An experimentally-accurate and complete room-temperature infrared HCN line-list for the HITRAN database

Georg Ch. Mellau\textsuperscript{a}, Vladimir Yu. Makhnev\textsuperscript{b,*}, Iouli E. Gordon\textsuperscript{c}, Nikolay F. Zobov\textsuperscript{b}, Jonathan Tennyson\textsuperscript{d}, Oleg L. Polyansky\textsuperscript{d,b}

\textsuperscript{a}Physikalisch-Chemisches Institut, Justus-Liebig-Universit"at Gießen, Heinrich-Buff-Ring 17, D-35392 Gießen, Germany
\textsuperscript{b}Institute of Applied Physics, Russian Academy of Sciences, 46 Ulyanov Street, Nizhny Novgorod, 603950, Russia
\textsuperscript{c}Harvard-Smithsonian Center for Astrophysics, Atomic and Molecular Physics Division, 60 Garden St, Cambridge, MA, USA
\textsuperscript{d}Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom

Abstract

A hydrogen cyanide line list (MOMeNT-90) developed for the HITRAN spectroscopic database covering 0–7500 cm\textsuperscript{-1} range ($\lambda > 1330$ nm) is presented. The line list is a combination of the variationally calculated line intensities with line centers obtained from experimentally derived energy levels. There are four features of this line list which distinguishes it from the previously calculated ones. First, the intensities are variationally calculated using a new, high-accuracy potential energy surface (PES) obtained via fitting the PES using experimental energy levels. Second, a new \textit{ab initio} dipole moment surface was calculated at a high level of quantum chemical theory. Based on the wave functions calculated with the new PES and use of the new dipole moment surface, line intensities are reported which of similar accuracy to those obtained experimentally. Third, the calculated states are mapped to the existing complete set of experimental eigenenergies, resulting in an assigned and complete HCN line list down to the HITRAN intensity threshold of $10^{-34}$ cm/molecule. Fourth, extensive validation of the line list is provided through line-by-line comparisons of the results with measured HCN spectra which confirms the accuracy of the intensities used to construct the line list. The line list is augmented with parameters needed to calculate line widths for pressure-dependent simulations.

Keywords: HCN, hydrogen cyanide, molecular spectra, infrared spectra, HCN line list

1. Introduction

Hydrogen cyanide (HCN) is a trace atmospheric species and is present in the atmosphere mainly as a product of biomass burning. Atmospheric HCN concentration is a clear indicator of forest fires \textsuperscript{[1]}. Combating forest fires is an important aspect in the fight against global warming. Accurate spectroscopic line lists are required for atmospheric HCN concentration monitoring.

HCN is an important astrophysical species. In the solar system it has been detected in the atmospheres of gas giants \textsuperscript{[2,3]} and Saturn’s moon Titan \textsuperscript{[4]} as well as comets \textsuperscript{[5,7]}. Interstellar HCN is well-known \textsuperscript{[5]}, and it forms an important component of the atmosphere of cool stars \textsuperscript{[9,10]} and circumstellar disks \textsuperscript{[11]}. Recently HCN has been detected in a number of exoplanets \textsuperscript{[12,14]}. The temperature of the molecular gas determines the type of line list required for an analytic-type spectroscopic analysis. Thus, there are “cold” (room temperature) line lists such as HITRAN \textsuperscript{[15]} or GEISA \textsuperscript{[16]} and “hot” line lists such as HITEMP \textsuperscript{[17,18]} or ExoMol \textsuperscript{[19]}. Ideally, the data used to produce a line list originate from high-precision, high-resolution spectroscopic measurements. Due to the sheer number of the transitions required, this is often not possible. As a result many line lists are a combination of experimental data sets and calculated \textit{ab initio} predictions.

The first \textit{ab initio} calculation of HCN band strength was performed by Jørgensen \textit{et al.} \textsuperscript{[20]} who generated an \textit{ab initio} line list for hot HCN which they used for stellar modelling \textsuperscript{[9]}. An improved line list by Harris \textit{et al.} \textsuperscript{[21]} was based on the \textit{ab initio} potential energy surface (PES) and dipole moment surface (DMS) of Van Mourik \textit{et al.} \textsuperscript{[22]}. The \textit{ab initio} dataset covers all states of the [H,C,N] molecular system up to the isomerization barrier with a typical error for low J rovibrational transitions of 1-20 cm\textsuperscript{-1} \textsuperscript{[23]}. Subsequent line lists \textsuperscript{[24-27]} have improved the original line list of Harris \textit{et al.} by replacing calculated energy levels with available empirical or experimental energy levels.

The best \textit{ab initio} HCN rovibrational eigenenergies based on the Born-Oppenheimer (BO) approximation with all possible correction terms are still not accurate enough for the needs of the scientific community using spectroscopy as an analytical tool. For small molecules, it is not a problem to use experimental spectroscopy to detect the fundamental vibrations and also many of the rovibrational states with a few quanta of vibrational excitation. Absorption
spectroscopy with very long path length or at high temperature can extend the range of vibrational and rotational excitation covered by experiment, but complete coverage of the eigenenergy ladder remains difficult to achieve with this spectroscopic technique.

In 1997, one of us [28, 29] showed that a special setup of a tube furnace type emission experiment, referred to as HOT GAs Molecular Emission (HOTGAME) spectroscopy, can cover the complete eigenenergy ladder of triatomic molecules such as HCN or CO2. The analysis of his first set of HCN/HNC spectra was performed using the spectroscopic programs and techniques available at that time [30–32]. The HOTGAME emission spectra are very dense and sample an extremely large dynamic range of line intensities. To extract both the transitions between the very highly excited states, and to obtain accurate line positions as well as intensities for the assigned lines, a dedicated analysis program was developed [33].

After further significant improvement of the experiment sensitivity, Mellau introduced the concept of an experimental “Complete set of eigenenergies” for HCN and HNC [23, 24]. This refers to a complete list of experimental eigenstates with an relative accuracy of better than 10^{-8}, where “complete” means that not a single eigenenergy is missing from the set up to a certain highly excited cutoff eigenenergy value. These original data sets were extended with additional band series [33, 34]. The line list reported in this work are based on these HCN molecular eigenenergy lists.

Recently, some of us have constructed an HCN PES capable of predicting HCN infrared transitions with an accuracy of about 0.037 cm^{-1} [37], about two orders of magnitude better than the original calculations of Harris et al. In this paper, a further improvement of the PES by a factor of about 3 is described. This improvement is based on theoretical insights gained by mapping the experimental eigenenergies with the ab initio data. In this work, we combine the intensities obtained with our greatly improved PES and DMS with the experimental energy levels to provide a spectroscopically accurate line list for HCN.

The paper is organized as follows. In the section 2 we present the details of the PES calculation. In the section 3 we describe the ab initio calculations of the line intensities of our line list which we call var-MOMeNT-90. In the section 4 we map the line list which we call var-MOMeNT-90 to the ab initio data set. Only such an extensive basis set could provide a good starting point for a fit of the experimental levels.

2. HCN potential energy surface (PES)

As a starting point of the fit of the PES to experimental energy levels we used the ab initio surface obtained in [38]. This PES was based on MRCI+P-r energies (multi-reference configuration interaction with Pople correction, relaxed reference) computed with an aug-cc-pCV6Z basis set. Only such an extensive basis set could provide a good starting point for a fit of the experimental levels.

2.1. Non-adiabatic effects on HCN spectrum

There are two mass-dependent corrections to Born-Oppenheimer approximation: BODC (Born-Oppenheimer Diagonal Correction, or adiabatic correction) and non-adiabatic correction. An adiabatic correction has been already determined for HCN molecule in the recent work on ab initio potential energy surfaces [38].

Non-adiabatic corrections for HCN, or indeed other small molecules, are harder to characterize and therefore more rarely considered. We investigated inclusion of this correction previously [37, 38] as they proved necessary to achieve a good fit of the empirical energy levels up to
$J = 10$; with them these levels were reproduced with a standard deviation of $\sigma = 0.0373$ cm$^{-1}$.

However, at the start of this project we compared the HCN energy level predictions for $J > 20$ with empirical ones due to Mellau et al. [23, 39] and noticed that rotational part of energy deviates from the experimental value for each state in a systematic manner. We suspected that this behavior was caused by our attempt to improve our energies by including a non-adiabatic correction using both atomic masses in the calculation and as part of the fit itself. This hypothesis led us to try to fit levels without making full use of the non-adiabatic correction on the assumption that these both components are not consistent with each other and their effect should be reduced. We used the nuclear mass for the hydrogen atom and retained atomic masses for the carbon and nitrogen atoms. This combination was used because the hydrogen mass plays the main role in non-adiabatic correction to rotational-vibrational states.

2.2. Nuclear motion calculations

Vibrational energy levels were calculated using the DVR3D program suite [27]. The parameters used are presented in Table 1; Morse-like oscillators were used for the radial basis functions. We include almost all corrections used in the ab initio PES: adiabatic and relativistic correction surfaces and a nonadiabatic correction (by choosing atomic masses), to get the best starting point for fitting. More details about corrections are given in Ref. [38].

2.3. Potential energy surface fitting

Our approach has already been described in detail [37]. Our fit started from the ab initio data of Ref. [38] fitted to a polynomial form:}

\[ V_{ai}^{HCN}(r_1, r_2, \theta) = \sum_{ijk} b_{ijk} s_1^i s_2^j s_\theta^k, \]  

\[ s_1 = r_1 - r_{HCN}^{1e}, \]

\[ s_2 = r_2 - r_{HCN}^{2e}, \]

\[ s_\theta = \cos \theta - \cos \theta_{HCN}^{i}, \]

where $r_1$ is the CH bond length, $r_2$ is the CN bond length and $\theta$ is H–C–N angle (likewise, as defined in [37]), and $(r_{HCN}^{1e}, r_{HCN}^{2e}, \theta_{HCN}^{i})$ corresponds to the equilibrium configuration of HCN (we retain HCN superscripts here for the consistency with paper [37] where both HCN and HNC systems are considered.). This fit to ab initio data (900 points, 54 constants, see [37] for details) gives a standard deviation of about 1.5 cm$^{-1}$, which can be compared to the fit global covering both HCN and HNC wells performed in [38], which gave a standard deviation of about 2.6 cm$^{-1}$. $V_{ai}^{HCN}(r_1, r_2, \theta)$ provides the starting point for a fit to the observed HCN energy levels:

\[ V(r_1, r_2, \theta)^{HCN} = V_{ai}^{HCN}(r_1, r_2, \theta) + \sum_{ijk} d_{ijk} s_1^i s_2^j s_\theta^k. \]  

Table 2: HCN vibration energy levels. Experimental values from work of Mellau [23]. All values are in cm$^{-1}$.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(0 0 0)</td>
<td>1411.413</td>
<td>1411.431</td>
<td>-0.017</td>
<td>-0.035</td>
</tr>
<tr>
<td>(0 0 1)</td>
<td>2096.845</td>
<td>2096.868</td>
<td>-0.023</td>
<td>-0.042</td>
</tr>
<tr>
<td>(0 0 2)</td>
<td>2802.958</td>
<td>2802.963</td>
<td>-0.004</td>
<td>-0.030</td>
</tr>
<tr>
<td>(0 0 3)</td>
<td>3311.477</td>
<td>3311.484</td>
<td>-0.007</td>
<td>-0.005</td>
</tr>
<tr>
<td>(0 0 4)</td>
<td>3502.121</td>
<td>3502.117</td>
<td>0.003</td>
<td>-0.001</td>
</tr>
<tr>
<td>(0 0 5)</td>
<td>4173.070</td>
<td>4173.100</td>
<td>-0.029</td>
<td>-0.071</td>
</tr>
<tr>
<td>(0 0 6)</td>
<td>4174.608</td>
<td>4174.601</td>
<td>0.007</td>
<td>0.013</td>
</tr>
<tr>
<td>(1 0 0)</td>
<td>4684.310</td>
<td>4684.323</td>
<td>-0.013</td>
<td>-0.029</td>
</tr>
<tr>
<td>(0 1 0)</td>
<td>4888.039</td>
<td>4888.052</td>
<td>-0.012</td>
<td>-0.033</td>
</tr>
<tr>
<td>(0 2 0)</td>
<td>5393.697</td>
<td>5393.689</td>
<td>0.008</td>
<td>0.013</td>
</tr>
<tr>
<td>(0 3 0)</td>
<td>5525.812</td>
<td>5525.824</td>
<td>0.012</td>
<td>0.021</td>
</tr>
<tr>
<td>(0 4 0)</td>
<td>6036.960</td>
<td>6036.968</td>
<td>-0.008</td>
<td>-0.019</td>
</tr>
<tr>
<td>(0 5 0)</td>
<td>6228.598</td>
<td>6228.601</td>
<td>-0.003</td>
<td>-0.004</td>
</tr>
<tr>
<td>(0 6 0)</td>
<td>6254.405</td>
<td>6254.401</td>
<td>0.006</td>
<td>0.056</td>
</tr>
<tr>
<td>(0 7 0)</td>
<td>6519.610</td>
<td>6519.612</td>
<td>-0.001</td>
<td>-0.011</td>
</tr>
<tr>
<td>(1 0 0)</td>
<td>6760.705</td>
<td>6760.729</td>
<td>-0.023</td>
<td>-0.037</td>
</tr>
<tr>
<td>(0 1 1)</td>
<td>6855.443</td>
<td>6855.429</td>
<td>0.013</td>
<td>-0.058</td>
</tr>
<tr>
<td>(1 0 1)</td>
<td>6951.812</td>
<td>6951.723</td>
<td>-0.004</td>
<td>-0.053</td>
</tr>
<tr>
<td>(0 2 1)</td>
<td>7369.443</td>
<td>7369.453</td>
<td>-0.010</td>
<td>0.041</td>
</tr>
<tr>
<td>(1 0 2)</td>
<td>7455.423</td>
<td>7455.390</td>
<td>0.033</td>
<td>-0.021</td>
</tr>
</tbody>
</table>

We took $(r_{1e}, r_{2e}, \theta_e) = (1.066$ Å, $1.153$ Å, $180.0^\circ)$ as the minimum in the HCN well from the ab initio PES. Mellau [23] reported experimental characterization of all 3822 eigenenergies of HCN up to 6880 cm$^{-1}$ relative to the ground state using high temperature hot gas emission spectroscopy. This data set was used in this work to fit the PES. The set of ab initio polynomial coefficients from eq. (2) presented in the Ref. [37], the set of optimized polynomial coefficients from eq. (2) is presented in the Supplementary Material (Table 1).

2.4. Accuracy of the fitted potential

As shown below, significant improvement in both vibrational and rovibrational energy levels reproduction has been achieved. In particular, vibrational energy values were reproduced with a standard deviation $\sigma = 0.0172$ cm$^{-1}$. The standard deviation for energy levels for the $J$s included to fit $(0, 2, 5, 9)$ roughly corresponds with the vibrational one. There are two reasons for this: the nature of the fit itself and the weak contribution of the rotational non-adiabatic correction to the vibrational levels.

The deviation from the experimental levels with $J > 10$ grows with $J$. Some high vibrational-rotational levels can deviate from the experiment by about 1 cm$^{-1}$. Even taking into account the fact that the nonadiabatic correction still requires improvement, it is worth noting such an unprecedented level of predictive ability of the highly excited ro-vibrational states of the new PES.
The var-MOMeNT-90 HCN line list is built on our new spectroscopically-detemined PES. This PES reproduces the low $J$ experimental energy levels of hydrogen cyanide with a standard deviation $\sigma = 0.0121$ cm$^{-1}$ (for levels with $J = 0, 2$ and 5, see Table 3 for more details). A fitted PES is important not only for the accuracy of the line positions, but also for the accuracy of the line intensities [40]. There is a difference of few percents between the line intensities obtained using a pure ab initio PES and Table 3: HCN rovibration energy levels. Experimental values from work of Mellau [23]. All values are in cm$^{-1}$.

<table>
<thead>
<tr>
<th>$J$</th>
<th>$\sigma$ [23]</th>
<th>$\sigma$ (this work)</th>
<th>$N_{\text{levels}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0382</td>
<td>0.0172</td>
<td>21</td>
</tr>
<tr>
<td>0&amp;2</td>
<td>0.0312</td>
<td>0.0133</td>
<td>90</td>
</tr>
<tr>
<td>0&amp;2&amp;5</td>
<td>0.0287</td>
<td>0.0121</td>
<td>199</td>
</tr>
<tr>
<td>0&amp;2&amp;5&amp;9</td>
<td>0.0330</td>
<td>0.0166</td>
<td>322</td>
</tr>
<tr>
<td>0&amp;2&amp;5&amp;9&amp;10</td>
<td>0.0373</td>
<td>0.0175</td>
<td>445</td>
</tr>
<tr>
<td>0&amp;2&amp;5&amp;9&amp;10&amp;20</td>
<td>0.0926</td>
<td>0.0340</td>
<td>568</td>
</tr>
</tbody>
</table>

For all fundamental transitions the DMS reproduces the measurements better than the previous work of van Mourik et al. [22]. The accuracy of the calculated line intensities for two of the fundamental transitions is about 1 %, this is within the experimental uncertainties of very accurate experimental intensity measurements. Since the line list for this work was constructed for the eigenenergy range 0-7200 cm$^{-1}$, some transitions involving eigenstates in the 7200-10300 cm$^{-1}$ range may be less accurate.

The final ab initio line list was prepared for the 296 K HITRAN temperature with an intensity cutoff at $10^{-34}$ cm/molecule. In the calculations $J_{\text{max}} = 90$ was used, but due to the selected sensitivity of the cold line list, the transitions present in the list have $J$ quantum numbers well below the $J = 90$ limit. On the other hand, this means that none of the bands is limited at high $J$ due to the incompleteness of the data set, the limit is always given by the cutoff sensitivity of the line list.

The overall comparison between the HITRAN2016 database [15] and the ab initio line list of this work is shown in Figure 1 and Figure 2.

3. The var-MOMeNT-90 HCN line list

The var-MOMeNT-90 HCN line list is built on our new spectroscopically-detemined PES. This PES reproduces the low $J$ experimental energy levels of hydrogen cyanide with a standard deviation $\sigma = 0.0121$ cm$^{-1}$ (for levels with $J = 0, 2$ and 5, see Table 3 for more details). A fitted PES is important not only for the accuracy of the line positions, but also for the accuracy of the line intensities [40]. There is a difference of few percents between the line intensities obtained using a pure ab initio PES and the same PES but spectroscopically tuned [37].

The spectroscopically tuned PES was used together with the new ab initio DMS previously reported [37] to calculate the line centers and line intensities using the DVR3D program suite [27]. The ab initio DMS was created for the HCN well to cover all transitions between the energy levels in the 0-7200 cm$^{-1}$ eigenenergy range. An aug-cc-pCV5z basis set with the Pople relaxed correction and a relativistic correction, mass-velocity plus first-order Darwin term (MVD1) were used to construct the new DMS.

For all fundamental transitions the DMS reproduces the measurements better than the previous work of van Mourik et al. [22]. The accuracy of the calculated line intensities for two of the fundamental transitions is about 1 %, this is within the experimental uncertainties of very accurate experimental intensity measurements. Since the line list for this work was constructed for the eigenenergy range 0-7200 cm$^{-1}$, some transitions involving eigenstates in the 7200-10300 cm$^{-1}$ range may be less accurate.

The final ab initio line list was prepared for the 296 K HITRAN temperature with an intensity cutoff at $10^{-34}$ cm/molecule. In the calculations $J_{\text{max}} = 90$ was used, but due to the selected sensitivity of the cold line list, the transitions present in the list have $J$ quantum numbers well below the $J = 90$ limit. On the other hand, this means that none of the bands is limited at high $J$ due to the incompleteness of the data set, the limit is always given by the cutoff sensitivity of the line list.

The overall comparison between the HITRAN2016 database [15] and the ab initio line list of this work is shown in Figure 1 and Figure 2.

4. Mapping the experimental data to the ab initio data

A one-to-one mapping between the experimental and the ab initio sets of energy levels is possible only after expanding the good quantum numbers used to label the ab initio eigenenergies with approximate vibrational quantum numbers. Matching between the data sets with a search for the nearest neighbor seems to be straightforward if one considers only the accuracy of the theoretical pure vibrational eigenenergies. However, this is actually a fallacy; even for the pseudo-theoretical eigenenergy list reported in this paper, such matching is not possible. With rotational excitation, this accuracy deteriorates by several orders of magnitude and the density of states increases considerably due to the sub-states with non-zero vibrational angular momentum.

In this work, we have used a universal mapping method developed in [23] that is based on the spectroscopic patterns of molecular eigenenergies. In an ongoing research project, this method is being implemented in Giessen as a computational package. The idea behind this project is to obtain an immediate and fully automatic response to the
quality of any molecular physics theory that computes a complete set of eigenenergies for a molecular system. The response is a high quality visualisation of the differences between measurement and theory. As described in section 2.1 the first visualisation output of the differences for the eigenenergies calculated with the initial PES allowed us to significantly improve the fitted PES and eigenenergy calculations (see differences between the two σ of the last row of the Table 5).

As a first application, the eigenenergy data set calculated in this work was mapped onto the experimental HCN data set for all vibrational states up to 10300 cm⁻¹. The result of this analysis is a J = ℓ, . . . 90 manifold of ab initio and experimental pair of rovibrational eigenenergies. For each vibrational state the experimental eigenenergies were extrapolated from the state-specific J_{max} detected in the spectrum to J = 90 with the state specific spectroscopic constants. A few vibrational states within the 7000-10300 cm⁻¹ eigenenergy wavenumber range were not measured; for these bands spectroscopically predicted eigenenergies were mapped to the ab initio data.

This complete rovibrational data set was used to generate the HCN line list.

5. The MOMeNT-90 HCN line list

This paper provides a line list for transitions that are missing from the HITRAN database or improve the line intensities present in the HITRAN2016 database. Our data set is limited first by the 10300 cm⁻¹ cutoff of the vibrational energies considered in this work. The second limitation is related to the sensitivity cutoff used in generating the pure ab initio line list, this was set for this work to 10⁻³⁴ cm/molecule. We have excluded from the line list some of the vibrational states predicted only spectroscopically.

After mapping the experimental data set to the theoretical one, spectroscopically defined band series were extracted from the data set. This allows a more precise control of the data and the validation of the data set band by band.

The natural way to partition the set of HCN molecular transitions is to fix the stretching quantum numbers v₁ and v₃ for the lower and upper states, respectively, and consider all possible transitions matching such a setting. The reported transitions are thus split into data files containing all transitions between states with fixed stretching quantum numbers v₁ and v₃.

The results are reported in two sets of line list files. One set contains the continuous list of line data for a given vibrational series. These files are named MAPDATA and can be imported directly by user analysis programs.

The same data are presented in a second file organized into spectroscopic bands. These files are named MAPRESULT and can be used for spectroscopic band-by-band analysis of HCN spectra. The bands in these files were mapped to the HITRAN database, and the transition present in HITRAN are listed directly below the lines containing the data of this work. The MAPDATA and MAPRESULT files are part of the supplementary data published with this work.

If a line is present in HITRAN, a final column is added with the difference between the line position compiled in this work and the line position from the HITRAN database. For a large portion of the bands compiled in this work that are also present in HITRAN, this column lists 0.00000 values for all band lines. These transitions appear to have already been compiled for HITRAN with the Melmon eigenenergy tables used also in this work, a reference to the corresponding experimental papers appears to be missing from the HITRAN2016 database.

Figures 3 and 4 show the overview plot for the subbands of three MAPDATA files. The highest line intensity of the bands are marked by the transition quantum number labels.

![Figure 3](image1.png)

Figure 3: Overview of the 0v₂0 → 0(v₂ + Δv₂)0 line list.

![Figure 4](image2.png)

Figure 4: Overview of the 0v₂0 → 1(v₂ + Δv₂)0 line list.
6. Validation of the new HCN transition data set

There is a large difference in the number of molecular eigenstates and the number of possible transitions between these states that are relevant for spectroscopic measurements. For a triatomic molecule - and especially for HCN - it is possible to experimentally sample the complete eigenenergy ladder by HOTGAME measurements. Once one has the experimental eigenergies, it is possible to calculate any possible transition frequency with experimental accuracy. It is not difficult to see that obtaining a complete set of accurate experimental line intensities for this set of transition frequencies is quite a different problem.

The accuracy of standard experimental intensity analysis is very limited for various technical and chemical reasons. A set of line intensities accurate to 1 percent can be considered an accurate experimental data set. This low relative accuracy, compared to the relative accuracy of $10^{-8}$ needed to obtain experimental grade results for line position, can be achieved by accurate \textit{ab initio} calculations \cite{22}; transition probabilities can be calculated with an accuracy similar to the majority of experimental results. The only problem here is the validation of the theoretical results.

To validate the theoretical line intensities calculated in this work, we compare simulated spectra with spectra recorded in Gießen. Here we present the comparison with the absorption spectrum labeled AA (a selected part of this spectrum is shown in Figure 6h), which has already been reported in reference \cite{43}. The line intensities of this absorption spectrum was revisited in reference \cite{44}. In the cited work, the AA absorption spectrum was used to calibrate the intensity axis of some emission spectra. Calibrated emission spectra can be used to extract accurate absolute line intensities. Such an analysis offers the prospect of covering wavenumber regions with a complete set of experimental line intensities, even for hot line lists. The absorption spectrum AA was compared in reference \cite{44} to the intensity measurements of Devy et al. \cite{45}. The result of this analysis was that the effective pressure of the AA experiment is slightly lower than the value measured with the pressure transducer due to the adsorption of the HCN molecules on the metallic surface of the cell.

In order to give an error bar for the observed line intensities, we have to consider two main effects. The temperature of the molecular gas is set to the average temperature in the laboratory during the measurement. The effect of this uncertainty corresponds to an intensity error of at most 1%. The difference between the recorded and the calibrated pressure corresponds to a relative error of the intensities of 8%. By calibrating the pressure, we think that this error is reduced to at least half of this value. We think that the experimental absorption spectrum should match our simulations within 5% (considering the \textit{ab initio} line intensity parameter error-free) and that this is a conservative estimation of the expected error bar. Pressure calibration was performed independently of the actual measurements.

Figure 1 is a standard method for comparing line lists. Only basic differences between line lists can be made visible with such a figure. Because of the sheer number of transitions, it is not obvious how to compare line lists or how to compare a line list with a measured spectrum. For example, one can pick the spectral region of a band and plot a measured spectrum and a simulated spectrum on the same axis. This is still not better, as can be seen in the Figure 6. The width of the lines in high resolution spectra is much smaller than the wavenumber spacing between the lines, and thus the experimental and simulated spectra overlap completely.

Figure 6 shows a possible visualization of a spectroscopic band, a \textit{band plot}. Here, a small region around each line of a subband is selected, and the axis regions between the lines are removed. The band plot allows a quick overview of the intensity agreement between experiment and theory. Here the problem remains that as the intensity of the lines decreases, the comparison becomes less clear.

Figure 6 shows the same band with the individual lines scaled by individual factors. This figure has now lost all typical band intensity and line spacing structure, but is probably the best visualisation of the intensity match.

Figure 6 shows another example of a band plot. Here, both low and high intensity lines of other bands overlap with the lines of the selected band. Figure 6 shows a \textit{corrected band plot}. In this plot, the simulated lines of other bands are shown (green curves) and these are used to correct the measured spectrum. For the lines labeled Q14, Q10 and Q9, this correction leads to a better intensity match between experiment and theory. Here the problem remains that as the intensity of the lines decreases, the comparison becomes less clear.

Figures 7 and 8 show the corrected band plots of the $0v_20 \rightarrow 0(v_2 + 1)1$ transitions covering the 2700-2900 cm$^{-1}$ wavenumber range. These bands are missing from the HI-
TRAN database. The validation figures show that the match between experiment and our new HCN line list is better than 5%.

Figures 9 and 10 show the corrected band plots of the 0v2→1(v2+1)0 transitions covering the 3150-3450 cm⁻¹ wavenumber range. These bands have been partially included in the HITRAN2016 database. The validation figures show that for the fundamental and first excited state our intensities agree with the HITRAN2016 database and the experiment better than 5%. For the higher excited states, the intensities in HITRAN2016 database are overestimated or missing, while the simulation based on the line list of this work which agrees to better than 5% with the measurements.

Figure 11 shows the band plot and the corrected band plot of the 05'0→15'0 PR e-e subband of the emission spectrum labeled ED [39]. First, we see that the even with the extended sensitivity of 10⁻³⁴ cm⁻¹/molecule relative to the HITRAN database not all overlapping bands are simulated. The high J lines of the subband are missing due to the chosen sensitivity. The highest intensity lines in the emission spectrum are slightly suppressed due to self-absorption. Again, we see a perfect match between our calculated intensity and the intensity of the lines in the experimental spectrum, this perfect agreement also holds for transitions between the highly excited states of the molecule.

In our data set, the 02→3 levels are spectroscopically predicted. There is a large difference of 25 cm⁻¹ between our line positions and those given in the HITRAN2016 database, which we assume is due to an error in HITRAN2016.

7. Linewidths

Apart from the transition wavenumbers νij, lower state energies and intensities Sij (along with quantum numbers of transitions states) it is necessary to have information about the line shape parameters f for modelling absorption κij in real applications, including atmospheric ones. For instance in the HITRAN database all of the transitions are given air- and self-broadened values. Additional information about definitions of the line shape parameters can be found in the HITRAN2004 paper [46] and to a wider extent on the HITRAN website (www.hitran.org).

In general, the broadening coefficients are rotationally and vibrationally dependent. However, as was shown by Yang et al. [47] for HCN, fits for states (0, 1, 0), (1, 0, 0) and pure-rotational states show that one can consider the vibrational dependence as negligible and use a single formula for the dependence of the broadening coefficient on the rotational number.

$$\gamma_{\text{air}}(p, T) = A_0 + A_1|m| + A_2|m|^2 + A_3|m|^3 + A_4|m|^4 \quad (3)$$

This formula was applied in HITRAN2008 [48] for all the lines with |m| > 40 (the previous polynomial used in HITRAN2004 was applicable only up to |m| = 29). The lines with |m| > 40 were assigned a constant half-width parameter of 0.0518 cm⁻¹ atm⁻¹, which corresponds to the polynomial value at |m| = 40. Since there is no evidence for pronounced vibrational dependence of the widths for the case of HCN so this is used for all the bands.

The form given in Eq. (3) is also used for $\gamma_{\text{self}}$ and $\gamma_{\text{air}}$. Coefficients for these constants were taken from HITRAN2004 paper [46]. As for $\gamma_{\text{air}}$, there are only limited measurements available and there currently not enough data for all bands represented in this work so we decided to not include this parameter.

In the future it would be productive to employ Pade approximants to be able to extrapolate the trend to the rotational quanta larger than 40. As demonstrated in Refs. 19 60 Pade approximants have better predictive abilities that polynomials and therefore do not require using a constant value beyond experimental values.

In this work we used PNNL spectra [51] to validate the line list, including the line-broadening parameters. Because PNNL spectra are N₂-broadening, Eq. (3) was used with nitrogen-associated parameters also derived from Yang et al. [47]. We used the HITRAN Application Programming Interface (HAPI) [52] to model the cross-sections. There are some examples of comparison PNNL spectra with modeled in this work (Figs. 12, 13, 14, 15). For demonstration purposes we also took H²CN line-list from upcoming HITRAN2020 [53] edition and add its absorption at the natural abundance. One should note that the line-mixing is not included for HCN therefore in some of the Q-branches one should not expect complete agreement between generated cross-sections and the pressure-broadened spectrum.

As one can see from the Figures very good agreement is generally observed. For example, the region of ν2 and 3ν2 bands is presented on Figure 12 where part of absorption listed in PNNL is underestimated by HITRAN2016 line list. We note that the peaks the simulated cross-sections are considerably higher than those of the PNNL cross-sections in the 2114-2120 cm⁻¹ region of Fig. 12. This region corresponds to the Q-branch of the 3ν2 band. The intensities of this branch are very sensitive to which DMS was used in the calculations, so some of them may be slightly overestimated by our calculations. But those transitions from Q-branch, which cause the most absorption, give the best match between the PNNL and our spectra.

The discrepancy between the PNNL spectra and the HITRAN2016 data is more noticeable in the lower absorption regions, as shown in Figure 13. Due to the weakness of the spectrum, the signal-to-noise ratio is lower here, making the band structure less clear. But it should be noted that its strongest part, the Q-band, is distinguishable and available for reliable comparison, as shown in Figure 13.

In addition to differences in intensities, differences in absorption may be due to missing isotopologue bands (Fig. 14) or weak bands (Fig. 15).
8. Conclusion

The very accurate MOMeNT-90 line list for the rovibrational spectrum of room temperature H\textsuperscript{14}C\textsuperscript{14}N for the region below 7500 cm\textsuperscript{-1} is presented. A unique attribute of this polyatomic line list comes from the fact that all the line positions, even for the very weak lines that are hardly observable experimentally, were derived from the experimental energy levels obtained from the corresponding high-temperature studies. The wavefunctions obtained with the new version of the spectroscopically-tuned \textit{ab initio} PES were used together with our accurate DMS to calculate accurate HCN line intensities at room temperature with a sensitivity of 10\textsuperscript{-34} cm/molecule. We demonstrated the accuracy of the line intensities using some selected absorption and emission spectra. Our new intensities are very different from the intensity data present in HITRAN.2016.

The creation of a hybrid line list requires an accurate mapping procedure of the theoretical line positions to the experimental ones. We think that our special mapping procedure is error-free and we were able to detect all such missassignments present in HITRAN.2016.

We provide band-by-band tables for all HCN transitions reported in this work. This dataset should support spectroscopic band-by-band analysis with our line list.

A recent scientific white paper \cite{55} signed by 82 members of the exoplanet community describes the need for atomic and molecular data for exoplanet research. The infrared line positions for important exoplanetary molecules such as HCN are required to be accurate to 0.1 km s\textsuperscript{-1}, which corresponds to 0.01 cm. 2021-0016 is also acknowledged.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded by RFBR through research projects number 18-02-00705 and 18-32-00698, ERC Advanced Investigator Project 883830 and the UK Natural Environment Research Council grant NE/1000767/1. OLP and GCM acknowledge the support of the Alexander von Humboldt Stiftung for the support of OLP visit to Giessen during the work on this paper. RF State Project 0030-2021-0016 is also acknowledged.

References

Figure 6: Line list validation figures: from full spectrum plot a) to corrected f) and scaled band plots d). The band plot figures allow a visual inspection of the quality of the line list. Experimental spectrum is drawn in blue, simulated spectra in red and green. Line P11e is plotted in yellow with an intensity parameter multiplied by 0.95 to facilitate visual estimation of the difference between measured and ab initio data.
Figure 7: Comparison between the line list of this work and an FT-IR spectrum. In these band plots, the individual lines of a subband are joined to allow a visual inspection of the line shapes. Experimental spectrum is drawn in blue, simulated spectra in red and green.
Figure 8: Comparison between the line list of this work and an FT-IR spectrum. In these band plots, the individual lines of a subband are joined to allow a visual inspection of the line shapes. Experimental spectrum is drawn in blue, simulated spectra in red and green. This figure shows some of the weak bands at the noise level.
Figure 9: Comparison between the line list of this work and an FT-IR spectrum. In these band plots, the individual lines of a subband are joined to allow a visual inspection of the line shapes. Experimental spectrum is drawn in blue, simulated spectra in red and green. Brown curves correspond to a simulation with HITRAN2016 line list.
Figure 10: Comparison between the line list of this work and an FT-IR spectrum. In these band plots, the individual lines of a subband are joined to allow a visual inspection of the line shapes. Experimental spectrum is drawn in blue, simulated spectra in red and green. Brown curves correspond to a simulation with HITRAN2016 line list.
Figure 11: Comparison between the line list of this work and an FT-IR emission spectrum. In these band plots, the individual lines of a subband are joined to allow a visual inspection of the line shapes. Experimental spectrum is drawn in blue, simulated spectra in red and green. The highest intensity lines in the emission spectrum are slightly suppressed due to self-absorption.

Figure 12: Cross-section of HCN and H$^{13}$CN (in natural abundance) from PNNL (black dots), HITRAN2016 (blue line) and MOMeNT-90 (this work)(red line) in region of $\nu_3$ and $3\nu_2$ bands. The H$^{13}$CN cross-section (from HITRAN2020 line list) was included to this work.
Figure 13: Cross-section of HCN and H\textsuperscript{13}CN (in natural abundance) from PNNL (black dots), HITRAN2016 (blue line) and MOMeNT-90 (this work) (red line) in region of ν\textsubscript{2} + 2ν\textsubscript{3} band. The H\textsuperscript{13}CN cross-section (from HITRAN2020 line list) was included to this work.

Figure 14: Cross-section of HCN and H\textsuperscript{13}CN (in natural abundance) from PNNL (black dots), HITRAN2016 (blue line) and MOMeNT-90 (this work) (red line) in region of 2ν\textsubscript{1} band. The H\textsuperscript{13}CN cross-section (from HITRAN2020 line list) was included to this work.
Figure 15: Cross-section of HCN and H$^{13}$CN (in natural abundance) from PNNL (black dots), HITRAN2016 (blue line) and MOMeNT-90 (this work) (red line) in region near 6910 cm$^{-1}$. The H$^{13}$CN cross-section (from HITRAN2020 line list) was included to this work.