this problem and our metric uses a median over a wide band.

### 6. Effect of Pressure Broadening on Cross-Sections

In this section, we systematically investigate the dependence of molecular cross-sections on the various parameters and assumptions involved in implementing pressure broadening with a given line list. Our goal here is to both quantify the uncertainties propagated into cross-sections due to various choices involved in their generation and to define a quantitatively meaningful approach to make those choices. In order to pursue this, here we focus on the \( \text{H}_2\text{O} \) molecule which has the most complete line list and broadening data currently available for exoplanetary atmospheres. We investigate the dependence of the cross-sections on the following key factors: (a) pressure and temperature, (b) average versus line-by-line treatment of broadening parameters, (c) spectral resolution (i.e. binning) of the cross-sections and (d) broadening agent. Note that the effect of the parameters of the Voigt line profile, namely the profile resolution and extent of the line wings, was investigated in the previous sections and here we adopt the optimal values and methodology discussed in Sections 4.2 and 5.

#### 6.1 Definition of change due to broadening

In order to assess the difference that any given aspect of broadening makes to the cross-sections, we need to formally define the number of points which require evaluation. We also find our method to give sufficiently accurate profile estimates for the final output grid spacing of 0.01 cm\(^{-1}\) and the instrument resolutions given here in later sections.
corresponding change quantitatively. Since the cross-sections for any molecule can span many orders of magnitude over a given spectral range and features very drastically with wavelength, defining a robust metric is challenging. For example, a simple metric such as a ‘mean difference’ across the entire spectral range available is often unreliable. On the other hand, focusing on lines with maximum error places undue emphasis on the lowest intensity lines which will see the highest fractional change but would be hard to observe. Conversely, focusing on the highest intensity lines is unreliable because they are not representative of the line population. Therefore, we use the median percentage difference across the entire spectral range as our metric of choice to quantify the change in cross-sections due to any aspect of broadening. The median percentage change in cross-section for each line is computed as

\[
\delta = \text{Median} \left\{ \frac{\sigma - \sigma_0}{\sigma_0} \right\} \times 100,
\]

where \(\sigma_0\) and \(\sigma\) are the cross-sections before and after incorporating the particular pressure broadening prescription, and the median is evaluated for cross-sections computed over the entire wavelength range of interest, 0.5–20 \(\mu\)m.

### 6.2 Effect of pressure and temperature

The effect of pressure broadening on cross-sections as a function of pressure and temperature is shown in Fig. 9. In the region of interest for exoplanet atmospheres, for pressures around \(10^{-4}–1\) atm and temperatures of 500–2000 K, we find notable changes to the cross-section as a function of resolution. For low resolutions (\(R < 100\)), we find that the differences introduced to cross-sections from pressure broadening are \(<1\) per cent across all \(P\) and \(T\) considered in this work. At resolutions of \(R = 5000\), similar to those expected from JWST, and representative hot Jupiter temperatures of \(T \sim 1000\) K, we find \(\delta < 1\) per cent at \(P = 0.1\) atm. At the highest resolution of \(R = 10^5\), we find that for temperatures of 1000 K and pressures of 0.1 atm, \(\delta = 60\) per cent for H₂O in an H₂ atmosphere. For lower temperatures of 500 K, this can increase dramatically giving \(\delta = 1000\) per cent.

Fig. 2 shows that at low temperatures (\(T \lesssim 500\) K) the wide Lorentzian will have more of an effect than the Gaussian profile. From this we expect to see that cross-sections (and full atmospheric models) at low temperatures are more affected by pressure broadening than those at high temperatures. Fig. 10 shows that, even before propagating through an atmospheric model, we can expect cool targets of \(~500\) K to be more affected by pressure broadening than hotter targets \(~1000\) K by \(100\) per cent or more at the highest resolutions. At resolution of \(R = 10^5\), we can expect the median difference between the two broadening cases (thermal only and pressure and thermal) for water to be \(100\) per cent beyond pressures of 0.1 atm at \(T = 500\) K as shown in Fig. 10. This would lead to uncertainties in molecular abundance of a similar factor. The implication here is that cooler targets are likely to be more affected by pressure broadening. When observing such cool targets with high-resolution instruments, pressure broadening will potentially limit the precision on abundance measurements. These cooler targets are likely to have complex atmospheric structure and incorporate different chemistry in their upper atmosphere, even including broadening from more complex molecules than molecular hydrogen. However, for cool targets, there may be other factors obscuring observations such as clouds and hazes.

In recent years, observational programmes have largely focused on hotter targets, mostly hot Jupiters, as their thick atmospheres and high temperatures lead to stronger spectral features and their short orbital periods make them easier to observe than other targets. Currently, observing lower mass, cooler planets is proving difficult; however, pressure broadening may affect us more in the future when observations become more sensitive to such planets. At temperatures of 1000 K and beyond, we still find an effect from pressure broadening though less strong. For these hotter targets, we might expect pressure broadening to have an effect in the 0.1–1 atm pressure regime of around 10–100 per cent. As discussed in Section 6.5, this depends also on the broadening agent and we find that self-broadening for water is much stronger than H₂.

### 6.3 Line-by-line versus mean broadening parameters

As currently there is a lack of line-by-line broadening parameters for several molecules, it is often unavoidable to rely on sparse broadening data when computing pressure-broadened cross-sections, as discussed in Section 3.3. When only sparse data are available, i.e. broadening parameters are available only for a few lines, in this work representative values for the broadening parameters are chosen based on the available data and the mean applied across all the lines. We find that the difference to cross-sections when using this method is up to \(~20\) per cent at the output grid spacing of 0.01 cm\(^{-1}\) for all pressures. In this section, we investigate the difference such
an approach makes to the cross-sections overall compared to the cases where broadening parameters are available for all the lines.

For purposes of demonstration, we use the latest line list of H$_2$O from the HITMPl data base which includes line-by-line broadening parameters with air-broadening. In one case, we calculate H$_2$O cross-sections over a wide range of temperatures and pressures using detailed line-by-line values for the air-broadening parameters. In another case, we adopt constant values for the broadening parameters averaged over the entire line list and apply those values for broadening every line. The median percentage difference in the cross-sections derived from the two cases over the entire line list at the native spacing of the line list with a grid spacing of 0.01 cm$^{-1}$ is shown in Fig. 11. At this resolution, we find that using mean broadening values can result in cross-sections that are inaccurate by up to 20 per cent for observable pressures (~0.1 bar) and low temperatures ($T \lesssim 500$ K) where pressure broadening is strongest. For lower pressures and higher temperatures, the effect is less pronounced in a ‘median’ sense. For lower resolutions, the differences reduce. Therefore, when numerous lines are available to calculate representative average values for the broadening parameters as in the present case, then the mean treatment of pressure broadening is a reasonable approximation to a detailed treatment for low-resolution observations.

Fig. 11 shows this difference to decrease as pressure increases. In high-pressure cases, the broadening becomes wide enough such that many profiles begin to overlap. This effectively smoothes the information and causes the differences in profile shapes to be less distinguishable. From Fig. 9, this effect is more clear as we see at high pressures much of the information from the individual transitions is lost. This effect occurs at pressures greater than 0.1 atm implying that for high pressures of $P \gtrsim 1$ the detailed broadening parameters for each transition may not affect cross-sections as much.

Contrary to the above scenario where we can construct a reasonable mean from many broadening parameters, detailed line-by-line broadening parameters are required to derive accurate cross-sections, especially from upcoming high-precision and high-resolution observations. In the above case, while the median error across all the lines is low overall, individual lines with significant deviation from the mean broadening values can result in more than 100 per cent difference in cross-sections which are relevant for interpreting high-resolution observations that rely on detecting specific lines (Snellen et al. 2010, 2014). Secondly, for several molecules relevant broadening values are available based on experimental data for only a few lines, the average of which may not be representative of the entire line population, leading to larger inaccuracies than found in the above example. Finally, recent and upcoming observations are already sensitive to cool and dense planetary and brown dwarf atmospheres (e.g. Fraile et al. 2014; Buenzli et al. 2015) with high-precision observations which necessitate accurate line-by-line broadening parameters.

Appendix A1 shows all molecules that are available from various sources with how many lines each contains which here we use as a proxy for how complete a line list is, with those containing only a few thousands of lines being most unreliable. HITMPl, HITRAN and GEISA have line-by-line pressure broadening parameters generated but only for air- and self-broadening, with only HITMPl containing values for the high temperatures relevant for exoplanetary atmospheres. In order to address the issue of pressure broadening accurately in the future, line lists will need to include broadening due to molecular hydrogen, relevant for giant planets, and be complete to high temperatures.

### 6.4 Effect of spectral resolution

One of the most important questions that can be answered in this work is how the difference created by pressure broadening to molecular cross-sections is influenced by spectral resolution. As discussed earlier, we use H$_2$O as our case study and consider cross-sections in...
the HST G141 bandpass (1.1–1.7 \( \mu m \)). Observations of exoplanetary spectra are conducted over a wide range of spectral resolution, ranging from broad-band photometric observations and low-resolution spectra \((R \lesssim 100)\), e.g. with HST, to very high resolution spectra with large ground-based facilities \((R \sim 10^5)\). Resolution can greatly influence how an observed spectrum can be interpreted and can break the degeneracies between different molecules. Here we discuss the effect resolution has on \( \delta \) as a function of pressure and temperature.

We find that at the highest resolutions and lowest temperatures \((R = 10^5\) and \(T = 500\) K), pressure broadening can introduce a difference to the final cross-section of \( \delta = 1000 \) per cent for \(P = 0.1\) atm. For lower resolutions of \(R = 5000\), similar to those that will be attainable with JWST, the differences become much smaller. However, at low temperatures \((T = 500\) K) and high pressures \((P = 1\) atm), a \( \delta \) of 40 per cent is found for H\(_2\)-broadening at \(R = 5000\). For low-resolution spectra \((R \lesssim 100)\) of exoplanets, which are possible with current instruments, for representative exoplanetary temperatures \((T = 500–2500\) K, \(P \lesssim 1\) atm) and H\(_2\)-rich atmospheres, we find the median difference in cross-sections introduced by various aspects of pressure broadening \((\delta)\) to be \( \lesssim 1\) per cent.

For illustration, Fig. 12 shows molecular cross-sections of several molecules binned to the spectral ranges and resolutions achievable with HST (WFC3 G102 and G141 grisms, \(R \sim 100\)) and JWST (NIRSPEC, \(R \sim 3000\)). We can see at the higher resolution of JWST, it is much easier to break the degeneracies and identify molecules. This is also easier in JWST due to the longer spectral range giving more potential to search for molecules.

Fig. 13 shows \( \delta \) as a function of \(P\) and \(T\). We find that \( \delta \) increases with increasing resolution, reaching differences up to 100 per cent or higher for \(R \sim 10^5\), \(P \gtrsim 0.1\) atm and \(T \lesssim 500\) K. For example, as shown in the upper panel of Fig. 13 considering a temperature of 500 K, we find that at a nominal pressure of 1 atm, \( \delta \) can be as high as 100 per cent for \(R = 10^4\) or more. On the other hand, for the highest resolutions possible today of \(R \sim 10^5\), \( \delta \gtrsim 100\) per cent even for pressures as low as 0.1 atm.

Similarly, the lower panel of Fig. 13 shows the variation in \( \delta \) with temperature for a nominal pressure of 1 atm, showing that \( \delta \) can be...
very high ($\gtrsim 100$ per cent) for $T < 1000$ K for $R > 10^4$. Consequently, we find that it is very important for atmospheric models to include pressure broadening when interpreting high-resolution spectra ($R \gtrsim 10^4$) of exoplanet atmospheres observable with current and upcoming facilities (e.g. VLT, Keck and E-ELT). Otherwise, the derived molecular abundances will be limited by a minimum uncertainty of more than 100 per cent due to inaccurate cross-sections. On the other hand, we find that for $R < 10^4$, $\delta$ is reduced reaching a maximum of $\sim 10$ per cent for pressures of relevance to exoplanetary atmospheric observations of $P \lesssim 0.1$ atm. Fig. 14 shows many slices across these plots with both $H_2$-broadening and self-broadening, the latter having an even larger effect on the cross-sections than $H_2$-broadening (as discussed in Section 6.5).

While our results indicate that it is important to include pressure broadening in cross-sections for interpreting high-resolution observations, it is nevertheless advisable to also include the same for low-resolution spectra as well. Even though $\delta$ is found to be at a maximum of $\sim 10$ per cent for $R < 10^4$ and $P \lesssim 0.1$, low-resolution spectra at very high precision with HST and JWST could allow retrieval of molecular abundances with uncertainties of a few per cent, in which case the 10 per cent uncertainty in the cross-sections could become a limiting factor. Secondly, it is to be noted again that $\delta$ is a metric of differences only in a median sense while individual lines could potentially contribute higher $\delta$ than the median value. Finally, while the current analysis focused on $\text{H}_2\text{O}$ with $H_2$-broadening, the same with other molecules could in principle lead to higher $\delta$ even at low resolutions which future studies need to investigate.

### 6.5 Effect of broadening agent

Another important factor in pressure broadening is the primary broadening agent in the planetary atmosphere. As shown in...
equation (4), the broadening agent governs the Lorentzian HWHM of the broadening profile. Here we find that for resolutions of \( R = 100 \,000 \) the difference between broadening in an \( \text{H}_2 \) atmosphere and broadening in a \( \text{H}_2 \text{O} \) atmosphere can be significant for water features such as those in the WFC3 1–1.7 \( \mu \text{m} \) bandpass. At low temperatures of 500 K, we find \( \delta_{\text{H}_2} \sim 1000 \) per cent and \( \delta_{\text{H}_2 \text{O}} \sim 10 \,000 \) per cent for pressures of 0.1 atm at \( R = 100 \,000 \). At higher temperatures, which have been shown to reduce the \( \delta \) found in the previous sections, we find that for \( T = 1000 \) K \( \delta_{\text{H}_2} \sim 60 \) per cent and \( \delta_{\text{H}_2 \text{O}} \sim 300 \) per cent in the same conditions. We find that across pressure, temperature and resolution, the \( \delta \) from self-broadening in \( \text{H}_2 \text{O} \) is on average 4 times greater than that from \( \text{H}_2 \).

Molecular line lists containing pressure broadening data, e.g. HITRAN or HITEMP, typically contain data for self and air as the broadening agents, motivated by the terrestrial applications which HITRAN was originally intended for. However, for giant planetary atmospheres, \( \text{H}_2 \) is the dominant broadener and is of particular relevance for studying atmospheres at high spectral resolution and photometric precision. Accurate line-by-line \( \text{H}_2 \) broadening data for high temperatures are still elusive for most molecules of interest though a few molecules have data available, particularly for \( \text{H}_2 \text{O} \) (Partridge & Schwenke 1997) and more recently for CO (Li et al. 2015).

Here, we investigate the effect of broadening agent on the median accuracy of molecular cross-sections for a representative case. We consider the case of \( \text{H}_2 \text{O} \) for which we have line-by-line broadening parameters with \( \text{H}_2 \), self and air as broadening agents (Partridge & Schwenke 1997; Freedman, personal communication). To illustrate the differences made by changing the broadening agent, we have used cases where the molecule is broadened only by a particular molecule self, air or \( \text{H}_2 \) (i.e. the partial pressure is 1 in each case). Fig. 15 shows the median percentage difference in cross-sections caused by each of the three scenarios compared to Gaussian-only broadening for an illustrative case with resolutions of \( 10^4 \) and \( 10^5 \) and \( T = 500 \) K.

Fig. 15 shows that it is important to carefully choose broadening agents before generating cross-sections for different planet types. First, self-broadening can cause significantly higher \( \delta \) values compared to air- or \( \text{H}_2 \)-broadening at observable pressures (\( P \sim 0.1–1 \) atm). Secondly, the differences between \( \text{H}_2 \)-broadening and air-broadening are relatively small in the \( \text{H}_2 \text{O} \) case. Therefore, while modelling \( \text{H}_2 \text{O} \)-rich atmospheres in the absence of any \( \text{H}_2 \)-broadening data for \( \text{H}_2 \text{O} \) molecules, though not ideal, it is more advisable to use air-broadening than self-broadening. In contrast, when modelling atmospheres of low-mass exoplanets, e.g. super-Earths that can have volatile-rich atmospheres such as \( \text{H}_2 \text{O} \)-rich or \( \text{CO}_2 \)-rich atmospheres, it is important to use cross-sections that are generated with the appropriate broadener. For example, for \( \text{H}_2 \text{O} \)-rich atmospheres, self-broadening of \( \text{H}_2 \text{O} \) should be considered in the cross-sections rather than air-broadening.

The effect of broadening agent can be substantial depending on other parameters. The effect of the broadening agent is naturally strongest in regions of parameter space where pressure broadening is expected to be strongest, namely at high pressures, low temperatures and high resolution. Fig. 14 shows the differences between self- and \( \text{H}_2 \)-broadening for various parameters. The differences begin to become significant at high pressures \( (P \gtrsim 0.1 \) atm) for \( R \gtrsim 10 \,000 \) and become substantial even at lower pressures for high resolutions. For very high resolution observations at \( R \gtrsim 10^5 \), self-broadening can lead to differences of 1000 per cent in cross-sections. This has a much greater effect than \( \text{H}_2 \)-broadening over a large range of pressures, making it critical to use broadening data with the appropriate broadening agents for interpreting observations at these resolutions. Finally, the differences in cross-sections induced by different broadening agents are particularly strong at lower temperatures \( (T \lesssim 1000 \) K) as the broadening has a greater effect at these temperatures.

Given the wide range of temperatures and pressures of exoplanetary atmospheres that are accessible to current and upcoming observations, line-by-line broadening parameters are required in molecular line lists for different broadening agents. In particular, there is a critical need for high-temperature (\( \gtrsim 500 \) K) \( \text{H}_2 \)-broadening data as the most observable atmospheres are those of hot and giant exoplanets with \( \text{H}_2 \)-rich atmospheres, for which very high resolution

Figure 15. Effect of broadening agent on cross-sections. The curves show median difference pressure broadening induces when compared with Gaussian-only cross-sections when broadening agent is changed. Here the \( \text{H}_2 \text{O} \) molecule is considered at \( R = 10 \,000 \) and 100 \,000 at a temperature of 500 K, chosen since lower temperatures provide the highest effect from pressure broadening. The figure illustrates that self-broadening is 400 per cent stronger on average than \( \text{H}_2 \)-broadening and that air- and \( \text{H}_2 \)-broadening are comparable in magnitude.
spectra ($R \sim 10^5$) are also being reported. In the mean time, it is advisable to use air-broadening where available for such atmospheres because while not ideal it provides closer cross-section estimates to H$_2$-broadening for H$_2$O and is an improvement on incorporating no pressure broadening at all. Future studies would also need to investigate if the same is true for other molecules as and when H$_2$-broadening data become available for those molecules. Finally, a realistic atmosphere will contain many different molecules and so contribute broadening from many different species. For smaller planets, we expect atmospheres that are more complex, containing more massive molecules with high abundance. This will affect the pressure broadening particularly as the partial pressure will no longer be 1 and there will be contributions from many species. In such cases, we expect the contribution to vary depending on the abundance of more massive broadening molecules with greater pressure broadening parameters. Our current work gives an estimate for the most extreme cases of H$_2$O- or H$_2$-dominated atmospheres.

### 7 EFFECT OF PRESSURE BROADENING ON TRANSMITTANCE

A thorough investigation of the effect of pressure broadening on fully modelled exoplanetary spectra is non-trivial and beyond the scope of the present study. Many factors such as the inhomogeneous $P$-$T$ structure and composition of the atmosphere will determine the final spectrum. Nevertheless, as a simpler exercise, here we nominally assess the effect of pressure broadening on the transmittance in a fiducial atmosphere represented by a uniform column of gas. As in other parts of this study we use H$_2$O as the only absorbing species, and we consider pressure broadening in an H$_2$-rich atmosphere.

Cross-sections are used in atmospheric codes to determine the resultant intensity transmitted through a column of gas. Usually, the column will undergo some changes in pressure, temperature and number density over the length of the column. These factors combine to give an optical depth $\tau$,$$
\tau = \int n \sigma_\lambda dz,
$$
where $z$ is length through the column, $n$ is the abundance (number density) of the molecule and $\sigma_\lambda$ is the cross-section which is a function of wavelength. The abundance of a molecule is determined by the temperature and pressure of the column by the simple gas law

$$P = nk_\text{B}T.$$  

The optical depth gives a measure of how much intensity will be transmitted through a column of gas based on these properties. For a source intensity ($I_0$), the resultant intensity ($I_\lambda$) at the end of the column is given by

$$I_\lambda = I_0 e^{-\tau},$$

where the scale factor $e^{-\tau}$ is the transmittance.

In order to make a meaningful comparison between the effects of altering the cross-section on emergent intensity as a function of pressure and temperature, we choose to fix $\tau$ to a single value. Here, we assume the column of gas to be at a given constant pressure and temperature, and hence constant density. The length of the column is allowed to vary in order to contain the same $\tau$ regardless of pressure and temperature. As $\tau$ is also a function of wavelength, its value alters depending on the particular molecular feature. To fix $\tau$, we take the value at the peak of the water feature near 1.4 $\mu$m in the WFC3 band. If only the cross-section $\sigma_\lambda$ is changed, between a pressure-broadened case $\sigma_{\text{b}}$ and an unbroadened case $\sigma_{\text{unb}}$, we can then find the effect of pressure broadening has on the transmitted intensity. This is given as

$$\frac{I_{\lambda b} - I_{\lambda u}}{I_{\lambda u}} = e^{\sigma_{\text{b}} - \sigma_{\text{unb}}} dz - e^{\sigma_{\text{b}} - \sigma_{\text{unb}}} dz.$$  

As we assume a gas of the same number density at each pressure and temperature point, then we can assume

$$A = \int \sigma_{\text{b}} dz = n x,$$

where $x$ is some distance scale. Substituting this we find

$$\Delta I_\lambda = \frac{I_{\lambda b} - I_{\lambda u}}{I_{\lambda u}} = (e^{\sigma_{\text{b}} - \sigma_{\text{unb}}} - 1).$$

The above expression gives the relative change to the intensity, and hence the transmittance, for a given abundance and path through a uniform column of gas induced by using cross-sections with pressure broadening compared to those without. Using the cross-section across the WFC3 bandpass of 1–1.7 $\mu$m and binning down to a given resolution, we can find the difference to the transmittance of the column of gas as a function of wavelength. As discussed above, the length of the column is fixed such that the maximum optical depth of the column in the given bandpass equals a fixed parameter ($\tau$), for a given density corresponding to a given temperature and pressure. We can then alter $\tau$ to investigate the optically thin and optically thick regimes as functions of pressure and temperature. We note that the change induced to transmittance ($\Delta I_\lambda$) across a given bandpass is higher at wavelengths with higher absorption, which are also of the wavelengths of interest to observations. We therefore consider the max($\Delta I_\lambda$) in the WFC3 bandpass as our metric of choice in evaluating the effect of pressure broadening on transmittance in that bandpass. This does not account for the effect of signal-to-noise on such observations zero transmittance ($e^{-\tau}$) implies no signal; however, here we consider only values of $e^{-\tau}$ (which can have values between 0 and 1) that are greater than 0.01.

The effect of pressure broadening on the transmittance in our idealized column of gas is similar to the effect on cross-sections discussed in the previous sections. Fig. 16 shows the fractional difference ($\Delta I_\lambda$) pressure broadening makes to the transmittance as a function of several key parameters: the optical depth ($\tau$), pressure ($P$) and resolution ($R$); a nominal temperature of 1000 K is chosen for illustration but the general temperature dependence is discussed below. At the outset, for low resolutions ($R \lesssim 100$), $\Delta I_\lambda$ is $\lesssim 0.1$ per cent across the HST WFC3 G141 bandpass (1.1–1.7 $\mu$m) for almost the entire range of parameters of relevance to exoplanetary atmospheres, particularly for $P < 1$ atm, $T = 500$–3000 K and $\tau < 5$. Naturally, however, $\Delta I_\lambda$ is higher for higher resolutions. Considering nominal values of $\tau < 1$, $P < 0.1$ atm and $T > 1000$ K, we find a maximum $\Delta I_\lambda$ in the WFC3 band to be 6 per cent for a JWST-like medium resolution of $R = 5000$, and 75 per cent for a VLT-like very high resolution of $R = 10^5$.

The $\Delta I_\lambda$ for each resolution increases with increasing pressure and lowering temperature, particularly for very high resolution. For our lowest $T$ of 500 K and $\tau = 1$, for $R = 5000$ $\Delta I_\lambda$ is $\lesssim 12$ per cent for $P < 0.1$ atm, and $\lesssim 65$ per cent for $P < 1$ atm. On the other hand, for the same $T$ and $\tau$, for $R = 10^5$ $\Delta I_\lambda$ is $\lesssim 100$ per cent for $P < 0.1$ atm and $\lesssim 2000$ per cent for $P < 1$ atm. As $\tau$ increases, the difference between the two cases increases as there is more material to modify the intensity. However, as $\tau$ is increased to very
Figure 16. Effect of pressure broadening on the transmittance in an idealized atmosphere (see Section 7). The contours show the maximum percentage difference in transmittance, i.e. the scale factor $e^{-n \sigma_x}$, induced by considering cross-sections with pressure broadening versus those with Gaussian-only thermal broadening in the $HST$ WFC3 bandpass. The relative difference is shown for a wide range in key parameters: pressure ($P$), spectral resolution ($R$) and maximum optical depth in the WFC3 bandpass ($\tau$); a nominal temperature of 1000 K is chosen for illustration but the general temperature dependence is discussed in Section 7. Here the individual points of the $T$, $P$ and $R$ grid have been linearly interpolated over for plotting purposes (see e.g. Hill et al. 2013).

Table 3. Molecules used to generate cross-section data base and all line list sources, including which broadening agents are given. (Note that HITRAN is recommended at room temperature of 296K.) Here we choose many different sources to compare the effect that completion has on our final cross-sections.

| Molecule | Source | Broadening Max T (K) |
|----------|--------|----------------------|
| H$_2$O   | BT2$^a$| Self, air 3000       |
| H$_2$O   | HITREM$^b$ | Self, air 4000     |
| H$_2$O   | HITRAN$^c$ | Self, air 296       |
| H$_2$O   | PS (1997)$^d$ | H$_2$ –        |
| CO$_2$   | HITREM$^b$ | Self, air 4000     |
| CO$_2$   | HITRAN$^c$ | Self, air 296       |
| CO       | HITREM$^b$ | Self, air 4000     |
| OH       | HITREM$^b$ | Self, air 296       |
| NO       | HITREM$^c$ | Self, air 296       |
| NO       | HITRAN$^c$ | Self, air 296       |
| CH$_4$   | YT10to10$^e$ | Self, air 2000   |
| CH$_4$   | HITRAN$^e$ | Self, air 296       |
| NH$_3$   | BYTE$^f$ | Self, air 1600      |
| NH$_3$   | HITRAN$^c$ | Self, air 296       |
| HCN      | Harris$^g$ | Self, air 4000     |
| HCN      | HITRAN$^c$ | Self, air 296       |
| C$_2$H$_2$ | HITRAN$^c$ | Self, air 296       |

$^a$Barber et al. (2006).
$^b$Rothman et al. (2010).
$^c$Rothman et al. (2013).
$^d$Partridge & Schwenke (1997).
$^e$Yurchenko & Tennyson (2014).
$^f$Yurchenko et al. (2011).
$^g$Harris et al. (2006).

high values, the medium becomes optically thick and no light is transmitted in certain wavelength regions.

This approach is simplistic as clearly it does not factor in the changes that could happen within the column in temperature and pressure; however, this does give us a first approximation of the difference induced by changing cross-sections on observations of transmission spectra of exoplanetary atmospheres. In reality, light travels through many layers of an exoplanetary atmosphere, with different temperatures, pressures and densities, before reaching the observer. The results above will hold for a specific pressure but full spectral models of exoplanetary atmospheres, both for transmission spectra and emission spectra, are required for a comprehensive investigation of the effect of pressure broadening discussed in the present work.

8 CROSS-SECTION DATABASE

In this work, we present a range of cross-sections for H$_2$O, CO$_2$, CO, CH$_4$, NH$_3$ and HCN from a range of sources shown in Table 3. These have been investigated with detailed, line-by-line calculations of the Voigt profile with pressure and thermal broadening simultaneously included which no other data base to date provides. The cross-sections span a temperature range of 300–3500 K and pressures of $10^{-4}$–$10^{-2}$ atm. Finally, the cross-sections have been created in a variety of resolutions.

The data in this work benefit not only from the addition of a further dimension of pressure with an accurate broadening profile, but also in being generated uniformly with the same code across
molecules. This ensures low and consistent systematic errors across our data. The full data base is represented in Fig. 17.

8.1 Molecular spectra of observational relevance

Fig. 17 shows cross-sections for each molecule that has currently been addressed in this work. Many of these molecules have had cross-sections computed for different line lists, broadening molecules and for a mean or detailed approach but here only the most complete line list cases are shown.

Each of the cross-sections contains strong molecular features relating to the particular molecule. These features are usually the most ideal regions for observations with low signal-to-noise. A list of the highest intensity features is given in Table 4 with their representative central wavelengths.

8.2 Temperature dependence of cross-sections

There is a strong dependence on temperature in these cross-sections where increasing the temperature raises the transition features in the spectrum, reducing the contrast between the peaks and troughs of each band of transitions. Molecules have a different temperature dependence based on their partition function. Individual molecules can have a stronger or weaker temperature dependence which can have an effect on our observations as in an atmosphere many molecular features will be combined. An example of the effect of such a case is shown in Fig. 18. CH₄ has a stronger temperature dependence than H₂O. This leads to changes in their relative contributions to the combined cross-section as temperature increases. While at low temperatures the two contributions are comparable, at very high temperatures the H₂O cross-section is stronger than the CH₄ cross-section. While it would still be possible to tell these two contributions apart in high-resolution spectra, it would be more difficult to assess the abundance of CH₄ in the presence of H₂O. This is a general effect from the partition functions, but there are subtleties depending on the individual transitions. For example, transitions with high values of the total angular momentum quantum number J, and high lower energy levels, become relatively much stronger at

| Molecule | Feature position (µm) |
|----------|-----------------------|
| H₂O      | 6.61, 5.90, 2.76, 2.67, 2.60, 1.87, 1.36, 1.13, 0.95 |
| CO₂      | 14.95, 4.23           |
| CO       | 4.57, 4.32, 2.68, 2.00 |
| HCN      | 14.00, 7.30, 6.93, 4.73, 3.86, 3.57, 3.00, 2.50, 1.53 |
| NH₃      | 15.96, 12.14, 10.37, 9.23, 6.67, 6.15, 3.00, 2.26, 1.95, 1.51, 1.22, 1.03 |
| CH₄      | 7.70, 7.40, 6.47, 3.42, 3.32, 3.21, 2.37, 1.67 |
| NO       | 5.32, 2.58            |
| OH       | 4.20, 2.25, 1.55      |
9 DISCUSSION AND SUMMARY

In this work, we present a systematic and quantitative investigation of the effects of various aspects of pressure broadening on molecular cross-sections for application to exoplanetary atmospheres. We first use H$_2$O as our primary molecule of choice for this investigation as it has the most complete absorption line data. The factors we investigate include the resolution and evaluation width of Voigt profiles, pressure versus thermal broadening, broadening agent, spectral resolution and completeness of broadening parameters. We investigate in detail the effect of pressure broadening both on the absorption cross-sections of H$_2$O under varied conditions and on the transmittance of a fiducial idealized atmosphere. We use the optimal methods resulting from this investigation to systematically and homogeneously generate a library of pressure-broadened absorption cross-sections for a wide selection of molecules of relevance to exoplanetary atmospheres across a wide range of temperature, pressure and spectral resolution.

This study allows us to address the question of the inaccuracies introduced to molecular absorption cross-sections from pressure broadening, both in the context of current and future observational capabilities. As new instruments come online with improved specifications, we will have access to a wealth of high-resolution data on exoplanet atmospheres. The interpretations of these data sets will be impacted by basic model inputs such as cross-sections which are directly degenerate with the molecular abundances derived using spectral retrieval methods. The comparisons presented here show in detail the magnitude of the errors we can expect in these fundamental inputs to atmospheric models across a range of parameters.

To generate cross-sections, we follow a prescription of mapping line intensities broadened by an appropriate function to a fine ‘sub-grid’ which finely samples the profile of the line. This is iterated over each transition from the source and binned to an output grid with a wider spacing (lower resolution) for further use. The lines are broadened by either a Gaussian-only model (which uses only thermal broadening) or a Voigt profile (which combines the Gaussian thermal and Lorentzian pressure broadening), evaluated using the Faddeeva package (Johnson et al.

When evaluating the Voigt profile on a grid, there are two clear sources of error. First, the grid spacing may be too wide, causing the evaluation of the contribution to each grid point to be poor, leading also to a poor normalization and misrepresenting the line transition intensity. Secondly, the wings of the profile may be cut off prematurely, leading to small fractions of intensity from the wings being missed. This is aggravated by the range of intensity values which span many orders of magnitude within a narrow wavenumber range. This leads to the wings of isolated, high-intensity profiles affecting the continuum level of low-intensity neighbours greatly, which can be underestimated when a cut-off is too narrow.

As discussed in Section 5, we present a method of accurately evaluating the Voigt profile on a fine grid that produces minimal errors in the final cross-section at resolutions of interest. This is achieved with a spacing that is adaptive in temperature and in pressure. The grid we adopt is found to be as accurate as the grid from (Hill et al. 2013) to within 0.2 per cent at pressures of 1 atm or less and gives a vast saving on computational time, particularly for high pressures. We pair this grid with a cut-off value, $\Delta v_c$, of 500 Voigt widths (raised to 1000 above pressures of 1 atm). $\Delta v_c$ describes the separation around the wavenumber centroid of the line transition to which the Voigt profile is evaluated. Having investigated a range of values at different resolutions, we find 500 Voigt widths to be sufficient both to provide good normalization for the profile and to evaluate far into the extensive Lorentzian wings. When compared with other values of $\Delta v_c$ from the literature, we find our value to be more accurate (10–100 per cent) at low pressures ($P = 0.1–0.001$ atm) due to the Voigt width adapting with both temperature and pressure as the Gaussian and Lorentzian components change. This combination of $\Delta v_c$ and an adaptive grid provides low errors for
all resolutions discussed in this work up to \( R = 100 \, 000 \). We find errors of less than 1 per cent (averaging \( \sim 0.2 \) per cent) in the final cross-sections at the peaks of transition features. This increases to 10 per cent within transition features at very low intensity; however, we find such transitions to be less likely to significantly affect observations and modelling results. Beyond this resolution, it may be wise to increase \( \Delta \nu_1 \) and use a finer sampling of the profile.

Here we use a standard Voigt profile, though works such as Ngo et al. (2012) have shown that a change in velocity of the broadening particle can alter the pressure-broadened profile shape. This would likely change the wing shape and alter the continuum from what we present here though we anticipate that difference to be small.

We find the effect of pressure broadening to be varied depending on the resolution, pressure and temperature. We choose to measure the change induced to \( \text{H}_2\text{O} \) cross-sections due to pressure broadening using the median percentage difference over the \( \text{HST} \) WFC3 bandpass (1.1–1.7 \( \mu \)m). This provides a reasonable estimate of the characteristic difference, though it is possible to induce higher changes for specific lines. Generally, the differences are larger for higher resolutions, higher pressures and lower temperatures. For low-resolution spectra (\( R \lesssim 100 \)) of exoplanets, which are possible with current instruments, for representative exoplanetary temperatures (\( T = 500–2500 \, \text{K} \), \( P \lesssim 1 \, \text{atm} \)) and \( \text{H}_2\text{O} \)-rich atmospheres, we find the median difference in cross-sections introduced by various aspects of pressure broadening (\( \delta \)) to be \( \lesssim 1 \) per cent. For higher resolutions (\( R \lesssim 5000 \)), including those attainable with \( \text{VLT} \) and \( \text{E-ELT} \), if cool \( \text{H}_2\text{O} \)-dominated targets were observed. For hotter targets of \( T = 2000 \, \text{K} \), this reduces to 15 per cent, though this is a median over wavelength and can be found to be higher (\( \gtrsim 100 \) per cent) for certain features. For spectral resolutions of \( R = 5000 \) (similar to that achievable with \( \text{VLT} \)), we find that \( \delta \) can be up to 40 per cent. On the other hand, for very high resolution spectra (\( R \sim 10^5 \)), pressure broadening can introduce \( \delta \gtrsim 100 \) per cent, reaching \( \gtrsim 1000 \) per cent for low temperatures (\( T \lesssim 500 \, \text{K} \)), high resolutions (\( R \sim 10^5 \)) and high pressures (\( P \sim 0.1–1 \, \text{atm} \)). Such a case could be found with instruments such as the \( \text{VLT} \) and \( \text{E-ELT} \) if cool \( \text{H}_2\text{O} \)-dominated targets were observed. For hot targets of \( T = 2000 \, \text{K} \), this reduces to 15 per cent, though this is a median over wavelength and can be found to be higher (\( \gtrsim 100 \) per cent) for certain features. For spectral resolutions of \( R = 5000 \) (similar to that achievable with \( \text{VLT} \)), we find that \( \delta \) reduces to \( \lesssim 5 \) per cent for hot targets of \( T = 2000 \, \text{K} \) at \( P = 0.1 \, \text{atm} \).

From this, we can see that even with very high resolutions, current hot Jupiter targets with temperatures of 800–2500 \( \text{K} \) will not be greatly affected by differences induced from pressure broadening with pressures of \( \gtrsim 0.1 \, \text{atm} \). A more significant change is found at pressures of 1 atm or above, though current observations of exoplanetary atmospheres typically probe pressures above 1 atm (Madhusudhan 2012). Data on cool targets at high resolution are currently a distant future prospect and we are unlikely to be affected by this level of uncertainty in the near future with such targets. However, even with the lower temperature end of hot Jupiter targets (\( T \sim 1000 \, \text{K} \)) and modern-day instruments such as the \( \text{VLT} \), pressure broadening can cause discrepancies in the cross-section of 30–200 per cent for \( \text{H}_2\text{O} \)-dominated atmospheres with \( P = 0.1–1 \, \text{atm} \).

Molecular cross-sections are degenerate with abundance in atmospheric models, and any error in cross-sections results in an uncertainty in our abundance measurements. From this work, we find that for cool targets (\( T \sim 500 \, \text{K} \)) at high resolutions (\( R \gtrsim 10^5 \)), we would expect uncertainties in the abundance measurements of at least 100 per cent purely from the cross-section inputs to atmospheric models over those that do not include pressure broadening for \( \text{H}_2\text{O} \)-dominated atmospheres. A true spectrum involves many cross-sections from an atmosphere with many layers of temperature and pressure and an observation through many optical depths which will compound this difference. For a true estimation of the difference pressure broadening will make to abundances, full, rigorous atmospheric models are needed.

Cross-sections have been created from a variety of sources which span different levels of completeness, i.e. the number of transitions for which line data are available, and temperature validity, i.e. the temperature up to which the intensity values and completeness can be trusted. In this work, we focus on \( \text{H}_2\text{O} \) as it is both currently detectable in hot Jupiters and well documented in line list sources. We find that our metric of finding the median percentage difference gives good results even when line lists have low completeness, such as the PS line list for \( \text{H}_2\text{O} \). Due to this, we are able to make comparisons between cross-sections generated from line lists of different sources.

Currently, significant efforts are being made into obtaining data on pressure broadening parameters for molecules in different gasses. However, there are some cases where no pressure broadening data are available for certain molecules. In other cases, it may be that a smaller, low-temperature line list source such as HITRAN has broadening parameters for fewer transitions but a more accurate and more complete line list from sources such as ExoMol does not. To investigate what the best approach is in such situations, a mean approach has been tested where broadening parameters are averaged and the mean parameter is applied to each profile. We find that, when taken across a wide wavelength range, the differences between cross-sections generated in a detailed manner and those with a mean broadening parameter applied are up to 20 per cent, even at very high resolution (\( R > 10^5 \)). However, when looking in detail in a narrow wavelength band, individual lines may be inaccurate at higher pressure due to slight differences in the broadening parameter from the mean. Despite this, using our metric of finding the median percentage difference, we find that mean broadening parameters are still useful to ascertain the magnitude of the change to the cross-section pressure broadening can induce.

We also investigated the influence of different broadening agents (self, air and \( \text{H}_2 \)) on the cross-sections using \( \text{H}_2\text{O} \) as a case study. Generally, self-broadening is significantly stronger than \( \text{H}_2\text{O} \)-broadening, by about a factor of 4 on average across our range of pressure, temperature and resolution. For \( \text{H}_2\text{O} \)-broadening, which is the dominant component for giant exoplanet atmospheres that are most amenable to spectroscopy, we see a smaller effect on cross-sections that due to self-broadening or air-broadening. We find in our current investigations using \( \text{H}_2\text{O} \) that where only self- and air-broadening are available, air-broadening produces a closer result to \( \text{H}_2\text{O} \)-broadening. As with other parameters discussed above, the differences induced to cross-sections due to the different broadening agents are <1 per cent for low resolution (\( R \lesssim 100 \)). For medium resolutions of \( R = 5000 \), we find that \( \delta \sim 10 \) per cent at \( T = 500 \, \text{K} \) and \( P = 0.1 \) for \( \text{H}_2\text{O} \) in an \( \text{H}_2\text{O} \) atmosphere (i.e. self-broadening), whereas \( \delta \sim 1 \) per cent for an \( \text{H}_2\text{O} \) atmosphere (i.e. \( \text{H}_2\text{O} \)-broadening). For hotter targets with higher resolution (\( T = 1000 \, \text{K} \), \( R = 100 \, 000 \)), we find that \( \delta \sim 100 \) per cent for \( P = 0.1 \, \text{atm} \). We find that when looking at the transitions of water in a water-dominated atmosphere, we expect the effect of pressure broadening to be more pronounced implying that pressure broadening is likely to be very important for hotter water-rich targets.

A partial pressure of one has been used here, assuming that only one broadening agent is present, though a combination of agents would be more physical, particularly with He included. Further investigation could be undertaken to find at what concentration other agents affect the broadening profile shape.

A final investigation has been undertaken to assess the difference including pressure broadening in cross-sections makes to the
transmitted intensity through a uniform column of gas, as a function of the pressure, temperature, spectral resolution and optical depth. Our investigation focused on an idealized column of H$_2$-dominated gas with H$_2$O as the only absorber in the HST WFC3 G141 bandpass (1.1–1.7 μm). The results follow the general trends of how each of these parameters influences the cross-sections themselves, as discussed above. For low resolutions ($R \lesssim 100$), we find the relative change in transmittance ($\Delta I/IO$) to be $\lesssim 1$ per cent across the HST WFC3 G141 bandpass (1.1–1.7 μm) for almost the entire range of parameters of relevance to exoplanetary atmospheres. For representative parameters of $\tau < 1$, $P < 0.1$ atm and $T > 1000$ K, $\Delta I/O$ can be up to 6 per cent for a $JWST$-like medium resolution of $R = 5000$, and 75 per cent for a VLT-like very high resolution of $R = 10^5$. $\Delta I/O$ can be even higher for higher $T$, lower $T$ and larger $\tau$. While for $R \lesssim 5000$ the $\Delta I/O$ are still below ~100 per cent, for very high resolutions ($R \sim 10^5$) $\Delta I/O$ can be as high as ~2000 per cent.

Ultimately, our present work suggests that incorporating pressure broadening to compute molecular cross-sections for atmospheric models will be necessary depending on the desired accuracy in molecular abundance estimates retrieved from the spectra. Across all the various factors considered in this work, for low-resolution observations ($R \lesssim 100$) of exoplanetary spectra that are currently possible, e.g. with $HST$, the median differences in cross-sections induced due to accurate pressure broadening are found to be $\lesssim 1$ per cent. For medium resolutions ($R \sim 5000$), similar to those possible with $JWST$, the differences are expected to be at the $\lesssim 40$ per cent level. For very high resolution spectra ($R \sim 10^5$), which are possible with current and future large ground-based telescopes such as VLT and E-ELT, significantly higher differences are possible of 100 per cent or even much larger, depending on other factors discussed above.

With atmospheric characterization becoming an ever more important part of exoplanet research, we can begin to see that pressure broadening will impact us in the future. With medium- and high-resolution spectra of exoplanets, both hot and cool, we can expect our abundance measurements to be affected in some way. Beyond that we may be able to detect and characterize the pressure in the atmospheres of other planets by finding regions of wavelength space particularly affected and using high-resolution spectra. This goal will be difficult to achieve even with a wealth of molecular data at our disposal as signal-to-noise ratios for such data are likely to be low. The structure and dynamics of a full atmosphere lead to a convolution of many profiles making characterization of pressure broadening exceedingly difficult. Other factors such as wind speed and Doppler broadening provide further barriers. These will be the ultimate challenges of the future when we will eventually be able to conduct very high resolution spectroscopy of cool low-mass exoplanets.

ACKNOWLEDGEMENTS
CH acknowledges the support from the Science and Technology Facilities Council (STFC) for PhD funding. We thank the groups that have spent time developing molecular line lists, particularly ExoMol, HITEMP and HITRAN, without which this work would not be possible. We also thank Richard Freedman, Iouli Gordon, Jonathan Tennyson and Sergey Yurchenko for helpful discussions. We would like to thank the anonymous reviewer for a very helpful review.

REFERENCES
Abramowitz M., Stegun I., 1964, Handbook of Mathematical Functions, 5th edn. Dover, New York

BROADENING IN EXOPLANETARY ATMOSPHERES 1447

Antony B. K., Gamache P. R., Szembeck C. D., Niles D. L., Gamache R. R., 2006, Mol. Phys., 104, 2791
Barber R. J., Tennyson J., Harris G. J., Tolchenov R. N., 2006, MNRAS, 368, 1087
Benneke B., Seager S., 2012, ApJ, 753, 100
Bernath P. F., 2014, Phil. Trans. R. Soc. A, 372, 30087
Birnbaum G., 1979, J. Quant. Spectrosc. Radiat. Transfer, 21, 597
Brogi M., Snellen I. A. G., de Kok R. J., Albrecht S., Birkby J., de Mooij E. W. J., 2012, Nature, 486, 502
Buenzli E., Saumon D., Marley M. S., Apai D., Radigan J., Bedin L. R., Reid I. N., Morley C. V., 2015, ApJ, 798, 127
Chamberlain J. W., 1978, Theory of Planetary Atmospheres: An Introduction to Their Physics and Chemistry, Wiley, NY
Deming D. et al., 2013, ApJ, 774, 95
Edwards D. P., Strow L. L., 1991, J. Geophys. Res., 96, 20
Faure A., Wiesenfeld L., Drouin B., Tennyson J., 2013, J. Quant. Spectrosc. Radiat. Transfer, 116, 79
Fischer J., Gamache R., Goldman A., Rothman L., Perrin A., 2003, J. Quant. Spectrosc. Radiat. Transfer, 82, 401
Fraine J. et al., 2014, Nature, 513, 526
Freedman R. S., Marley M. S., Lodders K., 2008, ApJS, 174, 504
Freedman R. S., Lustig-Yaeger J., Fortney J. J., Lupu R. E., Marley M. S., Lodders K., 2014, ApJS, 214, 25
Gamache R. R., Lynch R., Plateaux J. J., Barbe A., 1997, J. Quant. Spectrosc. Radiat. Transfer, 57, 485
Gamache R. R., Lynch R., Neshyba S. P., 1998, J. Quant. Spectrosc. Radiat. Transfer, 59, 319
Gamache R. R., Laraia A. L., Lamrououx J., 2011, Icarus, 213, 720
Gamache R. R., Lamrououx J., Laraia A. L., Hartmann J.-M., Boulet C., 2012, J. Quant. Spectrosc. Radiat. Transfer, 113, 976
Grimm S. L., Heng K., 2015, ApJ, 808, 182
Harris G. J., Tennyson J., Kaminsky B. M., Pavlenko Y. V., Jones H. R. A., 2006, MNRAS, 367, 400
Hill C., Yurchenko S. N., Tennyson J., 2013, Icarus, 226, 1673
Johnson S. G., Cervellino A., Wutke J., 2012, libcerf, Numeric Library for Complex Error Functions. Available at http://apps.jcns.fz-juelich.de/libcerf
Konopacky Q. M., Barman T. S., Macintosh B. A., Marois C., 2013, Science, 339, 1398
Kreidberg L. et al., 2014, ApJ, 793, L27
Kurucz R. L., 1992, Rev. Mex. Astron. Astroph., 23, 45
Lee J.-M., Fletcher L. N., Irwin P. G. J., 2012, MNRAS, 420, 170
Li G., Gordon I. E., Rothman L. S., Tan Y.-M., Hu S.-M., Kassi S., Campargue A., Medvedev E. S., 2015, ApJS, 216, 15
Line M. R. et al., 2013, ApJ, 775, 137
McCullough P. R., Crouzet N., Deming D., Madhusudhan N., 2014, ApJ, 791, 55
Madhusudhan N., 2012, ApJ, 758, 36
Madhusudhan N., Seager S., 2009, ApJ, 707, 24
Madhusudhan N. et al., 2011, Nature, 469, 64
Madhusudhan N., Crouzet N., McCullough P. R., Deming D., Hedges C., 2014a, ApJ, 791, L9
Madhusudhan N., Knutson H., Fortney J. J., Barman T., 2014b, Protostars and Planets VI. Univ. Arizona Press, Tucson, p. 739
Mihalas D., Auer L. H., Mihalas B. R., 1978, ApJ, 213, 739
Miller-Ricci Kempton E., Rauscher E., 2012, ApJ, 751, 117
Ngo N. H., Tran H., Gamache R. R., Hartmann J. M., 2012, Science, 339, 1398
Partridge H., Schwenke D. W., 1997, J. Chem. Phys., 106, 4618
Robert D., Bonamy J., 1979, J. Physique, 40, 409
Rothman L. S. et al., 1998, J. Quant. Spectrosc. Radiat. Transfer, 60, 665
Rothman L. S. et al., 2010, J. Quant. Spectrosc. Radiat. Transfer, 111, 2139
Rothman L. S. et al., 2013, J. Quant. Spectrosc. Radiat. Transfer, 130, 4
Schreier E., 2011, J. Quant. Spectrosc. Radiat. Transfer, 112, 1010

Downloaded from https://academic.oup.com/mnras/article-abstract/458/2/1427/2589008 by University of Cambridge user on 15 January 2021
APPENDIX A

A1 Available line list sources

Available sources for line lists are listed in Table A1 with each molecule that is offered. Some lists such as those from HITRAN, HITEMP and GEISA offer pressure broadening parameters as well as intensity and wavelength for transitions. Here we give the number of transitions for each list for each molecule. Where two sources are available, it is on the whole better to use one with more transitions as it is more complete, though there is also temperature validity to consider. We have used HITEMP’s water line list as the most accurate in this work as it is an update of the BT2 line list. ExoMol, however, does offer the most complete line lists with the highest temperature validity range, so where possible we recommend using their line lists though they do not currently provide pressure broadening parameters.

Table A1. A list of all available line list sources for different molecules with the number of lines available in each. In general, the more complete line lists are preferable not only as more features are evaluated within the molecular spectrum but because the continuum of low-intensity lines is better approximated.

| Molecule          | HITRAN | HITEMP | GEISA | ExoMol | Yueqi | Other | (Source)          |
|-------------------|--------|--------|-------|--------|-------|-------|------------------|
| AlO and isotopes  | –      | –      | –     | 5000 000 | –     | –     | –                |
| BeH               | –      | –      | –     | 7858   | –     | –     | –                |
| C2                | –      | –      | –     | 11 340 | –     | –     | –                |
| C2H2              | 12 613 | –      | 11 340 | –     | –     | –     | –                |
| C2H4              | 18 097 | –      | 18 378 | –     | –     | –     | –                |
| C2H6              | 43 592 | –      | 28 439 | –     | –     | –     | –                |
| C2HD              | –      | –      | 15 512 | –     | –     | –     | –                |
| C2N2              | –      | –      | 2577   | –     | –     | –     | –                |
| C3H4              | –      | –      | 19 001 | –     | –     | –     | –                |
| C3H8              | –      | –      | 8983   | –     | –     | –     | –                |
| C3H12             | 124 126| –      | 119 480| –     | –     | –     | –                |
| C6H6              | –      | –      | 9797   | –     | –     | –     | –                |
| CaH               | –      | –      | –     | 26 980 | 6000  | –     | –                |
| CH and isotopes   | –      | –      | –     | 291    | –     | –     | –                |
| CH4               | 60 033 | –      | 291    | –     | –     | –     | –                |
| CF4               | –      | –      | –     | 53 086 | –     | –     | –                |
| CH and isotopes   | –      | –      | –     | 53 086 | –     | –     | –                |
| C2H3Br            | 18 692 | –      | 36 911 | –     | –     | –     | –                |
| C2H3Cl            | 107 642| –      | 18 344 | –     | –     | –     | –                |
| C2H3CN            | –      | –      | 200    | –     | –     | –     | –                |
| C2H3CN            | 3572   | –      | 49 237 | –     | –     | –     | –                |
| CH3D              | –      | –      | 49 237 | –     | –     | –     | –                |
| C3H3              | –      | –      | 19 897 | –     | –     | –     | –                |
| C3H3              | 19 897 | –      | –     | –     | –     | –     | –                |
| CH4               | 336 830| –      | 240 858| 10 × 10^10 | –     | –     | –                |
| C1O               | 5721   | –      | 7230   | –     | –     | –     | –                |
| C1NO             | 21 988 | –      | 356 899| –     | –     | –     | –                |
| CN                | –      | –      | –     | 195 120| –     | –     | –                |
| CO                | 1019   | 113 631| 13 515 | –     | –     | –     | –                |
| CO2               | 169 292| 11 193 608| 413 524| –     | 573 881 316| –     | Tashkun & Perevalov (2011) |
| COF2              | 168 793| –      | 70 904 | –     | –     | –     | –                |
| CP                | –      | –      | –     | 13 825 | –     | –     | –                |
| CH                | 5129   | –      | –     | 93 040 | –     | –     | –                |
| CS                | –      | –      | –     | 93 040 | –     | –     | –                |
| FeH               | –      | –      | –     | 204 688| –     | –     | –                |
| GeH4              | –      | –      | 824   | –     | –     | –     | –                |
| H2                | 4017   | –      | –     | –     | –     | –     | –                |
| H2CO              | –      | –      | 37 050 | –     | –     | –     | –                |
| Molecule | HITRAN | HITEMP | GEISA | ExoMol | Yueqi | Other (Source) |
|----------|--------|--------|-------|--------|-------|----------------|
| H2O      | 142 045 | 114 241 164 | 67 504 | 505 000 000 | – | 296 000 000 \ Partridge & Schwenke (1997) |
| H2O2     | 126 983 | – | 126 983 | – | – | – |
| H2S      | 36 561 | – | 20 788 | – | – | – |
| HBr      | 3039 | – | 1294 | – | – | – |
| HC3N     | 180 332 | – | 179 347 | – | – | – |
| HCl      | 11 879 | – | 533 | – | 2588 | – |
| HCN      | 2955 | – | 81 889 | 34 418 408 | – | – |
| HCOOH    | 62 684 | – | 62 684 | – | – | – |
| HDO      | – | – | – | 700 000 000 | – | – |
| HF       | 10 073 | – | 107 | – | – | – |
| HI       | 3161 | – | 806 | – | – | – |
| HNC      | – | – | 5619 | – | – | – |
| HNO3     | – | – | 669 988 | – | – | – |
| HNO3     | 903 854 | – | – | – | – | – |
| HO2      | 38 804 | – | 38 804 | – | – | – |
| HOBr     | 2177 | – | – | – | – | – |
| HOCI     | 8877 | – | 17 862 | – | – | – |
| Kcl and isotopes | – | – | – | 70000000 | – | – |
| LiH      | – | – | – | 18 982 | – | – |
| MgH and isotopes | – | – | – | 6716 | – | – |
| N2       | 1107 | – | 120 | – | – | – |
| N2O      | 33 074 | – | 50 633 | – | – | – |
| NaCl and isotopes | – | – | – | 50000000 | – | – |
| NaD      | – | – | – | 167 224 | – | – |
| NaH      | – | – | – | 79 898 | – | – |
| NH       | – | – | – | 10 425 | – | – |
| NH3      | 45 302 | – | 29 082 | 1138 323 351 | – | – |
| NO       | 103 710 | 115 610 | 105 079 | – | – | – |
| NO+      | 1206 | – | 1206 | – | – | – |
| NO2      | 104 223 | – | 104 223 | – | – | – |
| O        | 2 | – | – | – | – | – |
| O2       | 1787 | – | 6428 | – | – | – |
| O3       | 261 886 | – | 389 378 | – | – | – |
| OCS      | 15 618 | – | 33 809 | – | – | – |
| OH       | 30 772 | 41 557 | 42 866 | – | – | – |
| PH       | – | – | – | 1.68 × 10^{10} | – | – |
| PH3      | 22 189 | – | 20 364 | – | – | – |
| PN and isotopes | – | – | – | 700 000 | – | – |
| ScH      | – | – | – | 1152 826 | – | – |
| SiO and isotopes | – | – | – | 92398 | – | – |
| SiO2     | 2889 065 | – | 92398 | – | – | – |
| SO2      | 72 460 | – | 68728 | – | – | – |
| SO3      | 10 188 | – | – | 174 674 257 | – | – |
| TiH      | – | – | – | 181 080 | 157 430 | Bernath (2014) |
| TiO      | – | – | – | 8325 354 | Kurucz (1992) |