An ab initio variationally computed room-temperature line list for $^{32}\text{S}^{16}\text{O}_3$

Daniel S. Underwood, Jonathan Tennyson and Sergei N. Yurchenko*

Ab initio potential energy and dipole moment surfaces are computed for sulfur trioxide ($\text{SO}_3$) at the CCSD(T)-F12b level of theory with appropriate triple-zeta basis sets. The analytical representations of these surfaces are used, with a slight correction, to compute pure rotational and rotation–vibration spectra of $^{32}\text{S}^{16}\text{O}_3$ using the variational nuclear motion program TROVE. The calculations considered transitions in the region 0–4000 cm$^{-1}$ with rotational states up to $J = 85$. The resulting line list of 174 674 257 transitions is appropriate for modelling room temperature $^{32}\text{S}^{16}\text{O}_3$ spectra. Good agreement is found with the observed infrared absorption spectra and the calculations are used to place the measured relative intensities on an absolute scale. A list of 10 878 experimental transitions is provided in a form suitable for inclusion in standard atmospheric and planetary spectroscopic databases.

I Introduction

Sulfur trioxide ($\text{SO}_3$) is a stable, planar, symmetric molecule whose electronic ground state is a closed shell. On earth it is a pollutant produced from smoke-stacks and other industrial exhausts. In the atmosphere $\text{SO}_3$ forms sulfuric acid with its association with acid rain. It is also corrosive inside combustion systems. In both cases the reactivity of $\text{SO}_3$ makes its production hard to monitor. $\text{SO}_3$ is produced naturally on earth as part of volcanic emissions. It is also thought to be a significant constituent of the atmosphere of Venus.

The infrared vibration–rotation spectrum of $^{32}\text{S}^{16}\text{O}_3$ (henceforth referred to as $\text{SO}_3$) has been extensively investigated in a series of papers by Maki and co-workers. Its “forbidden” rotational spectrum, for which centrifugal distortions can induce transitions, has been investigated using microwave Fourier-transform spectroscopy. However, although these works provide an extensive list of measured line frequencies, none of them report absolute transition intensities.

There has been limited theoretical work on $\text{SO}_3$. Dorney et al. reported force constants, while Martin computed an ab initio quartic force field using coupled cluster methods and reported theoretical estimates for the band origins of the low-lying vibrational states. Again, neither of these works considered transition intensities.

The lack of any absolute transition intensities for $\text{SO}_3$ places severe limitations on the use of its infrared spectrum for remote sensing applications or inclusion of this data in standard atmospheric and planetary spectroscopic databases. In this work we report the calculation of a new ab initio PES and associated dipole moment surface (DMS) for $\text{SO}_3$. These are used to not only produce theoretical spectra for $^{32}\text{S}^{16}\text{O}_3$, but also to place the relative intensity measurements of its infrared spectrum on an absolute scale.

II The potential energy and dipole moment surfaces

The ab initio PES was computed using the recently-proposed explicitly correlated F12 singles and doubles coupled cluster method including a perturbational estimate of connected triple excitations, CCSD(T)-F12b, in conjunction with the corresponding F12-optimized correlation consistent basis sets, namely the valence correlation-consistent functions aug-cc-pVTZ-F12 and aug-cc-pV(T+d)Z-F12 for oxygen and sulfur, respectively. We also utilized the XoPT, cc-pV5Z/JKFT1 and aug-cc-pvCV5Z/MP2F1 auxiliary basis sets for evaluating the many-electron integrals, the exchange and Fock operators, and the remaining electron repulsion integrals, respectively. The value of the geminal Slater exponent $\beta$ was chosen as 1.2. Relativistic corrections were introduced as scalar relativistic effects and the diagonal Born–Oppenheimer corrections (DBOCs) were also computed. Molpro.2010 was employed for all electronic structure calculations. We used a grid of 13 000 points with the bond lengths $r_i$ ($i = 1,2,3$) ranging from 1.1 to 2.0 Å, the interbond angles $\omega_k$ ($k = 1,2,3$).
between 70 and 123°, and the inversion angle \( \tilde{\rho} \) between 0 and 50°. Grid points were specified based on an algorithm which produced a higher concentration of points close to the equilibrium position in a coupled-cluster manner. Each point was processed using an MP2 calculation as a further filter, and those producing an MP2 result of 50 000 cm\(^{-1}\) or higher were disregarded. This grid is sufficient to approach dissociation. The T1 diagnostic test\(^{23} \) was applied to CCSD(T) is not very reliable for the higher energy regions which adequate for all geometries considered. The quality of the \( ab\ initio \) PES can be assessed by calculating the fundamental term values of the molecule. Employing the variational approach TROVE\(^{30} \) (see below) with the \( ab\ initio \) PES we obtain 1073.56, 499.89, 1396.75, and 529.74 cm\(^{-1}\) for \( \nu_1, \nu_2, \nu_3, \) and \( \nu_4 \) respectively. The corresponding experimentally derived values are listed in Table 1. While the bending modes are reproduced reasonably well, the stretching fundamental term values exhibit large deviations from the experiment, 5–9 cm\(^{-1}\). In order to improve the quality of the stretching modes we have used the following morphing procedure.\(^{31} \) We re-expand the \( ab\ initio \) PES (1) around an intermediate reference geometry \( r_e^\text{ref} \), which is now also an \( ab\ initio \) point with a root-mean-square (rms) error of 0.067 cm\(^{-1}\). Weight factors were set using the expression suggested by Partridge and Schwenke:\(^{29} \)

\[
N = \text{tanh}[-0.0005 \text{ cm} \times (V_{i} - 16 000 \text{ cm}^{-1})] + 1.00000002
\]

where \( V_{0}(\sin \tilde{\rho}) = \sum_{j=1}^{3} f_{0j}(\sin \rho_{j} - \sin \tilde{\rho})^{s} \) and \( F_{jk}(\sin \tilde{\rho}) = \sum_{j=0}^{3} f_{jk...}(\sin \rho_{j} - \sin \tilde{\rho})^{s} \). Here \( \sin \rho_{j} \) is the equilibrium value of \( \sin \tilde{\rho} \), \( s \) is a molecular parameter, and the quantities \( f_{0j}^{(s)} \) and \( f_{jk...}^{(s)} \) in eqn (6) and (7) are expansion coefficients. The linear stretching terms \( F_{10}^{(s)} \) and \( F_{ab}^{(s)} \) also vanish because of the symmetric arguments. The same form has been used to represent the PESs of NH\(_{3}\), PH\(_{3}\), SbH\(_{3}\), NH\(_{2}\), and BiH\(_{3}\). The potential parameters \( f_{jk...}^{(s)} \) were obtained through a weighted least squares fit to the \( ab\ initio \) points with a root-mean-square (rms) error of 0.067 cm\(^{-1}\). Weight factors were set using the expression suggested by Partridge and Schwenke:\(^{29} \)

\[
N = \text{tanh}[-0.0005 \text{ cm} \times (V_{i} - 16 000 \text{ cm}^{-1})] + 1.00000002
\]

This PES is represented by an analytical form given by the expansion\(^{24} \)

\[
V(\tilde{\xi}_1, \tilde{\xi}_2, \tilde{\xi}_3, \tilde{\xi}_4; \sin \tilde{\rho}) = V_{C} + V_{0}(\sin \tilde{\rho}) + \sum_{j} f_{j}(\sin \tilde{\rho}) \tilde{\xi}_j
\]

\[
+ \sum_{j\neq k} F_{jk}(\sin \tilde{\rho}) \tilde{\xi}_j \tilde{\xi}_k
\]

\[
+ \sum_{j<k<l} F_{jkl}(\sin \tilde{\rho}) \tilde{\xi}_j \tilde{\xi}_k \tilde{\xi}_l
\]

\[
+ \sum_{j<k<l<m} F_{jklm}(\sin \tilde{\rho}) \tilde{\xi}_j \tilde{\xi}_k \tilde{\xi}_l \tilde{\xi}_m
\]

where \( V_{0}(\sin \tilde{\rho}) = \sum_{j=1}^{3} f_{0j}(\sin \rho_{j} - \sin \tilde{\rho})^{s} \) and \( F_{jk...}(\sin \tilde{\rho}) = \sum_{j=0}^{3} f_{jk...}(\sin \rho_{j} - \sin \tilde{\rho})^{s} \). Here \( \sin \rho_{j} \) is the equilibrium value of \( \sin \tilde{\rho} \), \( s \) is a molecular parameter, and the quantities \( f_{0j}^{(s)} \) and \( f_{jk...}^{(s)} \) in eqn (6) and (7) are expansion coefficients. The linear stretching terms \( F_{10}^{(s)} \) and \( F_{ab}^{(s)} \) also vanish because of the symmetric arguments. The same form has been used to represent the PESs of NH\(_{3}\), PH\(_{3}\), SbH\(_{3}\), NH\(_{2}\), and BiH\(_{3}\). The potential parameters \( f_{jk...}^{(s)} \) were obtained through a weighted least squares fit to the \( ab\ initio \) points with a root-mean-square (rms) error of 0.067 cm\(^{-1}\). Weight factors were set using the expression suggested by Partridge and Schwenke:\(^{29} \)

\[
N = \text{tanh}[-0.0005 \text{ cm} \times (V_{i} - 16 000 \text{ cm}^{-1})] + 1.00000002
\]

This PES is represented by an analytical form given by the expansion\(^{24} \)

\[
V(\tilde{\xi}_1, \tilde{\xi}_2, \tilde{\xi}_3, \tilde{\xi}_4; \sin \tilde{\rho}) = V_{C} + V_{0}(\sin \tilde{\rho}) + \sum_{j} f_{j}(\sin \tilde{\rho}) \tilde{\xi}_j
\]

\[
+ \sum_{j\neq k} F_{jk}(\sin \tilde{\rho}) \tilde{\xi}_j \tilde{\xi}_k
\]

\[
+ \sum_{j<k<l} F_{jkl}(\sin \tilde{\rho}) \tilde{\xi}_j \tilde{\xi}_k \tilde{\xi}_l
\]

\[
+ \sum_{j<k<l<m} F_{jklm}(\sin \tilde{\rho}) \tilde{\xi}_j \tilde{\xi}_k \tilde{\xi}_l \tilde{\xi}_m
\]

where \( V_{0}(\sin \tilde{\rho}) = \sum_{j=1}^{3} f_{0j}(\sin \rho_{j} - \sin \tilde{\rho})^{s} \) and \( F_{jk...}(\sin \tilde{\rho}) = \sum_{j=0}^{3} f_{jk...}(\sin \rho_{j} - \sin \tilde{\rho})^{s} \). Here \( \sin \rho_{j} \) is the equilibrium value of \( \sin \tilde{\rho} \), \( s \) is a molecular parameter, and the quantities \( f_{0j}^{(s)} \) and \( f_{jk...}^{(s)} \) in eqn (6) and (7) are expansion coefficients. The linear stretching terms \( F_{10}^{(s)} \) and \( F_{ab}^{(s)} \) also vanish because of the symmetric arguments. The same form has been used to represent the PESs of NH\(_{3}\), PH\(_{3}\), SbH\(_{3}\), NH\(_{2}\), and BiH\(_{3}\). The potential parameters \( f_{jk...}^{(s)} \) were obtained through a weighted least squares fit to the \( ab\ initio \) points with a root-mean-square (rms) error of 0.067 cm\(^{-1}\). Weight factors were set using the expression suggested by Partridge and Schwenke:\(^{29} \)
has been adopted for the rest of calculations for \( ^{32}\text{S}^{16}\text{O}_3 \) presented below.

The DMS was calculated using the same level of theory as the PES and on the same grid of 13,000 points. The \( \text{ab initio} \) values were then expressed analytically using the symmetrized molecular bond (SMB) representation described in detail in ref. 32. The resulting dipole moment parameters obtained through a least squares fit reproduce the \( \text{ab initio} \) data with an rms error of 0.00013 D. In these fits the factor defined by eqn (8) was used to weight the geometries according to the corresponding energies. To our knowledge there are no experimental or \( \text{ab initio} \) dipole moment data in the literature that we could use to validate our DMS against. However, our experience of working dipole moments for different systems\(^{33-35}\) shows that high quality \( \text{ab initio} \) intensities in most cases are competitive with experimental measurements. The quality of the relative intensities calculated using our DMS is discussed below.

The potential energy (both \( \text{ab initio} \) and morphed) and dipole functions used in the present work are given in the ESI.\(^{\dagger}\)

### III Ro-vibrational calculations

#### A Basis set convergence and Hamiltonian optimisation

Ro-vibrational calculations for \( \text{SO}_3 \) were performed using the program TROVE\(^{30}\) employing the same computational procedures used for the ammonia molecule.\(^{36,37}\) These molecules share the same molecular symmetry group \( D_{3h}(M) \).\(^{36}\) In order to achieve results of high accuracy as well as to minimise the requirement for computational resources, it is necessary to optimise the size of the Hamiltonian matrix. This involves preliminary truncation of the basis set, as well as limiting the order of both the kinetic and potential components of the Hamiltonian expansion (see ref. 30). Deciding upon these expansion orders involves a trade-off between accuracy and computational expense; it is necessary to initially ensure that the need for computational resource is at a minimum, while maintaining a good degree of accuracy. TROVE employs a polyad number truncation which controls the size of the basis set. For \( \text{SO}_3 \) the polyad number is given by

\[
P = 2(n_1 + n_2 + n_3) + n_4 + n_5 + n_6. \tag{9}
\]

where \( n_i \) are the quanta associated with 1D basis functions, \( \phi_i \) whose product gives our vibrational basis set.\(^{30}\) Each of these basis functions is associated with an internal coordinate \( \xi_i \), and only functions for which \( P \leq P_{\text{max}} \) are included in the primitive basis set. Initial tests were carried out to measure the degree of convergence using different values for \( P_{\text{max}} \) and the orders of the kinetic and potential energy expansions. In this work we use a kinetic energy expansion of order 4, and a potential energy expansion of order 8; using a kinetic energy expansion order of 6 requires a more expensive calculation where convergence is already observed to be within 0.001 cm\(^{-1}\) when expanding to fourth order. In the present study, we find that the convergence is more sensitive to \( P_{\text{max}} \) and we obtain convergence to within 0.1 cm\(^{-1}\) when \( P_{\text{max}} \) is 12 (see Table 1), we therefore use a basis set based on this value.

As well as using the polyad number to truncate the size of the basis, we employ a further truncation technique by specifying an upper limit for the eigenvalue calculations, i.e. the basis set is constructed such that it is approximately associated with energy values up to a limit of \( E_{\text{max}} \). This is based on an estimate given by summing eigenvalues of our 1D basis functions before they are considered for matrix element calculations; the active space is constructed using basis functions whose eigenvalues sum together to have \( E \leq E_{\text{max}} \). For the present study we use \( E_{\text{max}}/\hbar c = 10,000 \text{ cm}^{-1} \).

These precautions are particularly important for the \( \text{SO}_3 \) molecule, as its larger mass (compared to, for example, \( \text{XH}_3 \) systems) gives rise to small rotational constants, which in turn requires calculations up to a high \( J \) value to ensure adequate coverage of transitions for a given temperature. This means that use of any unnecessary basis functions will prove computationally costly. In addition to this basis set minimisation, we can reduce the size of the Hamiltonian further by making use of the group theory. \( \text{SO}_3 \) belongs to the \( D_{3h}(M) \) molecular group symmetry, which spans six irreducible representations. The nuclear spin-0 Bosons of the constituent atoms allow only the molecular ro-vibrational wavefunctions of the \( A'_1 \) and \( A'_1 \) representations in order to satisfy the Pauli principle. This reduces both the number of Hamiltonian matrices we need to consider and, since \( E \) symmetry Hamiltonian matrices are larger, their size.

The two other factors which are important in our spectral calculations are (i) the wavenumber range of the desired synthetic spectrum and (ii) the temperature at which we wish to simulate it. The quality of a computed spectrum will become

<table>
<thead>
<tr>
<th>( J )</th>
<th>( K )</th>
<th>Obs.</th>
<th>Calc. (A)</th>
<th>Calc. (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0</td>
<td>2.0912</td>
<td>2.0826</td>
<td>2.0916</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>2.6115</td>
<td>2.6007</td>
<td>2.6119</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>6.9707</td>
<td>6.9419</td>
<td>6.9718</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>5.3998</td>
<td>5.3775</td>
<td>5.4006</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>8.8852</td>
<td>8.8484</td>
<td>8.8864</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>13.0675</td>
<td>13.0134</td>
<td>13.0694</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>18.1106</td>
<td>18.0737</td>
<td>18.1132</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>29.7958</td>
<td>29.6725</td>
<td>29.8002</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>25.0838</td>
<td>24.9800</td>
<td>25.0874</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>17.2297</td>
<td>17.1583</td>
<td>17.2319</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>38.3360</td>
<td>38.1774</td>
<td>38.3417</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>36.7655</td>
<td>36.6134</td>
<td>36.7709</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>32.0539</td>
<td>31.9211</td>
<td>32.0584</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>146.3334</td>
<td>145.7283</td>
<td>146.3548</td>
</tr>
<tr>
<td>30</td>
<td>3</td>
<td>144.7645</td>
<td>144.1644</td>
<td>144.7856</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>140.0574</td>
<td>139.4781</td>
<td>140.0777</td>
</tr>
<tr>
<td>20</td>
<td>9</td>
<td>132.2114</td>
<td>131.6644</td>
<td>132.2304</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>121.2253</td>
<td>120.7236</td>
<td>121.2425</td>
</tr>
<tr>
<td>20</td>
<td>15</td>
<td>107.0973</td>
<td>106.5538</td>
<td>107.1122</td>
</tr>
<tr>
<td>80</td>
<td>18</td>
<td>89.8252</td>
<td>89.4529</td>
<td>89.8372</td>
</tr>
<tr>
<td>80</td>
<td>78</td>
<td>1195.6590</td>
<td>1190.7160</td>
<td>1195.8090</td>
</tr>
</tbody>
</table>

Table 2: Rotational term values (cm\(^{-1}\)) of \( ^{32}\text{S}^{16}\text{O}_3 \) in the ground vibrational state: the theoretical term values were obtained using the \( \text{ab initio} \) (A) and morphed (M) PESs; the experimentally derived values (Obs.) are from ref. 8.
sensitive to $E_{\text{max}}$ as the temperature increases; we need to ensure that we calculate all energy states that are significantly populated at the given temperature. This can be checked using the temperature-dependent partition function:

$$Q(T) = \sum_i g_i \exp(-E_i/kT),$$

where $g_i$ is the total degeneracy of the ro-vibrational state $i$ with energy $E_i$, with the sum running over all energies at the absolute temperature $T$, and $k$ is Boltzmann’s constant. The total degeneracy is given by $(2j + 1)$ times the nuclear spin degeneracy, which for the present case of $^{32}\text{S}^{16}\text{O}_3$ is simply 1 for both the $A_1'$ and $A_1''$ symmetries, given that the nuclear spin of $^{16}\text{O}$ and $^{32}\text{S}$ is zero. For a given temperature, we can determine the contribution of various states to the value of $Q$ and then check that $Q$ converges to a specific value as $T$ tends to infinity; as $T$ increases we require a greater coverage of higher-lying energy states. At $T = 298.15$ K we find that $Q$ converges to better than 1% at $j = 85$, with a value of $Q = 8089.3$. Therefore calculations spanning all $J$’s up to 85 should be sufficient for simulating spectra at this temperature. Fig. 1 shows the value of $Q$ as a function of all energy levels having the $J$ quantum number up to a maximum value, $J_{\text{max}}$, at an absolute temperature of $T = 298.15$ K. As energies are included in the summation for increasing values of $J$, we see that the associated energy levels contribute less and less to the value of $Q$, until it converges to a limit.

An ab initio quartic force field was previously published by Martin based on coupled cluster methods. As an initial test of the TROVE procedure we used the force field given by Martin to compute fundamental term values using a well-converged polyad truncation scheme of $P_{\text{max}} = 16$. We found discrepancies between our results and the values published, particularly the value of the $v_2$ fundamental term value. We suspected that Marin’s $F_{22}$ value has a problem and substituted it with a scaled value from a previously published force-field; this results in a significant improvement. Table 3 illustrates the differences between the values and our final converged result for the fundamental vibrations, and we see that we converge to an improved degree using our morphed PES.

### B Intensity simulations

The simulation of absorption spectra can be broken down into two main parts: first, the eigenfunctions and eigenvalues of the numerically constructed Hamiltonian matrix are calculated using a diagonalisation procedure; second, these eigenfunctions are used to compute transition dipoles, line strengths $S$, Einstein $A$ coefficients and intensities $I$ for allowed transitions. For $^{32}\text{S}^{16}\text{O}_3$, the rigorous selection rules determining allowed electric dipole transitions are $|J_i' - J_f'| \neq 0, \pm 1$ ($J_i' + J_f' \geq 1$), and symmetry selection rules $A_i' \leftrightarrow A_i''$. The intensity of a transition between two states is given by the formula

$$I(f \rightarrow i) = \frac{8\pi^3 N_{\text{A}} c}{(4\pi\epsilon_0)^3} \frac{E_i/kT}{Q(T)} \left[ 1 - \exp\left(-\frac{\hbar c v_{\text{lab}}}{kT}\right) \right] S(f \rightarrow i),$$

where $I(f \rightarrow i)$ is the transition intensity for a transition from state $i$ with energy $E_i$ to state $f$ with energy $E_f$ with $\hbar c v_{\text{lab}} = E_f - E_i$, $Q$ is the partition function. $S(f \rightarrow i)$ is the line strength, which is defined by the following integration:

$$S(f \rightarrow i) = g_{\text{ ns}} \sum_{m_l,m_s} \sum_{X,Y,Z} |\langle \Phi_{f i}^{(l)} | \mu_{A} | \Phi_{f i}^{(l)} \rangle |^2$$

for a transition between the initial state $i$ with rovibrational wavefunction $\Phi_{f i}^{(l)}$ and the final state $f$ with wavefunction $\Phi_{f i}^{(l)}$. Here, $g_{\text{ ns}}$ is the nuclear spin statistical weight factor, and $\mu_{A}$ is the electronically averaged component of the molecular dipole moment along the space-fixed axis $A = X, Y, Z$. The sum also runs over the quantum numbers $m_l$ and $m_s$ which are projections of the total angular momentum $J$ (in units of $\hbar$) on the laboratory fixed $Z$ axis, for the initial and final states, respectively.

Maki et al. reported extensive high-resolution studies of a number of fundamental, combination and overtone bands of $^{32}\text{S}^{16}\text{O}_3$. Their principle aim was to obtain accurate wavenumber measurements, but relative intensities were also measured. In the present work we convert these data into absolute intensities by normalizing to the theoretical intensities obtained using TROVE at $T = 298.15$ K as described below.

The measurements available to us cover three spectral regions: 405-708 cm$^{-1}$ (focusing on the $v_3$, $v_4$, $2v_2 - v_3$, $v_2 + v_4 - v_2$, $v_2 + v_4 - v_4$, $2v_4^{(l=0)} - v_4$ and $2v_4^{(l=2)} - v_4$ bands), 1200-1680 cm$^{-1}$ ($v_3$), and 2500-3280 cm$^{-1}$ ($2v_3^{(l=2)}$). These measurements were each made at different values of pressure.

### Table 3 A comparison of the fundamental term values (cm$^{-1}$) computed with TROVE employing the new morphed PES (‘this work’), the original (‘Martin’) force constants$^a$ and the $F_{22}$-corrected$^d$ (‘Martin $F_{22}$’), see the text

<table>
<thead>
<tr>
<th>Obs.</th>
<th>Martin</th>
<th>Martin $F_{22}$</th>
<th>This work</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu_1$</td>
<td>1064.92</td>
<td>1063.36</td>
<td>1064.22</td>
</tr>
<tr>
<td>$\nu_2$</td>
<td>497.57</td>
<td>428.36</td>
<td>487.10</td>
</tr>
<tr>
<td>$\nu_3$</td>
<td>1391.52</td>
<td>1386.81</td>
<td>1386.85</td>
</tr>
<tr>
<td>$\nu_4$</td>
<td>530.09</td>
<td>527.35</td>
<td>527.32</td>
</tr>
</tbody>
</table>

$^a$ Martin’s force constants, with a substituted $F_{22}$ value from ref. 13.
Table 4 Comparison of calculated (TROVE) and experimental band centers and numbers of line transitions

<table>
<thead>
<tr>
<th>Band</th>
<th>Obs.</th>
<th>Calc.</th>
<th>$P_1$</th>
<th>$P_2$</th>
<th>TROVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu_2 - \nu_0$</td>
<td>497.57</td>
<td>498.48</td>
<td>773</td>
<td>1265</td>
<td>5422</td>
</tr>
<tr>
<td>$\nu_4 - \nu_0$</td>
<td>530.09</td>
<td>528.59</td>
<td>996</td>
<td>2052</td>
<td>12195</td>
</tr>
<tr>
<td>$\nu_2 + \nu_4 - \nu_2$</td>
<td>534.83</td>
<td>537.36</td>
<td>0</td>
<td>69</td>
<td>15147</td>
</tr>
<tr>
<td>$2\nu_2 - \nu_4$</td>
<td>497.45</td>
<td>496.88</td>
<td>112</td>
<td>704</td>
<td>7171</td>
</tr>
<tr>
<td>$\nu_2 + \nu_4 - \nu_4$</td>
<td>497.81</td>
<td>498.76</td>
<td>47</td>
<td>602</td>
<td>27182</td>
</tr>
<tr>
<td>$2\nu_2((l=0) - \nu_4$</td>
<td>530.36</td>
<td>528.79</td>
<td>116</td>
<td>775</td>
<td>31096</td>
</tr>
<tr>
<td>$\nu_5 - \nu_0$</td>
<td>529.72</td>
<td>527.91</td>
<td>39</td>
<td>455</td>
<td>13718</td>
</tr>
<tr>
<td>$\nu_1 - \nu_0$</td>
<td>1391.52</td>
<td>1387.45</td>
<td>2014</td>
<td>—</td>
<td>14441</td>
</tr>
<tr>
<td>$2\nu_1((l=2) - \nu_0$</td>
<td>2777.87</td>
<td>2770.29</td>
<td>1527</td>
<td>—</td>
<td>18115</td>
</tr>
</tbody>
</table>

$^a$ For the 405–708 cm$^{-1}$ region, measurements were made at two pressures; $P_1 = 0.409$ Torr, and $P_2 = 2.04$ Torr. For the $\nu_2$ and $2\nu_2$ bands only the lower pressure measurements at 0.16 Torr were used here. Pressure values are not recorded for microwave measurements.$^{12}$

The latter numbers are the subject of the following selection criteria: $J \leq 85$, intensity cut-off, $I_{\text{calc}} > 10^{-34}$ cm$^2$ per molecule, and the wavenumber window, 0–4000 cm$^{-1}$. Experimental lines with negative relative intensities were left out of the analysis.

To normalize the experimental intensities, the experimental relative data obtained from each spectral window and at each different pressure were scaled to match the theoretical values computed at $T = 298.15$ K. The scaling factors obtained through a minimization procedure using all selected experimental lines with non-zero intensity (see Table 5) are $6.571 	imes 10^{-21}$ cm$^2$ per molecule (405–708 cm$^{-1}$, 0.409 Torr), $1.838 	imes 10^{-21}$ cm$^2$ per molecule (405–708 cm$^{-1}$, 2.04 Torr), $4.823 	imes 10^{-20}$ cm$^2$ per molecule (1200–1680 cm$^{-1}$, 0.16 Torr), $1.328 	imes 10^{-21}$ cm$^2$ per molecule (2500–3280 cm$^{-1}$, 4.99 Torr). With these factors the ‘experimental’ intensities match the theoretical values reasonably well, for example for 405–708 cm$^{-1}$ the agreement is within about $7.02 \times 10^{-22}$ cm$^2$ per molecule and 5.05 $\times 10^{-22}$ cm$^2$ per molecule, at 0.409 and 2.04 Torr, respectively.

Having the absolute intensities derived, band intensities were estimated for each experimental band as the sum of individual line intensities. In Table 5 these ‘experimental’ band intensities $S_{\text{exp}}$ (cm per molecule) are compared to the theoretical values obtained by summing intensities $S_{\text{calc}}$ (a) from all TROVE lines from a given window and (b) only from lines with experimental counterparts present. This was done separately for each spectral range, and each measurement pressure therein. In Table 5 these quantities are referenced to as $S_{\text{tot}}^\text{calc}$ and $S_{\text{red}}^\text{calc}$ for the ‘total’ and ‘reduced’ band intensities, respectively, and compared to $S_{\text{exp}}$. The ratio $S_{\text{red}}^\text{calc}/S_{\text{exp}}$ also shown in Table 5 demonstrates the good quality of the scaling procedure employed as well as that of our dipole moment. For example at 0.409 Torr, the differences between ‘experimental’ and theoretical band intensities are within about 20% for all bands from the 405–708 cm$^{-1}$ region with the exception of $2\nu_2((l=2) - \nu_4$ (see also discussion below). It should be stressed here that only one scaling factor for all eight bands from this window was applied at a given pressure. The difference between the two theoretical band intensities $S_{\text{tot}}^\text{calc}$ and $S_{\text{red}}^\text{calc}$ gives a measure of the missing experimental transition data. According to Table 5 the measurements for even stronger bands are missing more than 50% of the total intensity.

Table 5 also shows theoretical values of vibrational transition moments defined as

$$\bar{\mu} = \sqrt{\bar{\mu}_x^2 + \bar{\mu}_y^2 + \bar{\mu}_z^2},$$ (13)

where

$$\mu_x = \langle \Psi_\text{ vib} | \mu_x | \Psi_\text{ vib} \rangle,$$ (14)

and $\Psi_\text{ vib}$ and $\Psi_\text{ vib}^{(i)}$ are the vibrational eigenfunctions of the ‘initial’ and ‘final’ states, respectively, variationally computed

Table 5 Vibrational band intensities $S_{\text{exp}}, S_{\text{calc}}, S_{\text{red}}^\text{calc}$ in cm$^2$ per molecule $\times 10^{-18}$, and calculated transition moments $\bar{\mu}_D$ in Debye. $P_1$ and $P_2$ refer to the different pressure measurements within the same wavenumber region (see Table 4). $S_{\text{tot}}^\text{calc}/S_{\text{exp}}$ is the ratio of the theoretical reduced and total band intensities (see the text). $S_{\text{red}}^\text{calc}$ is the theoretical band intensity computed by summing all TROVE lines. $S_{\text{exp}}$ is the experimental band intensity obtained from a summation over all experimental values after scaling factors are applied (see the text). $S_{\text{red}}^\text{calc}$ is the theoretical band intensity computed using only lines for which experimental counterparts exist. $N_{\text{red}}$ is the number of matched lines.

<table>
<thead>
<tr>
<th>Band</th>
<th>$S_{\text{exp}}$</th>
<th>$S_{\text{calc}}$</th>
<th>$S_{\text{red}}^\text{calc}$</th>
<th>$S_{\text{red}}^\text{calc}/S_{\text{exp}}$</th>
<th>$N_{\text{red}}$</th>
<th>$\bar{\mu}_D$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu_2$</td>
<td>2.987</td>
<td>1.537</td>
<td>3.705</td>
<td>2.915</td>
<td>1.559</td>
<td>0.98</td>
</tr>
<tr>
<td>$\nu_4$</td>
<td>4.258</td>
<td>3.112</td>
<td>5.949</td>
<td>4.310</td>
<td>3.149</td>
<td>1.01</td>
</tr>
<tr>
<td>$2\nu_2 - \nu_4$</td>
<td>0.116</td>
<td>0.470</td>
<td>0.661</td>
<td>0.101</td>
<td>0.411</td>
<td>0.88</td>
</tr>
<tr>
<td>$\nu_2 + \nu_4 - \nu_2$</td>
<td>0.062</td>
<td>0.322</td>
<td>0.528</td>
<td>0.052</td>
<td>0.251</td>
<td>0.84</td>
</tr>
<tr>
<td>$\nu_2 + \nu_4 - \nu_4$</td>
<td>0.026</td>
<td>0.260</td>
<td>0.581</td>
<td>0.022</td>
<td>0.215</td>
<td>0.84</td>
</tr>
<tr>
<td>$2\nu_2((l=2) - \nu_4$</td>
<td>0.112</td>
<td>0.589</td>
<td>0.873</td>
<td>0.102</td>
<td>0.485</td>
<td>0.91</td>
</tr>
<tr>
<td>$2\nu_2((l=0) - \nu_4$</td>
<td>0.026</td>
<td>0.222</td>
<td>0.405</td>
<td>0.015</td>
<td>0.81</td>
<td>0.57</td>
</tr>
<tr>
<td>$\nu_5 - \nu_0$</td>
<td>—</td>
<td>0.009</td>
<td>0.101</td>
<td>—</td>
<td>0.003</td>
<td>—</td>
</tr>
<tr>
<td>$\nu_1 - \nu_0$</td>
<td>39.490</td>
<td>—</td>
<td>44.440</td>
<td>39.490</td>
<td>—</td>
<td>2014</td>
</tr>
<tr>
<td>$2\nu_4$</td>
<td>0.093</td>
<td>—</td>
<td>0.119</td>
<td>0.093</td>
<td>—</td>
<td>1527</td>
</tr>
</tbody>
</table>
using TROVE and $\mu_x$ is the component of the molecular dipole moment along the molecular-fixed axis $x = x, y, z$.

Fig. 2 presents an overview of the simulated absorption ($T = 298.15$ K) using TROVE and experimental absorption spectra of $^{32}\text{S}^{16}\text{O}_3$ for the whole simulation range up to 4000 cm$^{-1}$. It reveals the gaps and limitations of the available experimental data. Fig. 3 shows the ‘forbidden’ rotational band as a stick spectrum. It should be noted that the microwave measurements of ref. 12 do not report any intensities. In Fig. 4 all eight bands from the 405–708 cm$^{-1}$ region are combined into one graph for each pressure to illustrate the quality of the corresponding experimental data. This figure suggests that the 0.409 Torr data are generally more reliable. This is also reflected by the ratio values $S_{\text{calc}}/S_{\text{exp}}$ from Table 5, which are significantly closer to 1 at the lower pressure. For data obtained at two pressures we therefore consider scaled intensity values obtained at the lower pressure to be more reliable. Finally, Fig. 5 presents a detailed comparison for all the bands from the three spectral regions studied in this work in the form of stick diagrams. Our intensities based on the $ab\text{ initio}$ DMS are in very good qualitative agreement with the experiment.

**IV Conclusions and discussions**

The shifted theoretical bands in Fig. 4 and 5 indicate that our morphed PES of SO$_3$ requires further improvement (see also band centers in Table 4). The theoretical $\nu_2$ frequencies have an rms deviation of 0.91 cm$^{-1}$ when compared with the experimental data, which is relatively small compared to the deviation for $\nu_3$ of 4.07 cm$^{-1}$. This is to be expected since the underlying $ab\text{ initio}$ PES was computed at a modest level of theory. We are planning to refine this surface by fitting to all experimental data available.

Table 5 outlines the quality of the intensity scaling procedure, in which the relative values of the experimental intensities were converted to absolute values (cm per molecule). For the 0.409 Torr and 2.04 Torr measurements our comparisons mostly agree to within 20%, with the exception of the $2\nu_4(\ell_4 = 0) - \nu_4$ band measured at 0.409 Torr which shows nearly a 50% difference, and the $\nu_1 - \nu_4$ band measured at 2.04 Torr with 80% uncertainty. The latter can probably be attributed to both the small number of lines in each case, and residual errors in the transition dipole. In the case of the $2\nu_4(\ell_4 = 0) - \nu_4$ band the comparison with the 2.04 Torr experiment yields a better value for $S_{\text{calc}}/S_{\text{exp}}$ than for the 0.409 Torr measurement, which suggests that the number of lines available at 0.409 Torr is too low (see also Fig. 4). The significance of the results presented in Table 5 and illustrated in Fig. 4 is that these give an estimation on the quality of the $ab\text{ initio}$ dipole moment surface as well as of the experimental data. Based on these numbers we can place a lower estimate on the quality uncertainty for our intensities for each band, for example the experiment and theory for the $\nu_2$ and $\nu_4$ bands agree at least to within 3%.
for both pressure measurements, while it is only 13% for $2\nu_2 - \nu_2$, and about 18% for the remaining bands.

Our complete room-temperature line list for SO$_3$ containing 174 674 257 transitions can be accessed online at www.exomol.com in the ExoMol format described in ref. 39. It includes the transition energies, Einstein coefficients $A(f \leftarrow i)$, and absorption intensities estimated for $T = 298.15$ K. Additionally, a list of 10 878 experimental transitions with absolute intensities obtained for $T = 298.15$ K is included in the ESI† part of this paper in a form suitable for standard atmospheric and planetary spectroscopic databases.

Our future work will be focused on the development of a hot line list for $^{32}$S$^{16}$O$_3$ for high temperature industrial applications as well as for modelling molecular opacity in atmospheres of (exo-)planets and cool stars as part of the ExoMol project (see www.exomol.com).

Acknowledgements

We thank Alexander Fateev for stimulating our interest in this molecule and for many helpful discussions, and Jeff Barber for supplying his experimental result. This work was supported by grant 10442 from energinet.dk under a subcontract from the Danish Technical University and the ERC under Advanced Investigator Project 267219.

References

39. M. line lists for the opacity of exoplanets and other atmospheres, Eighth International Conference on Atomic and Molecular Data and Their Applications: ICAMDATA-8, 2012.