ExoMol line lists – LVII: High accuracy ro-vibrational line list for methane (CH_4)

Sergei N. Yurchenko, Alec Owens, Kyriaki Kefala, Jonathan Tennyson,*
[1] *Department of Physics and Astronomy, University College London, Gower Street, WC1E 6BT London, UK*

11 January 2024

ABSTRACT

The MM ro-vibrational line list for methane ($^{12}CH_4$) is presented; MM covers wavelengths $\lambda > 0.83$ μm (wavenumbers up to 12 000 cm^{-1}) and contains over 50 billion transitions between 9 155 208 states with total angular momentum $J \leq 60$. MM was generated through solution of the nuclear motion Schrödinger equation using variational program TROVE for an empirically-derived potential energy surface (PES) and a new high-level *ab initio* dipole moment surface. The PES was constructed by fitting the ro-vibrational energies of CH_4 to a set of highly accurate, experimentally-derived energies. Molecular states are classified using the T_d(M) symmetry group and are fully assigned with rotation and vibration quantum numbers. The MM line list is adapted to high-resolution applications by replacing the calculated ro-vibrational energies with the experimentally-derived values where available, namely for 23 208 states with $J \leq 27$ below 9986 cm^{-1}. Doing so leads to over 1 000 000 experimentally-derived CH_4 lines compared to approximately 330 000 lines of $^{12}CH_4$ in the HITRAN database. The MM line list is shown to be more complete than the recent HITEMP methane line list. Methane spectra computed using MM across a broad range of temperatures and wavenumbers show excellent agreement with experiment. The MM line list supersedes the previous ExoMol methane line lists 10to10 and 30to10 both in terms accuracy and coverage. Together with the pre-computed ExoMolOP molecular atmospheric opacity tables, it is now the recommended CH_4 dataset in the ExoMol database (`www.exomol.com`).

Key words: molecular data - exoplanets - stars: atmospheres - stars: low-mass

1 INTRODUCTION

Methane was detected in the exoplanet HD189733b by Swain et al. (2008). Since then, there were many more tentative detections of CH_4 in hot exoplanetary atmospheres using both the low and high resolution methods (see (Beaulieu et al. 2011) and references therein). Although methane is expected to be have significant presence in H_2-rich atmospheres at temperatures below ∼600 K (Burrows & Sharp 1999; Lodders & Fegley 2002), there were no robust detection of CH_4 at these temperatures, see. e.g. Stevenson et al. (2010); Knutson et al. (2014); Benneke et al. (2019b,a); Blain et al. (2021). Very recently, methane was detected in the JWST spectra of K2-18b at $\sim 5\sigma$ confidence by Madhusudhan et al. (2023), thus resolving the problem of the lack of CH_4 in the atmospheres of low temperature exoplanets (Madhusudhan & Seager 2011).

Methane is one of the dominant constituents of atmospheres of Brown dwarfs (Bochanski et al. 2011; Canty et al. 2015). The quality of the line lists have been shown to play critical role for the atmospheric retrievals of these objects, both in therms of the completeness (Yurchenko et al. 2014) and accuracy (Hood et al. 2023). JWST now provides an ideal platform for the spectroscopy of CH_4 in even cooler objects, Y dwarfs (Lacy & Burrows 2023).

The first ExoMol line list for methane, 10to10 (Yurchenko & Tennyson 2014; Yurchenko et al. 2014), containing almost 10 billion transitions, was developed back in 2014 to meet the huge demand for methane spectroscopic data in the exoplanetary and cool star research. 10to10 revolutionised atmospheric retrievals, demonstrating the considerable importance of hot transitions at elevated temperatures (Yurchenko et al. 2014). Many exoplanetary atmospheric studies have utilised the 10to10 line list (Tennyson & Yurchenko 2021), such as for example, the discovery of carbon and nitrogen bearing species in the hot Jupiter exoplanet HD 209458 b (Giacobbe et al. 2021). Although 10to10 filled the gaps in methane high-temperature absorption by providing the necessary completeness, it was shown to be not sufficiently accurate in NIR, especially for high-resolution spectroscopic applications of brown dwarfs and exoplanets (Hood et al. 2023). The higher resolution Doppler spectroscopy technique (Snellen 2014; Birkby 2018) is now an integral part of modern exoplanetary atmospheric studies (Brogi & Line 2019; Brogi & Birkby 2021) and is providing definitive detections of molecules based on high-resolution laboratory spectroscopic data (Brogi et al. 2012).

The 10to10 line list was based on an empirically-refined

* The corresponding author: j.tennyson@ucl.ac.uk

potential energy surface (PES) constructed through fitting to high-resolution experimental data available from the spectroscopic literature at that time. Since then, a large number of high-quality laboratory methane data has been made available (see a comprehensive review of the experimental literature on the methane experimental spectroscopy in Kefala et al. (2024)). The line list production in the ExoMol database has also seen significant developments, most notably towards the high-resolution applications, see, for example, Bowesman et al. (2021). The current ExoMol line list production relies on the extraction of all meaningful experimental ro-vibrational line positions from the literature and subsequent processing them through the MARVEL (Measured Active Rotational Vibrational Energy Levels) algorithm (Furtenbacher et al. 2007; Császár et al. 2007; Furtenbacher & Császár 2012; Tóbiás et al. 2019; Tennyson et al. 2023). MARVEL produces a consistent dataset of highly accurate empirical-quality energy levels which serves two purposes: (i) to refine the PES for the improvement of the accuracy of the predicted line positions, and (ii) to replace the computed line list energy levels with the more accurate MARVEL values. The latter process is known as a 'MARVELisation' procedure. In this work, we use a newly established MARVEL dataset of $^{12}CH_4$ energy levels compiled from the most up-to-date experimental spectroscopic data (Kefala et al. 2024). Our MARVEL dataset of CH_4 contains 23 264 states determined from 82173 transition frequencies from 96 experimental sources. It is the most comprehensive experimentally-derived set of CH_4 energies that we know of and it is central to the production of the MM line list. Using the variational code TROVE (Yurchenko et al. 2007), we have computed 50 395 644 806 lines, covering 9 155 208 ro-vibrational states for $J \leq 60$, with transitions in the wavenumber range 0 to 12000 cm^{-1} for upper state energies up to $E' = 18\,000$ cm^{-1} and lower state energies $E'' = 10\,000$ cm^{-1}. The calculated CH_4 energies have been MARVELised using values from Kefala et al. (2024) and will be continually improved upon as new experimental data is published. The MM line list is recommended to replace the outdated 10to10 and 30to10 ExoMol line lists.

The importance of methane has led to the construction of many line lists (Borysow et al. 2003; Hargreaves et al. 2012; Yurchenko & Tennyson 2014; Nikitin et al. 2014; Hargreaves et al. 2015; Rey et al. 2017; Makhnev et al. 2018; Hargreaves et al. 2020). Here we concentrate on those complete enough to be reliable at higher temperatures. The TheoReTS database (Rey et al. 2016) contains the CH_4 line list of Rey et al. (2014, 2017) which has been shown to be accurate for a large range of temperatures (Wong et al. 2019), and to outperform the 10to10 line list in the accuracy of line positions, particularly at higher frequencies. The TheoReTS CH_4 line list has been very useful in the analysis of laboratory data (Rodina et al. 2019; Nikitin et al. 2020) and facilitated a series of methane data releases in the past few years, see e.g. Kefala et al. (2024). The line list of Rey et al. (2014, 2017) forms the basis of the HITEMP CH_4 line list (Hargreaves et al. 2020), a hot synthetic line list applicable for temperatures up to 2000 K. This line list has been adopted by numerous exoplanetary groups, especially those specialising in high-resolution retrievals, see, e.g. Gandhi et al. (2020). The TheoReTS line list also utilised an empirically constructed PES of CH_4 and was continuously improved by replacing the calculated ro-vibrational line positions with experimental values.

2 POTENTIAL ENERGY SURFACE

A new empirical PES of CH_4 was constructed for the present work. We started from the highly accurate *ab initio* potential energy surface (PES) of methane by Owens et al. (2016), which is given analytically as an expansion in terms of 1D coordinates. Morse oscillator functions describe the stretch coordinates,

$$\xi_i = 1 - \exp\left(-a(r_i - r_{\rm ref})\right) \; ; \quad i = 1, 2, 3, 4 \tag{1}$$

while symmetrised combinations of inter-bond angles are used for the angular terms,

$$\xi_5 = \frac{1}{\sqrt{12}}\left(2\alpha_{12} - \alpha_{13} - \alpha_{14} - \alpha_{23} - \alpha_{24} + 2\alpha_{34}\right), \tag{2}$$

$$\xi_6 = \frac{1}{2}\left(\alpha_{13} - \alpha_{14} - \alpha_{23} + \alpha_{24}\right), \tag{3}$$

$$\xi_7 = \frac{1}{\sqrt{2}}\left(\alpha_{24} - \alpha_{13}\right), \tag{4}$$

$$\xi_8 = \frac{1}{\sqrt{2}}\left(\alpha_{23} - \alpha_{14}\right), \tag{5}$$

$$\xi_9 = \frac{1}{\sqrt{2}}\left(\alpha_{34} - \alpha_{12}\right). \tag{6}$$

Here r_1, r_2, r_3, r_4 are the C–H bond lengths and α_{12}, α_{13}, α_{14}, α_{23} and α_{24} are the five $\angle(H_j\text{–C–}H_k)$ inter-bond angles where j and k label the respective hydrogen atoms; a and $r_{\rm ref}$ are the Morse and reference equilibrium structural parameter, respectively.

The coordinate transformation from the six bending coordinates α_{ij} to the five independent coordinates $\xi_5, \xi_6, \xi_7, \xi_8$, and ξ_9 defined by Eqs. (2–6) has to be complemented by the redundancy condition (Halonen 1997),

$$\begin{vmatrix} 1 & \cos\alpha_{12} & \cos\alpha_{13} & \cos\alpha_{14} \\ \cos\alpha_{12} & 1 & \cos\alpha_{23} & \cos\alpha_{24} \\ \cos\alpha_{13} & \cos\alpha_{23} & 1 & \cos\alpha_{34} \\ \cos\alpha_{14} & \cos\alpha_{24} & \cos\alpha_{34} & 1 \end{vmatrix} = 0. \tag{7}$$

The potential energy function was represented by a symmetry adapted expansion

$$V(\xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6, \xi_7, \xi_8, \xi_9) = \sum_{ijk\ldots} \mathrm{f}_{ijk\ldots} V_{ijk\ldots}, \tag{8}$$

where $V_{ijk\ldots}$ are symmetrised combinations of different permutations of the coordinates ξ_i that transform according to the A_1 representation of the $\mathcal{T}_{\rm d}$(M) molecular symmetry group (Bunker & Jensen 1998a),

$$V_{ijk\ldots} = \{\xi_1^i \xi_2^j \xi_3^k \xi_4^l \xi_5^m \xi_6^n \xi_7^p \xi_8^q \xi_9^r\}^{\mathcal{T}_{\rm d}}. \tag{9}$$

The *ab initio* PES of CH_4 (Owens et al. 2016) was then refined by fitting the expansion parameters $\mathrm{f}_{ijk\ldots}$ in Eq. (9) to the MARVEL energies of CH_4 (Kefala et al. 2024) using the simultaneous fitting procedure by Yurchenko et al. (2003) as implemented in TROVE (Yurchenko et al. 2007), see also Yurchenko et al. (2011) and Yurchenko (2023). In

this procedure, the newly determined PES is constrained to the original *ab initio* PES to ensure the refinement does not produce nonphysical deformations of the surface. Full details of the variational calculations are given below in Sec. 4.

Only the $J = 0, 1, 2, 3, 4, 5$ and 8 energies were used in the fitting and only 77 potential parameters were refined. The quality of the refinement is illustrated in Fig. 1, where the theoretical ro-vibrational energies of CH_4 computed with the refined PES are compared to the experimentally derived MARVEL energies of CH_4Àll 23 292 MARVEL energies from Kefala et al. (2024) covering rotational excitations up to $J = 27$ are reproduced with an rms error of 0.14 cm^{-1}. If we exclude MARVEL energies that only have one transition supporting them as possible outliers, i.e. no combination differences, the rms error for the resulting 16 894 values reduces to 0.11 cm^{-1}.

3 DIPOLE MOMENT SURFACE

The electric dipole moment surfaces (DMSs) of CH_4 representing its μ_x, μ_y, and μ_z dipole moment components were computed *ab initio* as first-derivatives of the electronic energy with respect to an external electric field. To this end, the central finite difference scheme was used with an external electric field of ± 0.005 a.u. along each x, y and y Cartesian coordinate axes. Calculations were done at the CCSD(T)/aug-cc-pVQZ level of theory in the frozen core approximation using the quantum chemistry package MOLPRO (Werner et al. 2015). We used the same nine-dimensional grid of 97 721 nuclear geometries covering energies up to $hc \cdot 50\,000\,cm^{-1}$ that was used to calculate the *ab initio* PES of Owens et al. (2016).

To construct the DMSs of CH_4 analytically, we followed the procedure of Yurchenko et al. (2013) and Owens et al. (2017) based on the symmetrised molecular bond (SMB) representation of DMSs of XY_4-type molecules. In this representation, the molecular dipole moment vector is projected onto the symmetrically independent reference vectors $\mathbf{n}_i$ $(i = 1, 2, 3)$ given by,

$$\mathbf{n}_1 = \frac{1}{2}\left(\mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3 - \mathbf{e}_4\right), \tag{10}$$

$$\mathbf{n}_2 = \frac{1}{2}\left(\mathbf{e}_1 - \mathbf{e}_2 - \mathbf{e}_3 + \mathbf{e}_4\right), \tag{11}$$

$$\mathbf{n}_3 = \frac{1}{2}\left(\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3 - \mathbf{e}_4\right), \tag{12}$$

where the unit vectors along the four C–H bonds are,

$$\mathbf{e}_i = \frac{\mathbf{r}_i - \mathbf{r}_0}{|\mathbf{r}_i - \mathbf{r}_0|}; \quad i = 1, 2, 3, 4. \tag{13}$$

Here, $\mathbf{r}_0$ is the position vector of the C nucleus and $\mathbf{r}_i$ is the position of the respective H_i atom. The three reference vectors $\mathbf{n}_i$ span the F_2 irreducible representation in $\mathcal{T}_d$.

The *ab initio* dipole moment vector μ is given by

$$\mu = \mu_x \mathbf{n}_1 + \mu_y \mathbf{n}_2 + \mu_z \mathbf{n}_3, \tag{14}$$

where μ_α $(\alpha = x, y, z)$ are the dipole moment surfaces (also of F_2 symmetry) taking the form,

$$\mu_\alpha(\xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6, \xi_7, \xi_8, \xi_9) = \sum_{ijk\ldots} \mathrm{F}^{(\alpha)}_{ijk\ldots} \mu^{F_2}_{\alpha,ijk\ldots}. \tag{15}$$

The expansion terms $\mu^{F_2}_{\alpha,ijk\ldots}$ are constructed as symmetrised F_2 symmetry combinations of different permutations of the coordinates ξ_i,

$$\mu^{F_2}_{\alpha,ijk\ldots} = \{\xi_1^i \xi_2^j \xi_3^k \xi_4^l \xi_5^m \xi_6^n \xi_7^p \xi_8^q \xi_9^r\}^{F_{2\alpha}}, \tag{16}$$

where we use,

$$\xi_i = (r_i - r^{\mathrm{ref}}) \exp\left(-\beta (r_i - r^{\mathrm{ref}})^2\right); \quad i = 1, 2, 3, 4, \tag{17}$$

for the stretches and the angular coordinates as in Eqs. (2 – 6). The four stretching coordinates span the $F_{2\alpha}$ representation of the $\mathcal{T}_d$(M) symmetry group. A sixth order expansion was employed. The expansion coefficients $\mathrm{F}^{(\alpha)}_{ijk\ldots}$ $\alpha = x, y, z$ were determined simultaneously through a least squares fitting to the *ab initio* data. A total of 265 parameters were used (including the reference parameter $r^{\mathrm{ref}} = 1.256$ Å) and reproduced the *ab initio* data with a root-mean-square (rms) error of 0.0049 Debye. All parameters defining the PES and DMSs of CH_4 are provided as a supplementary material along with Fortran 95 routines to construct the corresponding analytic representations. We also provide an input file for the TROVE program, which contains both the parameters and the subroutines, respectively.

4 VARIATIONAL CALCULATIONS

To calculate the ro-vibrational spectrum (or line list) of CH_4 we solve the Schrödinger equation for the motion of the C and H nuclei using the variational program TROVE (Yurchenko et al. 2007). TROVE is an open-access[1] nuclear motion code for arbitrary, medium size polyatomic molecules. It can provide a complete solution for modern spectroscopic high-resolution applications, including the refinement of the PES by fitting to experimental ro-vibrational energies, and computing extensive molecular line lists (Tennyson & Yurchenko 2017). The calculation setup is very similar to that employed in Yurchenko & Tennyson (2014), but here we utilise a more accurate kinetic energy operator (KEO) and PES, and a larger ro-vibrational basis set.

4.1 Kinetic energy operator

The KEO was represented in an Eckart frame, see, e.g. Yurchenko (2023), as a sixth-order expansion in terms of nine curvilinear coordinates given by,

$$\xi_i = r_i - r_e, \quad i = 1, 2, 3, 4, \tag{18}$$

and ξ_j, $i = 5, 6, 7, 8, 9$ as in Eqs. (2 – 6). Here, r_e is the equilibrium value of the bond length C–H_i. The expansion was performed using the automatic differentiation method implemented in TROVE (Yachmenev & Yurchenko 2015). Our previous CH_4 line list 10to10 was computed using a similar KEO expansion but in terms of linearised coordinates, which possessed slower convergence (Yachmenev & Yurchenko 2015).

[1] github.org/exomol

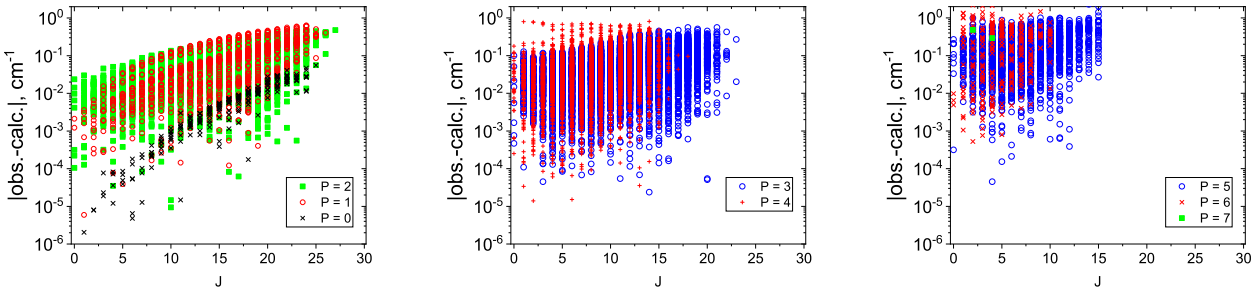

Figure 1. Residual fitting errors of the calculated (TROVE) energy term values (cm^{-1}) of CH_4 as obs.-calc. differences from the experimentally derived (MARVEL) values as function of J for different polyad numbers P.

4.2 Basis set

The vibrational basis wavefunctions were constructed using a multi-step contraction scheme detailed below, based on the the following polyad condition applied at each step

$$P = 2(v_1 + v_2 + v_3 + v_4) + v_5 + v_6 + v_7 + v_8 + v_9 \leq 14, \quad (19)$$

where v_i is a vibrational quantum number associated with a one-dimensional primitive basis function $\phi_{v_i}(\xi_i)$ $(i = 1, 2, 3, \ldots, 9)$. In the conventional normal mode representation this condition corresponds to

$$P = 2(n_1 + n_3) + n_2 + n_4, \quad (20)$$

where n_1, n_2, n_3 and n_4 are the standard normal mode quantum numbers used for methane (Albert et al. 2009). This is a significant improvement in the accuracy comparing to the 10to10 model, where a $P = 10$ basis set was used.

At contraction step 1, nine sets of one-dimensional primitive basis function $\phi_{v_i}(\xi_i)$ $(i = 1, 2, 3, \ldots, 9)$ are generated using the Numerov-Cooley (Cooley 1961; Noumerov 1924) method by solving the 1D Schrödinger equations for 1D Hamiltonian operators obtained by freezing all but one vibrational coordinate in the 9D Hamiltonian operator ($J = 0$) at their equilibrium values.

At step 2, the 9D coordinate space is divided into the three reduced sub-spaces, (i) $\{\xi_1, \xi_2, \xi_3, \xi_4\}$, (ii) ξ_5, ξ_6, and (iii) ξ_7, ξ_8, ξ_9 as dictated by $\mathcal{T}_\mathrm{d}$ symmetry; each of the sub-spaces is symmetrically independent and can be processed separately. For each of these sub-spaces, three reduced Hamiltonian operators were constructed as follows,

$$\hat{H}^{(1)}(\xi_1, \xi_2, \xi_3, \xi_4) = \langle 0_5, 0_6, 0_7 | \langle 0_8, 0_9 | \hat{H} | 0_8, 0_9 \rangle | 0_5, 0_6, 0_7 \rangle, \quad (21)$$

$$\hat{H}^{(2)}(\xi_5, \xi_6) = \langle 0_1 | \langle 0_2 | \langle 0_3 | \langle 0_4 | \langle 0_7, 0_8, 0_9 | \hat{H} | 0_7, 0_8, 0_9 \rangle | 0_4 \rangle | 0_3 \rangle | 0_2 \rangle | 0_1 \rangle, \quad (22)$$

$$\hat{H}^{(3)}(\xi_7, \xi_8, \xi_9) = \langle 0_1 | \langle 0_2 | \langle 0_3 | \langle 0_4 | \langle 0_5, 0_6 | \hat{H} | 0_5, 0_6 \rangle | 0_4 \rangle | 0_3 \rangle | 0_2 \rangle | 0_1 \rangle, \quad (23)$$

where $|0\rangle_k$ are ground state vibrational basis functions $\phi_0(\xi_k)$ $(k = 1, \ldots, 9)$. These yield the following three eigenvalue problems,

$$\hat{H}^{(i)}(\xi^{(i)}) \Phi^{(i)}_{\lambda_i}(\xi^{(i)}) = E_\lambda \Phi^{(i)}_{\lambda_s}(\xi^{(i)}). \quad (24)$$

Because of the property of the Hamiltonian operator to commute with the symmetry operations of the group in question, see Yurchenko (2023), such as $\mathcal{T}_\mathrm{d}(\mathrm{M})$, all eigenvectors of Eq. (24) are expected to transform according to some irreducible representations $\Gamma \in \mathcal{T}_\mathrm{d}(\mathrm{M})$. TROVE uses standard irreducible representation techniques together with a symmetry sampling method (Yurchenko et al. 2017) to classify these eigenfunctions according to the group symmetry.

The three sets of eigenfunctions $\Phi^{(\mathrm{i})}_{\lambda_1}(\xi_1, \xi_2, \xi_3, \xi_4)$, $\Phi^{(\mathrm{ii})}_{\lambda_2}(\xi_5, \xi_6)$ and $\Phi^{(\mathrm{iii})}_{\lambda_3}(\xi_7, \xi_8, \xi_9)$ resulting from these solutions were symmetrised and assigned local mode quantum numbers (i) v_1, v_2, v_3, v_4, (ii) v_5, v_6, and (iii) v_7, v_8, v_9 and vibrational symmetry Γ_vib. The final 9D vibrational basis functions $\Phi^{\Gamma_a}_{\lambda_1,\lambda_2,\lambda_3}$ were then formed as symmetry adapted combinations of products $\Phi^{(1)}_{\lambda_1}$, $\Phi^{(2)}_{\lambda_2}$ and $\Phi^{(3)}_{\lambda_3}$,

$$\Phi^{\Gamma_a}_{\lambda_1,\lambda_2,\lambda_3} = \{\Phi^{(1)}_{\lambda_1} \times \Phi^{(2)}_{\lambda_2} \times \Phi^{(3)}_{\lambda_3}\}^{\Gamma_a},$$

where Γ_a is an irreducible representation of $\mathcal{T}_\mathrm{d}(\mathrm{M})$,

$$\Gamma_a = \{A_1, A_2, E_a, E_b, F_{1x}, F_{1y}, F_{1z}, F_{2x}, F_{2y}, F_{2z}\}.$$

The pure vibrational ($J = 0$) 9D Schrödinger equations were solved on this contracted symmetry adapted basis set Φ^Γ_λ

$$\hat{H} \Phi^{J=0,\Gamma}_\lambda = E^{J=0,\Gamma}_\lambda \Phi^{J=0,\Gamma}_\lambda.$$

At step 3, the $J = 0$ eigenfunctions $\Phi^{J=0,\Gamma}_\lambda$ were turned into a vibrational basis set. To this end, all vibrational matrix elements of the Hamiltonian as well as of the DMSs, originally computed in the primitive basis set, were transformed to the $J = 0$ representation. In order to reduce the otherwise impractically large $J = 0$ basis set even more, it is further truncated using the intensity basis set pruning (IBSP) introduced by Owens et al. (2017), see also Mant et al. (2018) who it applied it to C_2H_4 calculations. The idea of IBSP is to remove specific vibrational functions of states with only a marginal contribution to the absorption opacity (smaller than some intensity threshold I_max).

The vibrational absorption band intensities were computed between all $J = 0$ states of CH_4. These intensities (cm/molecule) are shown in Fig. 2, where we used a high

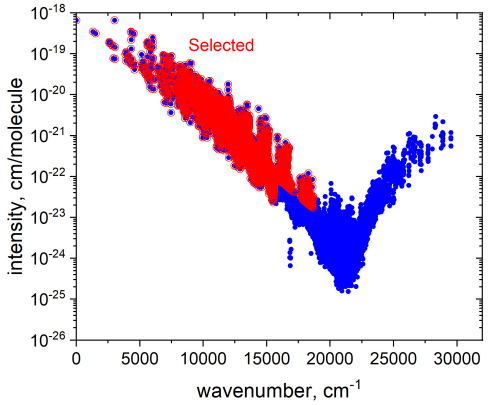

Figure 2. Intensity pruning: Blue filled circles are pure vibrational intensities (cm/molecule) at $T = 3000$ K as a function of the upper state energy term values. Red open circles indicate states selected for the final vibrational basis set in ro-vibrational calculations.

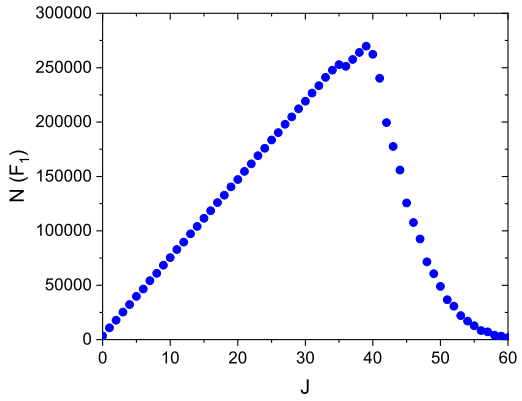

Figure 3. The dimension of the F_1 Hamiltonian matrices as a function of J.

$T = 3000$ K as a reference value describing highest excitation. The intensities at $\tilde{E} > 22000$ cm^{-1} appear to grow in a nonphysical manner, indicating some numerical issues in the TROVE calculations for very high excitations. Before this effect kicks in, the intensities of the high overtone transitions reduce exponentially by 4 orders of magnitude down to 10^{-22} cm/molecule at $T = 3000$ K The corresponding upper energies above some wavenumber threshold (chosen to be hc 18 700 cm^{-1}) are therefore omitted from the final vibrational basis set as not very significant spectroscopically.

We also removed certain states with weaker band intensities above hc 15 600 cm^{-1} shown in Fig. 2, using an empirical threshold of $I_{\rm vib} < 1.2 \times 10^{-18} \exp(-0.0006 \cdot \tilde{E}'{\rm cm})$, where the coefficients were selected to reduce the basis set to a manageable size for the high-performance computing facilities available to us (see below). The final pruned vibrational ($J = 0$) basis set consisted of 28 758 functions: 1336 of A_1, 1025 of A_2, 2439×2 of E, 3419×3 of F_1 and 3754×3 of F_2 symmetry.

At step 4, the $J = 0$ eigenfunctions were used to construct the ro-vibrational basis functions,

$$\Phi^{J,\Gamma}_{i_{\rm vib},n} = \left\{ \Phi^{(J=0),\Gamma_{\rm vib}}_{i_{\rm vib}} \times |J, \Gamma_{\rm rot}, n\rangle \right\}^{\Gamma}, \qquad (25)$$

to solve the ro-vibrational eigen-problem for $J \geq 0$, in the so-called $J = 0$ representation. In Eq. (25), Γ, $\Gamma_{\rm vib}$ and $\Gamma_{\rm rot}$ are the total, vibrational and rotational symmetries in $\mathcal{T}_{\rm d}$(M), respectively, $i_{\rm vib}$ is a TROVE vibrational index to count the $\Phi^{J=0,\Gamma_{\rm vib}}_{i_{\rm vib}}$ functions regardless of their symmetry, n is a rotational counting number, see Section 4.3. For further details of the TROVE symmetry-adaptation and contraction procedure, see Yurchenko et al. (2017).

The ro-vibrational basis, and hence the corresponding Hamiltonian matrices, were further pruned by applying an energy threshold as follows

$$\tilde{E}^{\rm approx}_{v,J} = \tilde{E}^{(J=0),\Gamma_{\rm vib}}_{v} + \tilde{E}^{(v=0),\Gamma_{\rm rot}}_{n} < 25\,000 \text{ cm}^{-1},$$

where $\tilde{E}^{(J=0),\Gamma_{\rm vib}}_{v}$ is a $J = 0$ eigenvalue and

$$\tilde{E}^{(v=0),\Gamma_{\rm rot}}_{n} = BJ(J+1)$$

is a rigid rotor energy estimate. The sizes of the matrices $N_{F_1}(J)$ for $\Gamma = F_1$ as a function of J are illustrated in Fig. 3. The dimensions of the other matrices are approximately related to $N_{F_1}(J)$ as

$$N_{F_1}(J) \approx N_{F_2}(J) \approx 3\,N_{A_1}(J) \approx 3\,N_{A_2}(J) \approx \frac{3}{2}\,N_E(J).$$

In $\mathcal{T}_{\rm d}$(M), the ro-vibrational basis set consists of ten irreducible representations A_1, A_2, E_a, E_b, F_{1x}, F_{1y}, F_{1z}, F_{2x}, F_{2y}, F_{2z}, where only A_1, A_2, E_a, F_{1x} and F_{2x} need to be processed to provide unique solutions. For each J, five ro-vibrational Hamiltonian matrices were constructed and diagonalised using the eigensolver DSYEV from LAPACK (Anderson et al. 1999) as implemented in Intel's MKL libraries.

4.3 Rotational basis set and symmetrisation

The rotational basis $|J, \Gamma_{\rm rot}, n\rangle$ used for the spherical top CH_4 in Eq. (25) are not the standard Wang combinations. The latter cannot form symmetry adapted combinations that transform irreducibly according to $\mathcal{T}_{\rm d}$(M). This is because some of the equivalent rotations of the group cannot be related to the standard rotational operations associated with the molecular frame axes R_x, R_y or R_z. This is discussed in detail by Bunker & Jensen (1998a), Alvarez-Bajo et al. (2011) and Yurchenko (2023).

In this work, the irreducible representations Γ were defined using the standard irreducible technique, see, e.g. Bunker & Jensen (1998a), where the reducible representations are constructed by applying the equivalent rotations $R(\alpha, \beta, \gamma)$ of $\mathcal{T}_{\rm d}$(M) to the rigid rotor wavefunctions $|J, k, m\rangle$ as given by

$$R(\alpha, \beta, \gamma)|J, k, m\rangle = \sum_{k'=-J}^{J} D^{(J)*}_{k',k}(\alpha, \beta, \gamma)|J, k', m\rangle. \qquad (26)$$

Here, α, β and γ are the Euler angles, k and m are the projections of the rotational angular momentum on the molecular fixed z and laboratory fixed Z axes, respectively, and $D^{(J)*}_{k',k}(\alpha, \beta, \gamma)$ are the Wigner D functions. The angles of the equivalent rotation operations $R(\alpha, \beta, \gamma)$ of $\mathcal{T}_{\rm d}$(M) were taken from Alvarez-Bajo et al. (2011), see their Table 4.

The symmetry adapted rotational basis functions were then constructed as linear combinations of $|J, k, m\rangle$ involv-

ing a range of k values

$$|J, \Gamma_{\rm rot}, n\rangle = \sum_{k=-J}^{J} T_{n,k}^{(J,\Gamma_{\rm rot})} |J, k, m\rangle, \quad (27)$$

where n is simply a counting number for a given J and $\Gamma_{\rm rot}$ and the rotational quantum number k can no longer be used for classification as a rotational label and m is omitted in the left hand side. The procedure has been established by Yurchenko et al. (2017) and given in detail by Yurchenko (2023).

4.4 Vibrational band centre correction

The $J = 0$ contraction scheme allows one to introduce an additional, post-fitting systematic correction to the vibrational energies ('band centres') following the empirical basis set correction (EBSC) approach from Yurchenko et al. (2009). This is done by replacing the diagonal matrix elements of the vibrational Hamiltonian operator $\hat{H}_{\rm vib}$ with the empirically-derived values for the vibrational energies $\tilde{E}_v^{(J=0),\Gamma_{\rm vib}}$, where available. Indeed, $\hat{H}_{\rm vib}$ is diagonal on the $J = 0$ basis set and such a simple empirical shift is propagated to all the ro-vibrational Hamiltonian matrices with $J > 0$ and hence to the corresponding ro-vibrational energy levels obtained.

Empirical corrections were applied to 198 $J = 0$ energy levels as the best averaged estimate relative to the MARVEL energies below $J = 8$. The typical shifts were within 0.1 cm^{-1}, with the largest being 0.78 cm^{-1} and the rms difference being 0.48 cm^{-1}. The full set of calculated $J = 0$ energies with the empirical corrections is provided as supplementary material.

5 THE MM LINE LIST

The ro-vibrational Einstein A coefficients of CH_4 were computed with TROVE using the ro-vibrational eigen-functions and *ab initio* electric DMS for all transitions between 0 and 12 000 cm^{-1}, covering states up to $J = 60$, lower state energies below $hc\cdot$ 12 000 cm^{-1}, and upper states energies below $hc\cdot$ 18 000 cm^{-1}. The Einstein A coefficients and energy levels form the new line list for CH_4, called MM (MARVELous Methane), comprising of 50 395 644 806 lines and covering 9 155 208 ro-vibrational states. The ExoMol line list format assumes a States file (`.states`), containing the state ID, energy term values (cm^{-1}), total state degeneracies, state uncertainties, lifetimes, quantum numbers and labels etc, and Transition files (`.trans`), containing the upper/lower state IDs and Einstein coefficients (s^{-1}). The MM transitions are sorted and divided into 1200 files in 100 cm^{-1} windows. Extracts from the States and Transition files are given in Tables 3 and 4, respectively.

5.1 Quantum numbers and labels

In order to help future users of our data, a variety of methane quantum numbers (QN) and labels are provided as part of the MM States file, see Table 3. Most importantly, the ro-vibrational states of CH_4 are classified by the rigorous quantum numbers: the total angular momentum J and the total symmetry $\Gamma_{\rm tot}$, either A_1, A_2, E, F_1, or F_2 in the molecular symmetry group $\mathcal{T}_{\rm d}$(M) (Bunker & Jensen 2004). As non-rigorous quantum numbers, we provide 'experimental' (MARVEL), 'normal mode' and 'local mode' (TROVE) sets of quantum numbers as follows.

Table 1. Symmetries of the normal modes of methane in $\mathcal{T}_{\rm d}$(M) and their fundamental wavenumbers.

Mode	Symmetry	$\tilde{\nu}$ cm^{-1}	Type
1	A_1	2916.45	symmetric stretch
2	E	1533.33	asymmetric bend
3	F_2	3019.49	asymmetric stretch
4	F_2	1310.76	asymmetric bend

The local mode set of quantum numbers is the basis for other QN schemes. It consists of the TROVE quantum numbers v_1, v_2, ..., v_9 representing the excitation indices of the 1D primitive basis set functions $\phi_{v_i}(\xi_i)$ and the vibrational symmetry $\Gamma_{\rm vib}$. The QN in TROVE are assigned using the largest basis set contribution approach, where quantum labels of a basis function with the largest magnitude of the eigencoefficient are assumed as the QN of the eigenstate. The primitive quantum numbers v_1, v_2, ..., v_9 are therefore propagated through our multi-step contraction procedure described above, by applying the largest contribution assignment at each contraction step. Having a representation of all 9 quantum modes in the state assignment is extremely useful, however, at high vibrational excitation it often leads to ambiguous quantum numbers, repeating or missing QN descriptions. An additional vibrational QN in TROVE is the counting index $i_{\rm vib}$ in $\Phi_{i_{\rm vib}}^{J=0,\Gamma_{\rm vib}}$ as in Eq. (25); $i_{\rm vib}$ is a unique identifier of the vibrational eigenfunctions, which are also $J = 0$ basis functions, and therefore can be very useful (see below).

The normal mode set consists of the following 15 labels:

$$\Gamma_{\rm tot}, J, n, \Gamma_{\rm rot}, n_1, n_2, L_2, n_3, L_3, M_3, n_4, L_4, M_4, \Gamma_{\rm vib}, N_{J,\Gamma_{\rm tot}},$$

where $\Gamma_{\rm rot}$ and $\Gamma_{\rm vib}$ are the symmetries of the rotational and vibrational wavefunction, respectively, n is the rotational quantum number as in Eq. (27), and $n_1, n_2, L_2, n_3, L_3, M_3, n_4, L_4, M_4$ are the normal mode vibrational quantum numbers (see the mode designation in Table 1). The normal mode QN scheme is based on the properties of the harmonic oscillator functions, 1D for ν_1, isotropic 2D for ν_2 and isotropic 3D for ν_3 and ν_4. We follow Yurchenko & Tennyson (2014) and use the absolute values of the vibrational angular momentum quantum numbers $L_i = n_i, n_i - 2, \ldots, 0(1)$. $M_i \leq L_i$ is a multiplicity index used to count states within a given n_i, L_i set, see Boudon et al. (2006). The advantage of this scheme is that the angular momentum indices L_i (for the E-type ν_2 mode) and M_i (for the F-type ν_3 and ν_4 modes) provide a one-to-one correspondence to the symmetry of the corresponding component as listed in Table 2. In fact, such a correspondence is common for the 2D E-type symmetry isotropic harmonic oscillators, see e.g. Bunker & Jensen (1998b); Yurchenko et al. (2005), while the 3D isotropic harmonic oscillator quantum numbers (Hougen 2001) are less common.

The normal mode assignment is in fact reconstructed using the local mode quantum numbers. The easiest way to correlate them is at step 2 of the TROVE calculations, when the three reduced eigenvalue problems in Eqs. (21–23) are solved.

Table 2. Symmetry classification of the 2D and 3D isotropic harmonic oscillators for the corresponding vibrational angular momentum index L_2 and the multiplicity index M_i ($i = 3, 4$, $0 \leq M_i \leq L_i$), respectively. n is an arbitrary integer, $n = 1, 2, 3, \ldots$.

L_2	Symmetry 2D	M_i	Symmetry 3D
0	A_1	0	A_1
$3\,n$	A_1, A_2	$12n + 0$	A_1
$3\,n \pm 1$	E	$12n + 6$	A_2
		$6n \pm 2$	E
		$4n + 3$	F_1
		$4n + 1$	F_2

Indeed, the local mode eigensolutions of the 4D Hamiltonian $\hat{H}^{(1)}$ in Eq. (21) need to match the combinations of stretching 1D and 3D harmonic oscillators only; the eigenvalues of the 2D bending $\hat{H}^{(2)}$ are easily matched to energies of a 2D isotropic harmonic oscillator; and the eigenvalues of the 3D bending reduced problem for $\hat{H}^{(3)}$ are then correlated solutions of a 3D harmonic oscillator. Once the correlation is established, it is propagated through the multi-step assignment procedure along with the local mode QN. Moreover, the unique vibrational state index $i_{\rm vib}$ can also be used to efficiently re-assign the TROVE local mode QN to any other QN scheme, which is possible even in the final data set.

The CH_4 experimental quantum number set usually consists of J, $\Gamma_{\rm tot}$, polyad number P in Eq. (19) and a counting number n, counting different states within the same J, $\Gamma_{\rm tot}$ and P. It is an extremely simple, and therefore powerful, set for correlating methane spectroscopic data from different sources. It has been adopted by the community as the methane spectroscopic standard and it was used to classify the experimental data in the recent MARVEL work on CH_4 (Kefala et al. 2024). The disadvantage of the experimental QN scheme is that it loses any connection with the physics, such as, e.g., vibrational state, which can is required, e.g., when it is important to know if a given state has a bending or stretching nature, whether it belongs to a forbidden or combination band, and so on.

An important aspect of the ExoMol States file is the state degeneracies $g_i = g_{\rm ns}(2J + 1)$. For $^{12}CH_4$, considering the H nuclear spin, the nuclear statistical weight, $g_{\rm ns}$, takes the symmetry-dependent values 5, 5, 2, 3, and 3 for A_1, A_2, E, F_1, F_2 respectively.

The electric dipole transitions obey the following symmetry-determined selection rules:

$$A_1 \leftrightarrow A_2\ , E \leftrightarrow E\ , F_1 \leftrightarrow F_2 \tag{28}$$

with the standard rotational angular momentum, **J**, selection rules:

$$J \leftrightarrow J \pm 1\ \ , J' + J'' \neq 0. \tag{29}$$

5.2 Partition function

A partition function of CH_4 computed using the MM line list is provided as part of the line list package, covering the temperature range from 0 to 2500 K on a grid of 1 K. In Fig. 4, the MM partition function is compared to the HITRAN partition function generated using the TIPS 2021 package (Gamache et al. 2021). The MM partition function closely agrees with TIPS until about $T = 2000$ K, after which it slowly deviates, underestimating the more complete TIPS partition function. Hence, $T = 2000$ K is our estimated upper temperature limit for completeness of the MMline list.

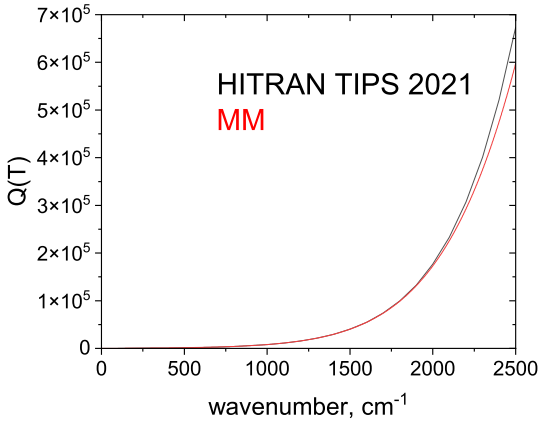

Figure 4. Partition functions of CH_4: HITRAN TIPS 2021 (Gamache et al. 2021) and MM.

5.3 Opacities

Temperature- and pressure-dependent opacities of CH_4 have been generated using the ExoMolOP procedure (Chubb et al. 2021) for the MM line list. These are provided for four exoplanet atmospheric retrieval codes ARCiS (Min et al. 2020), TauREx (Al-Refaie et al. 2021), NEMESIS (Irwin et al. 2008) and petitRADTRANS (Molliére et al. 2019). For the line broadening, we assumed a 85% H_2 and 15% He atmosphere and Voigt line profile with the following parameters: $\gamma_{\rm H_2} = 0.12$ cm^{-1}, $n_{\rm H_2} = 0.5$, $\gamma_{\rm He} = 0.05$ cm^{-1} and $n_{\rm He} = 0.5$.

6 SPECTROSCOPIC ILLUSTRATIONS

We use the MM line list to simulate spectra of CH_4 to demonstrate the accuracy of the line positions, the quality of the line intensities, as well as its completeness especially at high temperatures. To this end, we compare simulated spectra either with experimental or theoretical spectra from the literature or other databases.

We start with a comparison against the HITRAN spectrum of $^{12}CH_4$ at $T = 296$ K. Figures 5 and 6 provides detailed comparisons of MM and HITRAN 'stick' spectra at $T = 296$ K, where absorption coefficients (cm/molecule) are plotted as sticks as functions of the wavenumber covering the region between 2700 and 12000 cm^{-1}.There is very good agreement between MM and HITRAN below 6200 cm^{-1}, where all the lines shown are MARVELised, also nicely illustrating the completeness of the CH_4 MARVEL energies in this region. The MM spectrum below 2700 cm^{-1} agrees with HITRAN as well or better. The incompleteness of the CH_4 MARVEL dataset, and therefore of the assigned experimental data available in the literature, is evident above 8200 cm^{-1}. The agreement of the MM line positions with HITRAN also deteriorates for polyads six and seven, which

Table 3. Extract from the `.states` file of the CH_4 MM line list.

i	$\tilde{E}$ (cm^{-1})	g_i	J	unc	$\tilde{\tau}$	$\Gamma_{\rm tot}$	P	n	n_1	n_2	l_2	n_3	l_3	m_3	n_4	l_4	m_4	$\Gamma_{\rm v}$	n	$\Gamma_{\rm r}$	$i_{\rm v}$	C_i	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	Ma/Ca	$\tilde{E}_{\rm Ca}$ (cm^{-1})
11527	3924.922540	15	1	0.002010	2.0200E-03	A1	3	1	0	0	0	0	0	0	3	3	3	F1	1	F1	15	-1.00	0	0	0	0	0	0	0	0	3	Ma	3924.949931
11528	4135.704754	15	1	0.002010	2.0200E-03	A1	3	2	0	1	1	0	0	0	2	2	1	F1	1	F1	18	-1.00	0	0	0	0	0	1	0	0	2	Ma	4135.816306
11529	4327.267355	15	1	0.701000	7.0100E-01	A1	3	3	0	0	0	1	1	1	1	1	1	F1	1	F1	27	1.00	0	1	0	0	0	0	0	0	1	Ca	4327.267355
11530	4379.046748	15	1	1.101000	1.1010E+00	A1	3	4	0	2	2	0	0	0	1	1	1	F1	1	F1	29	-1.00	0	0	0	0	0	2	0	0	1	Ca	4379.046748
11531	4548.890837	15	1	0.002010	2.0200E-03	A1	3	5	0	1	1	1	1	1	0	0	0	F1	1	F1	32	-1.00	0	1	0	0	0	1	0	0	0	Ma	4548.876003
11532	5237.031268	15	1	0.005000	5.0080E-03	A1	4	1	0	0	0	0	0	0	4	4	3	F1	1	F1	42	-1.00	0	0	0	0	0	0	0	0	4	Ma	5236.993580
11533	5405.815045	15	1	0.005000	5.0080E-03	A1	4	2	0	1	1	0	0	0	3	1	1	F1	1	F1	45	0.99	0	0	0	0	0	1	0	0	3	Ma	5405.790618
11534	5445.903757	15	1	0.007500	7.5050E-03	A1	4	3	0	1	1	0	0	0	3	3	3	F1	1	F1	48	0.73	0	0	0	0	0	1	0	0	3	Ma	5445.996059
11535	5470.022619	15	1	1.101000	1.1010E+00	A1	4	4	0	1	1	0	0	0	3	3	1	F1	1	F1	50	-0.74	0	0	0	0	0	1	0	0	3	Ca	5470.022619
11536	5622.856290	15	1	0.009000	9.0040E-03	A1	4	5	0	0	0	1	1	1	2	2	1	F1	1	F1	59	-0.93	0	1	0	0	0	0	0	0	2	Ma	5622.906629
11537	5643.812311	15	1	0.009000	9.0040E-03	A1	4	6	0	0	0	1	1	1	2	2	2	F1	1	F1	61	-0.93	0	1	0	0	0	0	0	0	2	Ma	5643.847392
11538	5660.316674	15	1	1.301000	1.3010E+00	A1	4	7	0	2	2	0	0	0	2	2	1	F1	1	F1	65	1.00	0	0	0	0	0	2	0	0	2	Ca	5660.316674
11539	5760.164670	15	1	0.007570	7.6320E-03	A1	4	8	1	1	1	0	0	0	1	1	1	F1	1	F1	71	-1.00	0	1	0	0	0	1	0	0	1	Ma	5760.218373
11540	5838.639168	15	1	0.007570	7.6320E-03	A1	4	9	0	1	1	1	1	1	1	1	1	F1	1	F1	74	0.98	0	1	0	0	0	1	0	0	1	Ma	5838.686695
11541	5856.277154	15	1	0.007570	7.6320E-03	A1	4	10	0	1	1	1	1	1	1	1	1	F1	1	F1	80	-0.98	0	1	0	0	0	1	0	0	1	Ma	5856.350922
11542	5881.226199	15	1	1.501000	1.5010E+00	A1	4	11	0	3	3	0	0	0	1	1	1	F1	1	F1	83	0.98	0	0	0	0	0	3	0	0	1	Ca	5881.226199
11543	5922.076567	15	1	1.101000	1.1010E+00	A1	4	12	0	3	1	0	0	0	1	1	1	F1	1	F1	85	0.98	0	0	0	0	0	3	0	0	1	Ca	5922.076567
11544	6071.512360	15	1	1.201000	1.2010E+00	A1	4	13	0	2	2	1	1	1	0	0	0	F1	1	F1	92	-1.00	0	1	0	0	0	2	0	0	0	Ca	6071.512360

i:	State counting number.
$\tilde{E}$:	State energy (in cm^{-1}).
g_i:	Total degeneracy, equal to $g_{\rm ns}^i(2J+1)$.
J:	Total angular momentum.
unc:	Uncertainty (in cm^{-1}).
τ:	lifetime (s^{-1}).
$\Gamma_{\rm tot}$:	Total symmetry in $\mathcal{T}_{\rm d}$(M).
P:	Polyad number.
n:	Polyad counting number (see text)
$n_1 - n_4$:	Normal mode vibrational quantum numbers.
l_2, l_3, l_4:	Vibrational angular momenta quantum numbers.
m_3, m_4:	Multiplicity index quantum numbers.
$\Gamma_{\rm v}$:	Symmetry of the vibrational contribution in $\mathcal{T}_{\rm d}$(M).
n:	Rotational counting number, see Eq. (25).
$\Gamma_{\rm r}$:	Symmetry of the rotational contribution in $\mathcal{T}_{\rm d}$(M).
$i_{\rm v}$:	vibrational state ID.
Ma/Ca:	'Ma' is for MARVEL and 'Ca' is for Calculated .
C_i:	Largest coefficient used in the assignment.
$v_1 - v_9$	Local mode vibrational quantum numbers.
$\tilde{E}_{\rm Ca}$:	State energy term values in cm^{-1}, Calculated (TROVE).

again reflects the limited experimental data used in the refinement of our PES. Our intensities, which are based on *ab initio* DMSs, agree reasonably well with the HITRAN data even at very high wavenumbers, which is illustrated in Fig. 7, where we compare the HITRAN and MM line intensities as a ratio $I_{\rm HITRAN}/I_{\rm calc}$ averaged over lines within 58 vibrational bands for 204127 (out of 309683) lines we could establish the correlation with. The averaging included all sub-bands corresponding to different symmetries and vibrational angular moment quantum numbers l_2, l_3, l_4 of each vibrational state $(\nu_1, \nu_2, \nu_3, \nu_4)$. The majority of points are safely within 5%.

Figure 8 shows the coverage of the MARVELisation of MM in more detail. Using 23 292 MARVELised states only, the $T = 296$ K MM spectrum of CH_4 results in 1 371 651 experimental-quality lines. This is much larger than the 82173 experimentally assigned transitions used to produce the MARVEL energy set. The HITRAN 2020 line list for $^{12}CH_4$ contains 309 864 lines. Figure 8 is a great illustration of the MARVEL methodology that leads to over one million transitions of experimental quality. It should be noted, however, that even with this extension, the experimental dataset is far from complete, even at room temperature. As is also seen in Fig. 8, the spectroscopic range above 7000 cm^{-1} is poorly represented by experiment, especially by assigned experiment. Although HITRAN provides good coverage up to 9000 cm^{-1} and some coverage up to 11 500 cm^{-1}, the vast majority of these lines are only provided with approximate lower state energies needed for the temperature variation of the HITRAN intensities. We hope that our line list will help assign experimental data on CH_4 and in turn lead to improved theoretical models.

Currently, the most comprehensive (accurate and complete) line lists for CH_4 are from TheoReTS (Rey et al. 2017), MeCaSDa@VAMDC (Ba et al. 2013) and HITEMP (Hargreaves et al. 2020). The latter is a compilation of the $T = 296$ K HITRAN line list for $^{12}CH_4$ and the TheoReTS line list list by Rey et al. (2017). Figure 9 shows simulations of absorption spectra of CH_4 at $T = 1000$ K, $T = 1500$ K and $T = 2000$ K at low resolution (1 cm^{-1}) generated using the HITEMP, TheoReTS and MM line lists. The TheoReTS spectra, based on the TheoReTS line list for CH_4 by Rey et al. (2017), were computed by combining two contributions, (i) strong lines and (ii) temperature dependent continuum (Hargreaves et al. 2015). While there is generally good agreement

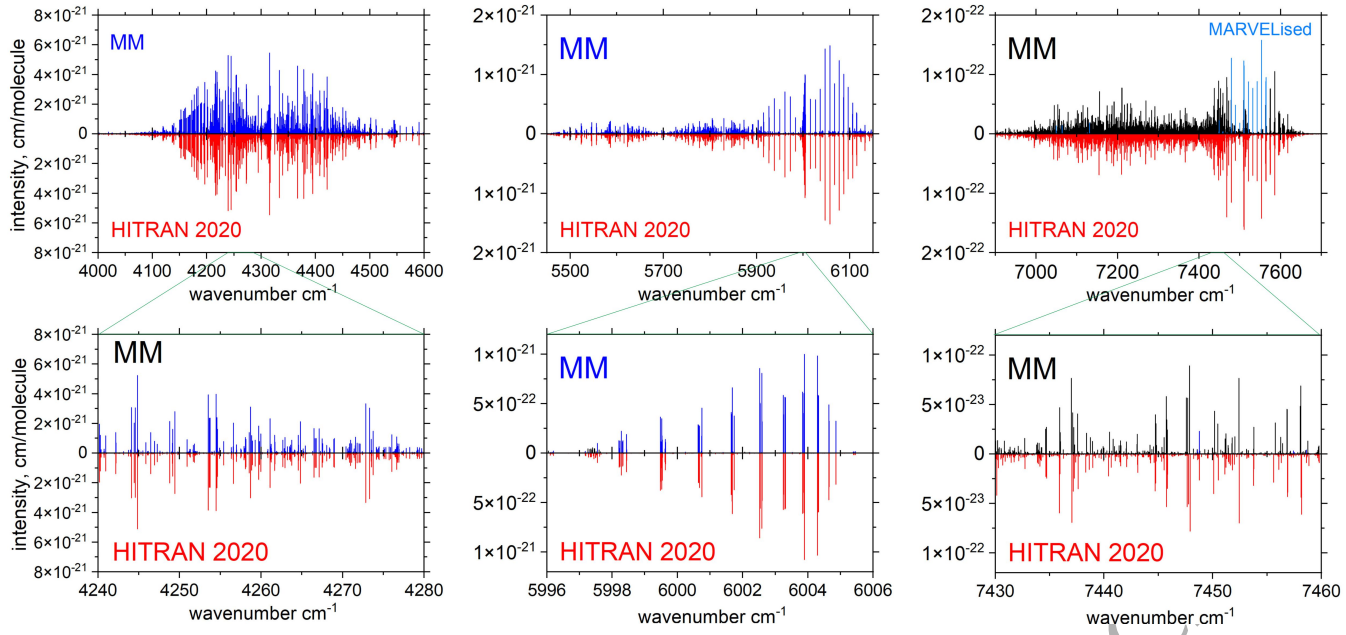

Figure 5. Absorption stick spectra of CH_4 calculated using the MM line list at $T = 296$ K and compared to HITRAN (polyads 3,4,5). All blue sticks (light or dark) correspond to MARVELised transitions, while the black sticks correspond to the calculated MM transitions. Insets zoom in in some selected regions.

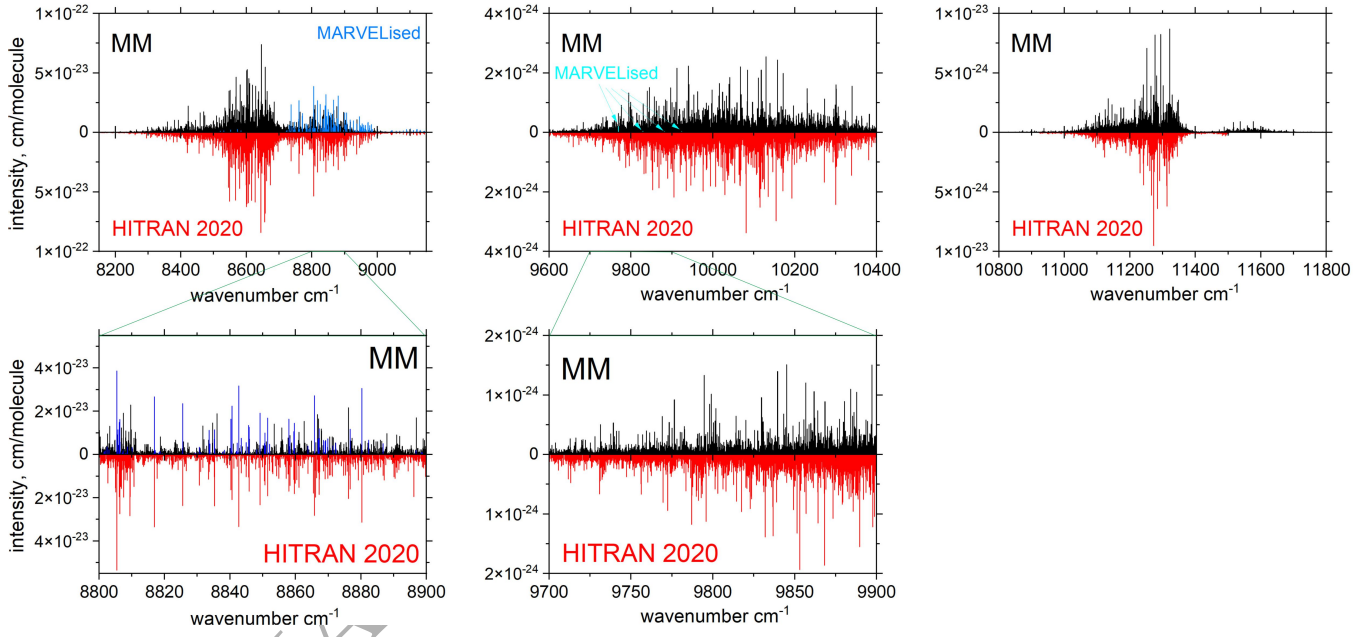

Figure 6. Absorption stick spectra of CH_4 calculated using the MM line list at $T = 296$ K and compared to HITRAN (polyads 6,7,8). All blue sticks (light or dark) correspond to MARVELised transitions, while the black sticks correspond to the calculated MM transitions. Insets zoom in in some selected regions.

between the three spectra at $T = 1000$ K and $T = 1500$ K, where MM is reasonably close both to HITEMP and TheoReTS, at the higher temperature of $T = 2000$ K all three line lists show significant deviation from each other especially at higher frequencies as well as between the bands. These discrepancies can be most likely attributed to the level of completeness of the line lists, with the TheoReTS based on the largest and HITEMP on the smallest number of transitions in these regions.

Another important benchmark is provided by Wong et al. (2019), who reported experimental cross sections of $^{12}CH_4$ at a set of temperatures from 295 K to 1000 K in the range

Table 4. Extract from the `.trans` file of the CH_4 MM line list.

f	i	A_{fi}
2818679	2713208	4.2438E-05
1576056	1766199	2.1991E-04
206172	313386	4.9898E-04
975346	1195595	1.5380E-03
5940830	6346754	1.0693E-04
69748	36029	4.1670E-04
774558	972721	1.3866E-05
3552129	3812515	9.2443E-04
1768478	1572686	1.4323E-04
3442284	3179845	9.3747E-06
288440	195011	3.1192E-07
1968365	1650415	2.3694E-04
1576697	1244901	2.5908E-03

f: Upper state counting number;
i: Lower state counting number;
A_{fi}: Einstein-A coefficient (in s^{-1}).

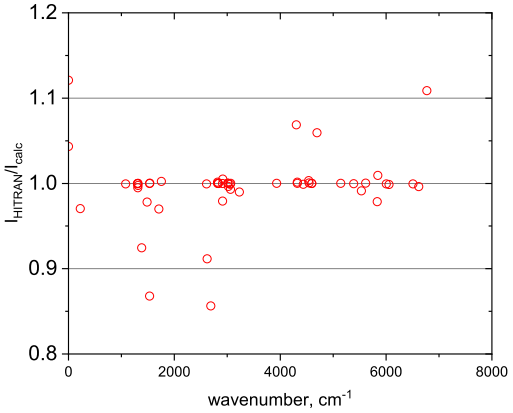

Figure 7. Average intensity ratio $I_{\rm HITRAN}/I_{\rm calc}$ between the HITRAN and the MM line intensities with 58 bands.

5500 cm^{-1} to 8900 cm^{-1}. In their paper they demonstrated excellent agreement of the TheoReTS line list with the experimental cross sections across the entire range of wavenumbers and temperatures. In Fig. 10, we compare the MM simulations at $T = 1000$ K on a grid of 0.0075 cm^{-1} to the CH_4 cross sections at the resolution of 2 cm^{-1} by Wong et al. (2019). The agreement with the experiment for this hot spectrum is comparable with the quality of TheoReTS demonstrated in Wong et al. (2019). It is interesting that the experimental cross sections in 8200–8900 cm^{-1} region are systematically higher than both TheoReTS and MM. We currently do not have an explanation for this discrepancy.

7 CONCLUSION

An extensive ro-vibrational line list for the main isotopologue of methane MM is presented. The line list is meant to replace the ExoMol 10to10 and 30to10 CH_4 line lists from 2014; it is both more complete and more accurate. It contains over 50 billion transitions and covers the frequency range from 0 to

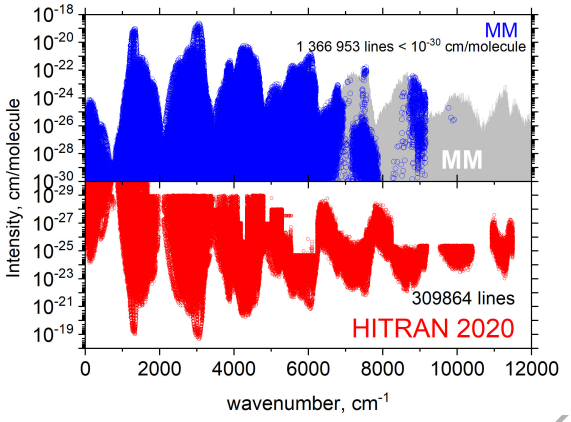

Figure 8. Comparison of the HITRAN (bottom) and MM (top) $T = 296$ K spectra. The blue points on the top display show line intensities (cm/molecule) computed for the MARVELised states only, while the grey area indicates the full coverage in the MM line list.

12 000 cm^{-1} with rotational excitation up to $J = 60$. With the lower energy cutoff of 10000 cm^{-1}, it should be complete up to 1500 K, with lower wavenumbers (∼7000 cm^{-1}), complete up to 2000 K. In fact, we show that at high temperatures (> 1000 K) it is significantly more complete than the HITEMP line list for CH_4. The line list is based on a new empirical potential energy surface, which was refined by fitting to the experimentally derived energies of CH_4, and on a new kinetic energy operator, represented by a Taylor expansion in terms of nine curvilinear coordinates. For intensities, a new high level CCSD(T) dipole moment of CH_4 was used. The accuracy of the line list is further improved via the MARVELisation procedure, where the computed energies are replaced by the experimentally derived values (from Kefala et al. (2024)). This is the ExoMol strategy to replace energies, not line positions, which provides a more consistent as well as more global solution. Indeed, using the MARVELised energy levels only, i.e. states levels characterised by experimental quality, we were able to generate near 1 000 000 new lines of that quality.

A room temperature spectrum computed with MM compares well with HITRAN CH_4 data, both in terms of the line positions and intensities, including the very many unassigned transitions in HITRAN. Indeed, Kefala et al. (2024) list about 50 high resolution measurements of CH_4 spectra which are either complete unassigned or only partially assigned. Our line list can undoubtedly help provide quantum number assignments for a significant number of these works as well providing updates for HITRAN. Our improvement is not the last word in the line list provision for methane. Analysis of the high wavenumber CH_4 spectra, especially above > 7000 cm^{-1}, will lead to more experimentally derived energies and facilitate further improvement of the PES of CH_4. We hope the turn-around time between the experiment, analysis and theory can now be made more efficient and quicker than 10 years ago.

In order to facilitate the usage in the exoplanetary applications, as part of the MM ExoMol line list package, we also provide a set of atmospheric opacities for four mainstream codes, ARCiS, TauREx, NEMESIS and petitRADTRANS.

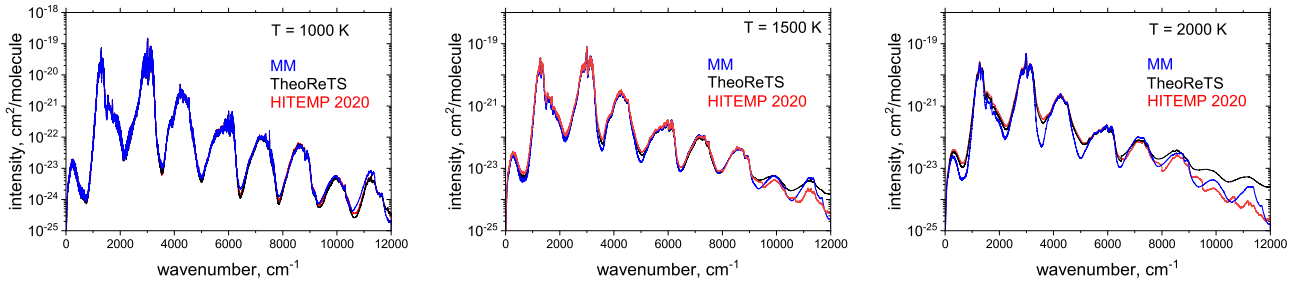

Figure 9. Absorption spectra of CH_4 simulated using the MM, TheoReTS and HITEMP line lists at $T = 1000$ K, $T = 1500$ K and $T = 2000$ K on a grid of 1 cm^{-1}. A Gaussian line profile with a half-width-at-half maximum (HWHM) of 1 cm^{-1} was used.

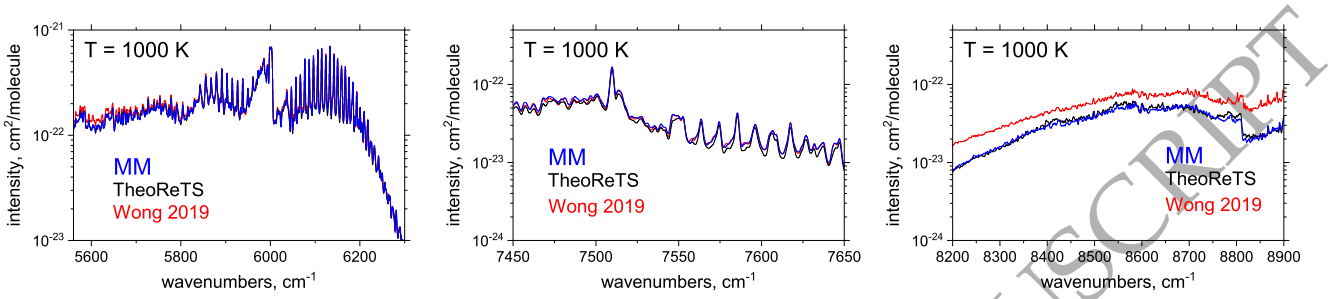

Figure 10. Comparison of the theoretical high temperature ($T = 1000$ K) MM cross sections of CH_4 with the experimental cross sections by Wong et al. (2019). Here we use the 2 cm^{-1} resolution data set of Wong et al. (2019), while the theoretical cross sections were generated using a Lorentzian line profile with a HWHM of 1 cm^{-1}.

A feature of the MARVEL procedure is its active nature which means that once new sets of assigned transitions become available these can be added to the existing MARVEL dataset and the MARVEL procedure can be re-run. The ExoMol data structure means that line lists can then be updated in a straightforeward fashion with out the need for the expensive re-calculation transiton probabilities (Einsten-A coefficients); updates like this have already been made for number of key molecules, for example AlO (Bowesman et al. 2021), H_3^+ and its isoptopologues (Bowesman et al. 2023) and twice for formaldehyde (H_2CO) (Abdoulanziz et al. 2021; Germann et al. 2024). The current MM line list is highly accurate in far- and mid-infrared region but loses accuracy at near-infrared wavelengths. Not due to lack of measured methane spectra in this region but the lack of assignments of these spectra. As we and others address this issue and assign these spectra, the results will be used to systematically improve the MM line list.

ACKNOWLEDGEMENTS

We thank Michael Rey for providing access to the extended TheoReTS CH_4 line lists. This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme through Advance Grant number 883830 and the STFC Projects No. ST/M001334/1 and ST/R000476/1. The authors acknowledge the use of the Cambridge Service for Data Driven Discovery (CSD3) as part of of the STFC DiRAC HPC Facility (www.dirac.ac.uk), funded by BEIS capital funding via STFC capital grants ST/P002307/1 and ST/R002452/1 and STFC operations grant ST/R00689X/1. AO acknowledges additional support from the Engineering and Physical Sciences Research Council [grant number EP/S021582/1].

DATA AVAILABILITY

During the reviewing process, the line list will be available at http://exomol.com/repository/.

The states, transition, opacity and partition function files for the CH_4 line lists can be downloaded from www.exomol.com. The open access programs TROVE and ExoCross are available from github.com/exomol.

SUPPORTING INFORMATION

`CH4_band_centres_corrections` is a list of TROVE $J = 0$ energy term values and corresponding empirical energy shifts (EBSC).

`SM_ch4_MM_pes_dms.zip` contains subroutines and input files for the refined PES and *ab initio* DMS of CH_4.

`12C-1H4__MM__model.inp` is a TROVE input file containing the potential parameters as well as the TROVE configuration parameters.

Supplementary data are available at MNRAS online. This includes the spectroscopic model in the form of the TROVE input file (), containing all parameters as well as the experimentally derived energy term values of CH_4 used in the fit.

REFERENCES

Abdoulanziz A., Argentin C., Laporta V., Chakrabarti K., Bultel A., Tennyson J., Schneider I. F., Mezei J. Z., 2021, J. Appl. Phys., 129, 053303

Al-Refaie A. F., Changeat Q., Waldmann I. P., Tinetti G., 2021, ApJ, 917, 37

Albert S., Bauerecker S., Boudon V., Brown L. R., Champion J.-P., Loëte M., Nikitin A., Quack M., 2009, Chem. Phys., 356, 131

Alvarez-Bajo O., Lemus R., Carvajal M., Perez-Bernal F., 2011, Mol. Phys., 109, 797

Anderson E., et al., 1999, LAPACK Users' Guide, third edn. Society for Industrial and Applied Mathematics, Philadelphia, PA

Ba Y. A., et al., 2013, J. Quant. Spectrosc. Radiat. Transf., 130, 62

Beaulieu J. P., et al., 2011, ApJ, 731, 16

Benneke B., et al., 2019a, Nature Astronomy, 3, 813

Benneke B., et al., 2019b, ApJL, 887, L14

Birkby J. L., 2018, Handbook of Exoplanets, pp 1485–1508

Blain D., Charnay B., Bézard B., 2021, A&A, 646, A15

Bochanski J. J., Burgasser A. J., Simcoe R. A., West A. A., 2011, AJ., 142, 169

Borysow A., Champion J. P., Jorgensen U. G., Wenger C., 2003, in Hubeny I., Mihalas D., Werner K., eds, ASP Conference Series Vol. 288, Stellar atmosphere modeling. pp 352–356

Boudon V., Rey M., Loëte M., 2006, J. Quant. Spectrosc. Radiat. Transf., 98, 394

Bowesman C. A., Shuai M., Yurchenko S. N., Tennyson J., 2021, MNRAS, 508, 3181

Bowesman C. A., et al., 2023, MNRAS, 519, 6333

Brogi M., Birkby J., 2021, in 2514-3433, ExoFrontiers. IOP Publishing, pp 8–1 to 8–10, doi:10.1088/2514-3433/abfa8fch8

Brogi M., Line M. R., 2019, ApJ, 157, 114

Brogi M., Snellen I. A. G., de Kok R. J., Albrecht S., Birkby J., de Mooij E. J. W., 2012, Nature, 486, 502

Bunker P. R., Jensen P., 1998a, Molecular Symmetry and Spectroscopy, 2 edn. NRC Research Press, Ottawa

Bunker P. R., Jensen P., 1998b, Molecular Symmetry and Spectroscopy, 2 edn. NRC Research Press, Ottawa

Bunker P. R., Jensen P., 2004, Fundamentals of Molecular Symmetry. IOP Publishing, Bristol

Burrows A., Sharp C. M., 1999, ApJ, 512, 843

Canty J. I., et al., 2015, MNRAS, 450, 454

Chubb K. L., et al., 2021, A&A, 646, A21

Cooley J. W., 1961, Math. Comp., 15, 363

Császár A. G., Czakó G., Furtenbacher T., Mátyus E., 2007, Annu. Rep. Comput. Chem., 3, 155

Furtenbacher T., Császár A. G., 2012, J. Mol. Struct., 1009, 123

Furtenbacher T., Császár A. G., Tennyson J., 2007, J. Mol. Spectrosc., 245, 115

Gamache R. R., Vispoel B., Rey M., Nikitin A., Tyuterev V., Egorov O., Gordon I. E., Boudon V., 2021, J. Quant. Spectrosc. Radiat. Transf., 271, 107713

Gandhi S., et al., 2020, MNRAS, 495, 224

Germann M., et al., 2024, J. Quant. Spectrosc. Radiat. Transf., 312, 108782

Giacobbe P., et al., 2021, Nature, 592, 205

Halonen L., 1997, J. Chem. Phys., 106, 831

Hargreaves R. J., Beale C. A., Michaux L., Irfan M., Bernath P. F., 2012, ApJ, 757, 46

Hargreaves R. J., Bernath P. F., Bailey J., Dulick M., 2015, ApJ, 813, 12

Hargreaves R. J., Gordon I. E., Rey M., Nikitin A. V., Tyuterev V. G., Kochanov R. V., Rothman L. S., 2020, ApJS, 247, 55

Hood C. E., Fortney J. J., Line M. R., Faherty J. K., 2023, ApJ, 953, 170

Hougen J., 2001, Methane Symmetry Operations . NIST, Gaithersburg, MD.

Irwin P. G. J., et al., 2008, J. Quant. Spectrosc. Radiat. Transf., 109, 1136

Kefala K., Boudon V., Yurchenko S. N., Tennyson J., 2024, J. Quant. Spectrosc. Radiat. Transf.

Knutson H. A., Benneke B., Deming D., Homeier D., 2014, Nature, 505, 66

Lacy B., Burrows A., 2023, ApJ, 950, 8

Lodders K., Fegley B., 2002, Icarus, 155, 393

Madhusudhan N., Seager S., 2011, ApJ, 729, 41

Madhusudhan N., Sarkar S., Constantinou S., Holmberg M., Piette A. A. A., Moses J. I., 2023, Carbon-bearing Molecules in a Possible Hycean Atmosphere (arXiv:2309.05566)

Makhnev V. Y., Kyuberis A. A., Zobov N. F., Lodi L., Tennyson J., Polyansky O. L., 2018, J. Phys. Chem. A, 122, 1326

Mant B. P., Yachmenev A., Tennyson J., Yurchenko S. N., 2018, MNRAS, 478, 3220

Min M., Ormel C. W., Chubb K., Helling C., Kawashima Y., 2020, A&A, 642, A28

Molliére P., Wardenier J. P., van Boekel R., Henning T., Molaverdikhani K., Snellen I. A. G., 2019, A&A, 627, A67

Nikitin A. V., et al., 2014, J. Quant. Spectrosc. Radiat. Transf., 154, 63

Nikitin A. V., et al., 2020, J. Quant. Spectrosc. Radiat. Transf., 253, 107061

Noumerov B. V., 1924, MNRAS, 84, 592

Owens A., Yurchenko S. N., Yachmenev A., Tennyson J., Thiel W., 2016, J. Chem. Phys., 145, 104305

Owens A., Yachmenev A., Thiel W., Tennyson J., Yurchenko S. N., 2017, MNRAS, 471, 5025

Rey M., Nikitin A. V., Tyuterev V. G., 2014, ApJ, 789, 2

Rey M., Nikitin A. V., Babikov Y. L., Tyuterev V. G., 2016, J. Mol. Spectrosc., 327, 138

Rey M., Nikitin A. V., Tyuterev V. G., 2017, ApJ, 847, 105

Rodina A. A., et al., 2019, J. Quant. Spectrosc. Radiat. Transf., 225, 351

Snellen I., 2014, Phil. Trans. Royal Soc. London A, 372, 20130075

Stevenson K. B., et al., 2010, Nature, 464, 1161

Swain M. R., Vasisht G., Tinetti G., 2008, Nature, 452, 329

Tennyson J., Yurchenko S. N., 2017, Int. J. Quantum Chem., 117, 92

Tennyson J., Yurchenko S. N., 2021, Astronomy & Geophysics, 62, 6.16

Tennyson J., Furtenbacher T., Yurchenko S. N., Császár A. G., 2023, J. Quant. Spectrosc. Radiat. Transf.

Tóbiás R., Furtenbacher T., Tennyson J., Császár A. G., 2019, Phys. Chem. Chem. Phys., 21, 3473

Werner H. J., et al., 2015, MOLPRO, version 2015.1, a package of ab initio programs, http://www.molpro.net

Wong A., Bernath P. F., Rey M., Nikitin A. V., Tyuterev V. G., 2019, ApJS, 240, 4

Yachmenev A., Yurchenko S. N., 2015, J. Chem. Phys., 143, 014105

Yurchenko S., 2023, Computational Spectroscopy of Polyatomic Molecules. CRC Press, doi:https://doi.org/10.1201/9780429154348

Yurchenko S. N., Tennyson J., 2014, MNRAS, 440, 1649

Yurchenko S. N., Carvajal M., Jensen P., Herregodts F., Huet T. R., 2003, Chem. Phys., 290, 59

Yurchenko S. N., Carvajalz M., Jensen P., Lin H., Zheng J. J., Thiel W., 2005, Mol. Phys., 103, 359

Yurchenko S. N., Thiel W., Jensen P., 2007, J. Mol. Spectrosc., 245, 126

Yurchenko S. N., Barber R. J., Yachmenev A., Thiel W., Jensen P., Tennyson J., 2009, J. Phys. Chem. A, 113, 11845

Yurchenko S. N., Barber R. J., Tennyson J., Thiel W., Jensen P., 2011, J. Mol. Spectrosc., 268, 123

Yurchenko S. N., Tennyson J., Barber R. J., Thiel W., 2013, J. Mol. Spectrosc., 291, 69

Yurchenko S. N., Tennyson J., Bailey J., Hollis M. D. J., Tinetti G., 2014, Proc. Nat. Acad. Sci., 111, 9379
Yurchenko S. N., Yachmenev A., Ovsyannikov R. I., 2017, J. Chem. Theory Comput., 13, 4368