Supplementary Information

Toward standardization of Raman Spectroscopy: Accurate wavenumber and intensity calibration using rotational Raman spectra of H_2, HD, D_2 and vibration-rotation spectrum of O_2

Ankit Raj,[1] Chihiro Kato,[2] Henryk A. Witek,[1,3,*] and Hiro-o Hamaguchi[1,3,†]

[1] *Department of Applied Chemistry and Institute of Molecular Science, National Chiao Tung University, Hsinchu 30010, Taiwan*

[2] *Kanagawa Institute of Industrial Science and Technology (KISTEC), Ebina 243-0435, Kanagawa, Japan*

[3] *Center for Emergent Functional Matter Science, National Chiao Tung University, Hsinchu 30010, Taiwan*

* hwitek@mail.nctu.edu.tw

† hhama@nctu.edu.tw

Contents

List of Figures

List of Tables

List of symbols

γ : polarizability anisotropy, $\gamma = \alpha_{\parallel} - \alpha_{\perp}$

σ : Standard deviation

Sec 1. Description of correction for obtaining spectra as photons/wavenumber

Photon counting measurements with CCDs coupled with dispersive spectrographs typically produce linearly changing wavelength over the horizontal pixels. Thus, photons are sampled over a non-linear wavenumber scale. Correction to obtain the correct number of photons per wavenumber when measurement is made over an approximately linear wavelength scale is discussed below. Similar corrections have been reported earlier with reference to electronic spectra.[1–4]

$$\nu = \frac{1}{\lambda} \tag{1}$$

$$\delta\nu = -\frac{1}{\lambda^2}\ \delta\lambda \tag{2}$$

$$\frac{\delta\nu_s}{\delta\nu_0} = \frac{\frac{1}{\lambda_s^2}\ \delta\lambda_s}{\frac{1}{\lambda_0^2}\ \delta\lambda_0} \tag{3}$$

$$\frac{\delta\nu_s}{\delta\nu_0} = \left(\frac{\nu_s}{\nu_0}\right)^2 \left(\frac{\delta\lambda_s}{\delta\lambda_0}\right) \tag{4}$$

where ν_0 is some reference wavenumber and ν_s is any other wavenumber in the spectra. The factor $(\frac{\nu_s}{\nu_0})^2$, corrects for the number of photons of the spectra measured across a linear wavelength scale. The term $(\frac{\delta\lambda_s}{\delta\lambda_0})$ is the small correction arising from the non-linearity of the wavelength over the pixels. In the present work, ν_0 has been chosen as 18789.9 cm^{-1} (in absolute wavenumbers which corresponds to 0 cm^{-1} in relative wavenumbers). Subsequently, $\delta\lambda_s$ is taken as the forward wavelength spacing across the complete wavelength range and $\delta\lambda_0$ is taken as the corresponding wavelength spacing for the laser wavelength.

$\frac{\delta\nu_s}{\delta\nu_0}$ can also be computed directly using the wavenumber step (or spacing) in the wavenumber array representing the x-axis, in a similar way as described for $(\frac{\delta\lambda_s}{\delta\lambda_0})$.

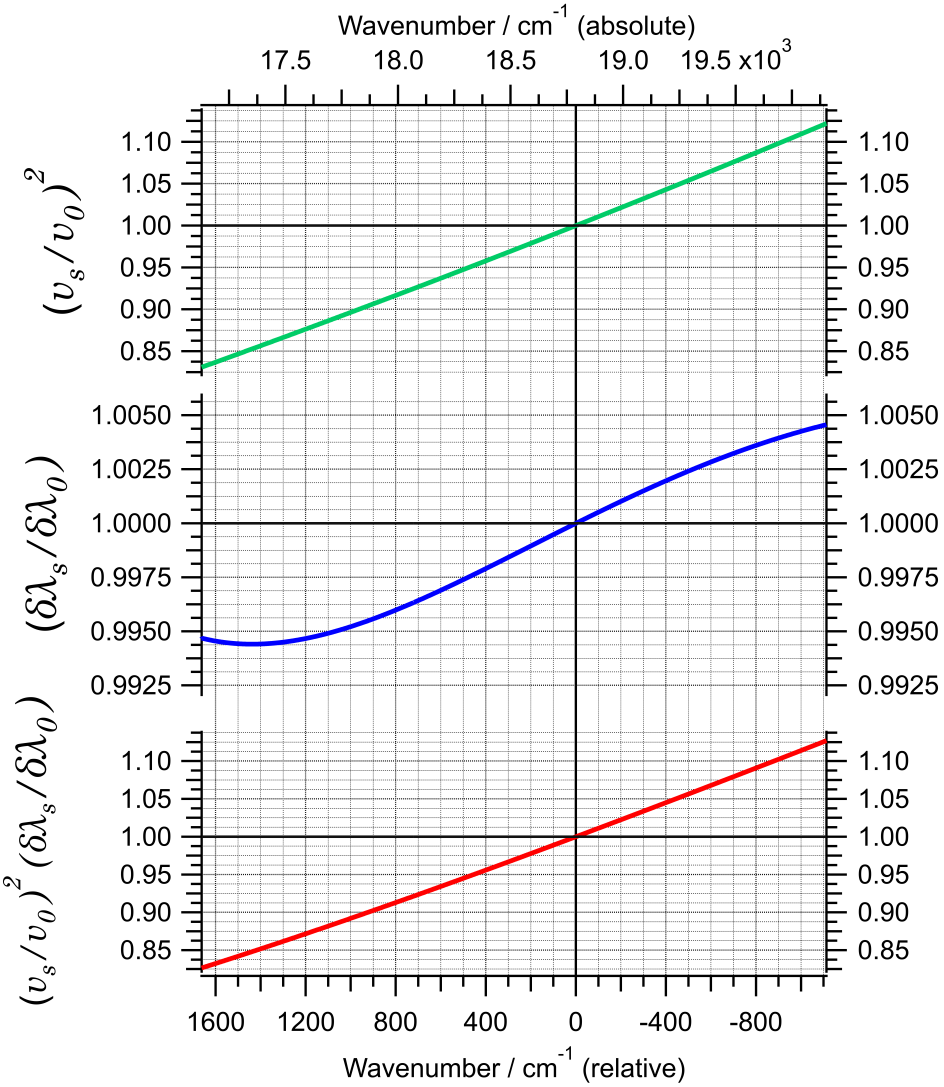


FIG. S1: Correction to obtain photons/wavenumber from spectra measured over photons/wavelength. The first term, $(\frac{\nu_s}{\nu_0})^2$ on the right hand side of Eqn. (4) is plotted on the top (in green). The second term $(\frac{\delta\lambda_s}{\delta\lambda_0})$ is plotted in the center in blue. The first term, in green, is large compared to the second term by as much as 30 times. The total correction, $\frac{\nu_s}{\nu_0} = (\frac{\nu_s}{\nu_0})^2 \frac{\delta\lambda_s}{\delta\lambda_0}$ is plotted in the bottom in red color. The above plots are specific to our spectrograph and grating position.

Sec 2. Concept of the fitting scheme for intensity calibration

Consider a pair of (x, y) data, described as $y = f(x)$, which is perturbed in terms of the y-values by a perturbation defined as $p(x)$ to produce say $c(x)$. The perturbation $p(x)$ is equivalent to the wavelength dependent sensitivity of the spectrometer. The objective is to determine $p(x)$ when the y-values for $f(x)$ and $c(x)$ are known for sparse pairs of x-points.

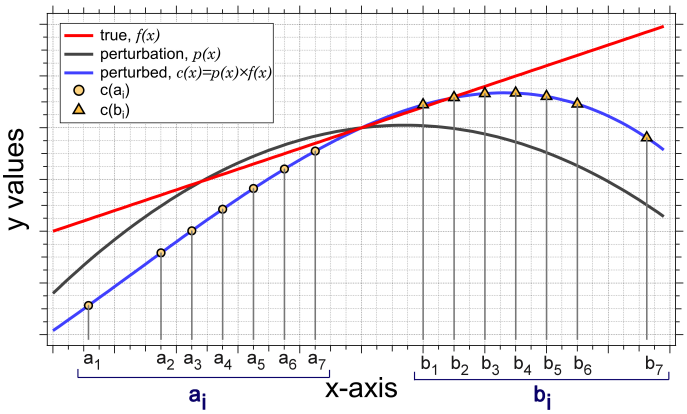


FIG. S2: Plot describing the concept of the fitting used in the present work. Ratio of y-values at pair of x-points (a_i, b_i), in the true, $f(x)$ and perturbed functions, $c(x)$ are used to determine the perturbation, $p(x)$ modeled as a polynomial. The choice of x-positions, a_i and b_i is arbitrary and these can be randomly located.

The ratio of the y-values at two x-points (a_i, b_i) in the true data, $f(x)$ expressed as $R^{true}_{a_i,b_i}$ is affected by the perturbation to yield $R^{pertb.}_{a_i,b_i}$. This change is directly related to the perturbation function by the ratio of $p(x)$ at a_i and b_i which can be given as $R^{p(x)}_{a_i,b_i}$.

$$\frac{R^{pertb.}_{a_i,b_i}}{R^{true}_{a_i,b_i}} = R^{p(x)}_{a_i,b_i} \tag{5}$$

where $R^{pertb.}_{a_i,b_i}$=$c(a_i)/c(b_i)$ and $R^{true}_{a_i,b_i}$=$f(a_i)/f(b_i)$. The perturbation, $p(x)$ can be modeled as a polynomial function with coefficients k_n to obtain,

$$\frac{R^{pertb.}_{a_i,b_i}}{R^{true}_{a_i,b_i}} = \frac{1 + k_1 a_i + k_2 a_i^2 + k_3 a_i^3 + ...}{1 + k_1 b_i + k_2 b_i^2 + k_3 b_i^3 + ...} \tag{6}$$

The difference of the LHS to RHS in Eqn. (6) can be used in a least squares minimization procedure to obtain the coefficients k_n.

$$\text{Residual} = \sum_i (\text{LHS} - \text{RHS})^2 \tag{7}$$

$$\textbf{Min}(\text{ Residual }) : k_n \text{ as fit parameters} \tag{8}$$

Error in the LHS of Eqn. (6) is given as

$$\sigma_{\text{LHS}_i} = \frac{c(a_i)}{c(b_i)} \sqrt{\left(\frac{\sigma_{c(a_i)}}{c(a_i)}\right)^2 + \left(\frac{\sigma_{c(b_i)}}{c(b_i)}\right)^2} \tag{9}$$

where $c(a_i)$ and $c(b_i)$ are the values of $c(x)$ at a_i and b_i, respectively. Similarly, $\sigma_{c(a_i)}$ and $\sigma_{c(b_i)}$ represent the error in the values of $c(x)$ at a_i and b_i, respectively. Weighted non-linear fitting can be performed using the weight, $w_i = 1/\sigma^2_{\text{LHS}_i}$ in

$$\text{Residual} = \sum_i w_i (\text{LHS} - \text{RHS})^2 \tag{10}$$

where LHS and RHS refer to the terms in Eqn. (6).

A. Implementation and benchmarking

The above procedure was implemented in an Igor[5] procedure and Python module. The Igor procedure uses the inbuilt function '*optimize*' while Python module uses the '*optimize.least_squares*' or '*optimize.minimize*' module from the SciPy library[6–8]. See Examples in the GitHub repository[9] for the computer programs.

The programs were tested with artificial function $f(x)$; known perturbation, $p(x)$ as quadratic, cubic and quartic functions and the final data, $c(x) = f(x)p(x) + noise$. Noise vectors of different magnitudes having normal distribution were checked to understand the effect of noise on the obtained fitting results.

Weighted non-linear fit was performed with 500 random noise vectors and several sets of initial guesses to check the expected performance with actual data. Scaling was applied to the coefficients to ensure numerical accuracy. Details about the test are given below.

```
x=[-10, 10]
f(x) = a + bx, a=0.96, b=0.035

p(x): quadratic = 1 + (k1/1e2)*x + (k2/1e2)*x^{2}
               k1=1.25, k2=-1.05
      cubic = 1 + (k1/1e2)*x + (k2/1e2)*x^{2} + (k3/1e3)*x^{3}
                 k1=1.25, k2=-1.05, k3=-1.1
      quartic = 1 + (k1/1e2)*x + (k2/1e2)*x^{2} + (k3/1e3)*x^{3} + (k4/1e4)*x^{4}
                 k1=1.25, k2=-1.05, k3=-1.1, k4=1.0

c(x) = f(x)*p(x) + noise
     noise = 0.1 to 15% of the y-value of c(x) for x=-10
     500 random noise vectors used

sampled points : 7 pairs of values from c(x), randomly chosen
                              a_{i}=[0.552763, 1.557788, 2.562814, 4.572864, 5.075376, 7.58793, 9.296482]
                              b_{i}=[-0.452261, -1.457286, -2.462311, -3.467336, -4.974874, -5.97989, -7.989949]

Tolerances in the fit coefficients : absolute error = 1e-9
                                     absolute error in residual function value = 1e-9

Initial guess : quadratic : (k1 = 1.0, k2 = -0.9),
                            (k1 = 1.4, k2 = 0.1)
                cubic : (k1 = 1.0, k2 = -0.9, k3 = -0.75),
                        (k1 = 1.4, k2 = 0.1, k3 = 0.2)
                quartic : (k1 = 1.0, k2 = -0.9, k3 = -0.75, k4 = 1.2),
                          (k1 = 1.4, k2 = 0.1, k3 = 0.2, k4 = -0.85)
```

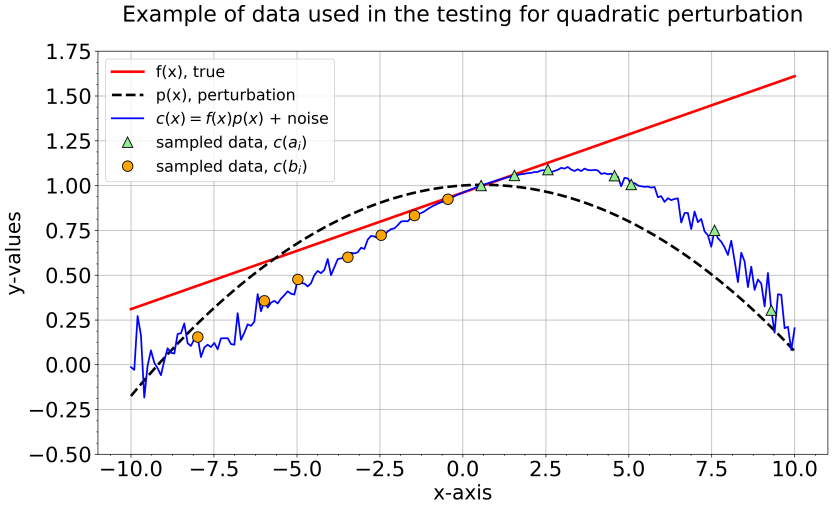


FIG. S3: Plot showing the artificial data used in the testing of the fitting. Perturbation introduced is a quadratic function. The noise included in the data has larger magnitude near the edges (-10 and +10) and this trend is somewhat analogous to the experimental data for rotational Raman intensities analyzed in the present work.

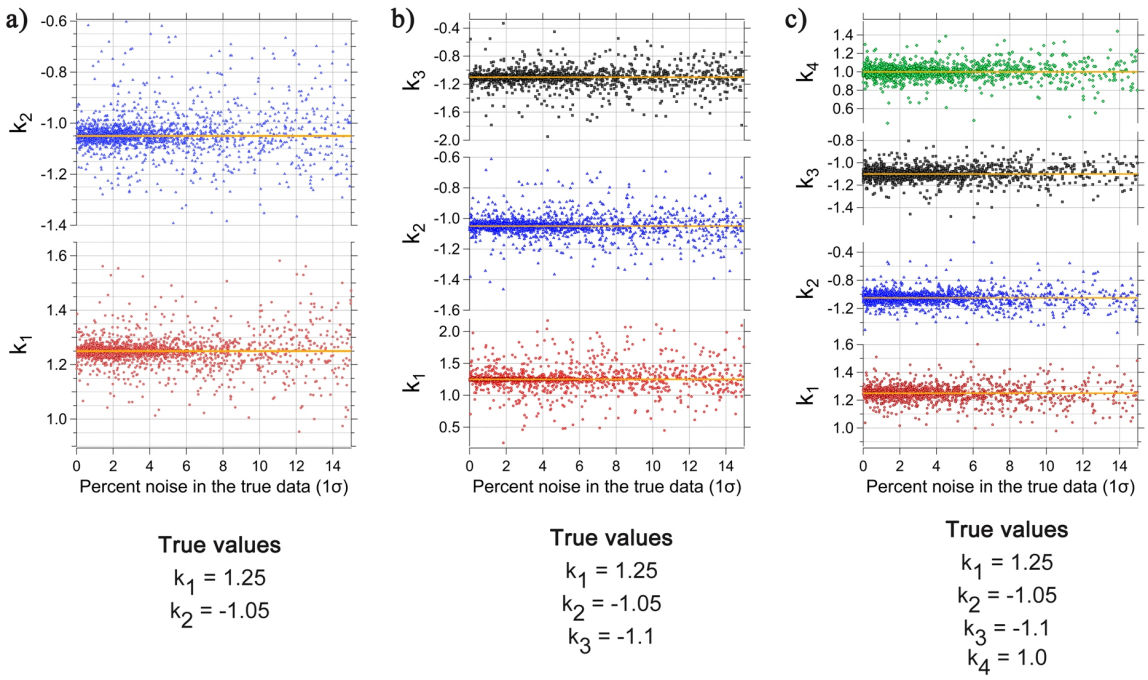


FIG. S4: Plot showing the coefficients from the fitting procedure for randomly generated noise vectors in the case of, a) quadratic , b) cubic and c) quartic perturbation functions, tested with artificial data (details given on Page 5). True values are shown in the plot using horizontal lines in dark yellow color and are also listed in the bottom.

Results

Benchmarking results are summarized in Fig. S4. These results show that for the artificial dataset tested, converged fit coefficients with smaller deviation from the true value are obtained for error <6–8%. These coefficients show a smaller spread around the true values (shown in yellow horizontal line). For error >6–8% the converged results show large variation and significant spread around the true values. Using constraints on the numerical values of the residual, best fit result can be obtained. It is necessary to mention here that the scale of 6–8% error, below which good results are obtained, is specific to this data set. The results of fit will change depending upon the number of sampling points, their location, the initial guess and the convergence criteria.

Considerations regarding the fit

1. Too large and too small fit coefficients affect the numerical accuracy. Scaling of the fit coefficients is suggested so that magnitude of the coefficient is close to 1, which helps in reducing such issues.

2. Sequential probe for the coefficients of the mathematical function describing the perturbation (or the wavelength dependent sensitivity here) can be performed, where the results from quadratic polynomial model can be used as initial guess for cubic polynomial model, and so on. (For present case, see details in Table. S12 and Fig. S13)

3. Inclusion of weight proves useful when data has uncertainties (and hence when dealing with smaller band intensities from rotational Raman spectra).

4. Multiple initial guesses must be tested to check whether a real minimum is found. Numerical values of the residual must be checked to verify this.

Sec 3. Error analysis of band area from rotational Raman bands

Error in the band area of rotational Raman bands are required for determining the weight which are used in the fitting procedure. Analysis of the error was undertaken to include actual uncertainties in the measured intensities for computing the weight for the fitting used for determination of the second correction to intensity. Details of this analysis are presented below.

Actual error in the experimental band area were determined by measurement and analysis of a large set of rotational Raman spectra from H_2. The exposure of the measurements ranged from 0.5–1800 sec with 56 spectra each, to generate data set with varying S/N ratio. After removal of background, the band area were determined using two approaches : integration and Voigt function fit. Change of band area over the 56 spectra was used to determine the 3σ error shown as percent relative to the absolute band area, in red (for integration) and green (for fitting) in Fig. S5. Here, σ represents the corrected standard deviation for limited samples. For a comparison, the $\sqrt{I}$ where I=band area, as a measure of error was also determined and shown in Fig. S5 in blue (3σ) and grey (1σ).[1]

These results show that the actual error in the band area (red or green data points) is smaller than the error estimated using the straightforward square root scheme (in blue). For smaller band area, the actual error is greater than the one determined from the $\sqrt{I}$. Error from fitting and integration approaches are very close, although fitting appears to give slightly smaller error for smaller band areas. In the present work, the

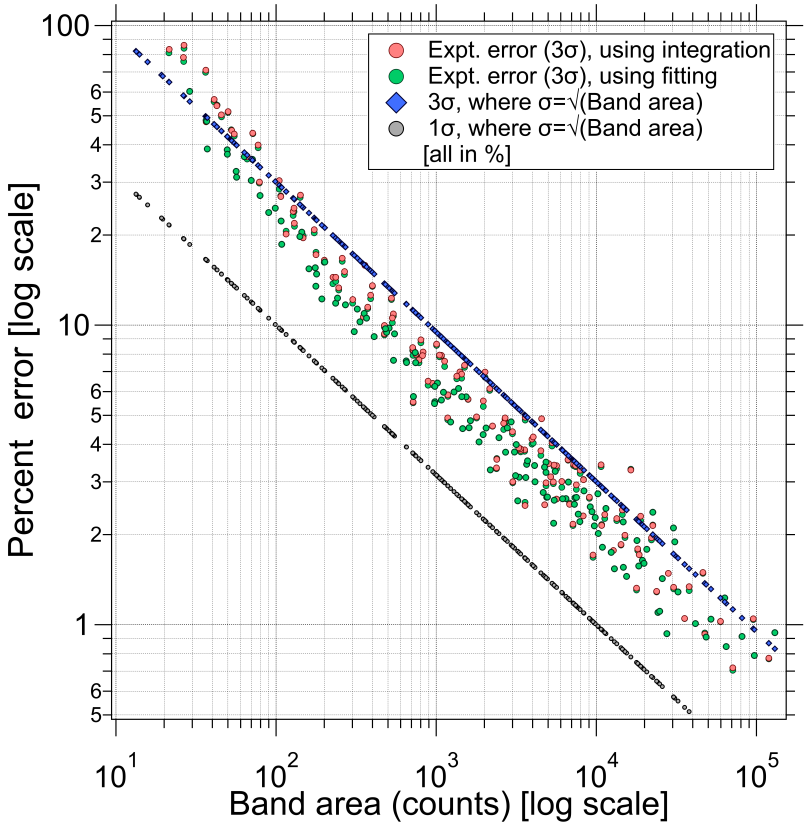


FIG. S5: Plot of experimental and estimated error over the magnitude of band area (both shown in log scales). The actual experimental error (3σ) is obtained from set of Raman spectra of H_2 measured with varying exposures each having 56 spectra. Band area was determined using integration and fitting, shown in red and green color. The usual estimates of error in photon measurements, $\sqrt{I}$ where I=band area are shown in blue (3σ) and grey (1σ). These were determined using the integrated band area. Actual experimental error is found to be always smaller than the estimates by few percent for larger band areas, whereas opposite trend is seen for smaller areas coming from weaker bands. *This set of result is specific for our spectrometer and largely depends on the CCD, its measurement parameters and the overall spectrometer alignment.

[1] 3σ covers 99.7% likelihood while for 1σ it is 68.2%.

red and green data points were fit to quadratic function under least squares minimization. The fit function and coefficients were used to determine the error in the experimental band area used for intensity calibration work.

Since pairs of bands are analyzed in this work, the total error in the band area ratios is given by, $\sigma_{total,i} = \frac{I_{J+2}}{I_{J-2}}\sqrt{(\frac{\sigma_{exp}}{I_{J+2}})^2 + (\frac{\sigma_{exp}}{I_{J-2}})^2}$. The preceding equation is Eqn. (7) from the main manuscript. In the fitting procedure, the errors are included by providing weights to data points from the respective rotational states. Weights used in the present work are shown in Fig.S6 for the 4 gas species and for each of the relevant rotational state used.

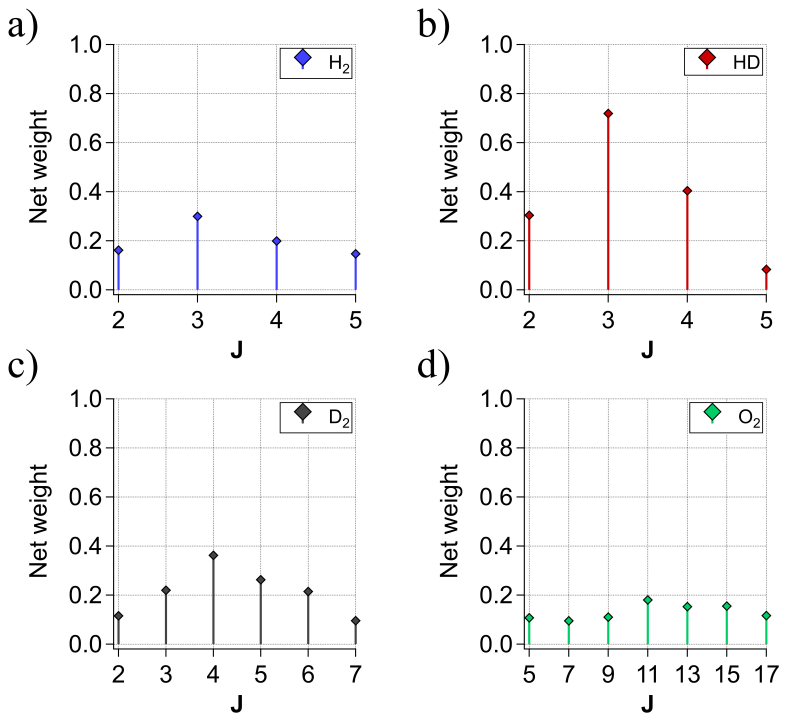


FIG. S6: Plot showing the weights applied to intensity ratios from different rotational states in the fitting analysis for determination of the final correction for intensity calibration.

Sec 4. Transitions wavenumbers and vibration-rotation correction factors for O_2

TABLE S7: Transition wavenumbers and vibration-rotation correction factors for bands originating from common rotational states, for pure rotation ($\Delta v = 0$) and the vibration-rotation transitions accompanying the fundamental ($\Delta v = 1$) for O_2 in the ground electronic state ($^3\Sigma_g^-$).

	$v' = v = 0$			$v' = 1, v = 0$		
J_i	ν_{O0} [a]	ν_{S0} [a]	$F^{\gamma}_{0,0}(J)$ [b]	ν_{O1} [a]	ν_{S1} [a]	$F^{\gamma}_{1,0}(J)$ [b]
3	-14.31	25.86	1.00028	**1541.972**	**1581.772**	0.96575
5	-25.86	37.35	1.00044	**1530.315**	**1592.865**	0.94673
7	-37.35	48.83	1.00059	**1518.537**	**1603.819**	0.92812
9	-48.83	60.31	1.00075	**1506.632**	**1614.635**	0.90991
11	-60.31	71.77	1.00091	**1494.622**	**1625.316**	0.89211
13	-71.77	83.23	1.00107	**1482.486**	**1635.857**	0.87471
15	-83.23	94.66	1.00122	**1470.239**	**1646.252**	0.85772
17	-94.66	106.09	1.00138	**1457.878**	**1656.511**	0.84113
19	-106.09	117.50	1.00154	1445.408	1666.620	0.82495
21	-117.50	128.89	1.00169	1432.828	1676.583	0.80917
23	-128.89	140.25	1.00185	1420.142	1686.397	0.79378
25	-140.25	151.60	1.00200	1407.349	1696.059	0.77880
27	-151.60	162.92	1.00215	1394.453	1705.568	0.76420
29	-162.92	174.21	1.00230	1381.456	1714.922	0.75000
31	-174.21	185.48	1.00245	1368.358	1724.119	0.73617
33	-185.48	196.71	1.00260	1355.161	1733.157	0.72273
35	-196.71	207.91	1.00275	1341.868	1742.033	0.70967
37	-207.91	219.08	1.00290	1328.480	1750.747	0.69697
39	-219.08	230.22	1.00305	1314.999	1759.296	0.68464
41	-230.22	241.31	1.00319	1301.427	1767.679	0.67267
43	-241.31	252.37	1.00333	1287.765	1775.893	0.66105
45	-252.37	263.38	1.00347	1274.016	1783.936	0.64979

[a] Transition wavenumbers in regular text have been calculated by solving the radial nuclear equation for $^{16}O_2$, using the the potential curve from Bytautas et al.[10] The maximal error in these calculated transition wavenumbers is around 0.15 cm^{-1}. Values in bold face, for J=3–17 for O1- and S1-branch of $^{16}O_2$ obtained from the work of Edwards et al.[11]. These have error of ∼0.009 cm^{-1} or less.

[b] Wavelength independent correction factors for vibration-rotation interaction in O_2 have been calculated from the expansion of the Herman-Wallis factors from Ref. [12]. In terms of the ratio of square of matrix elements of polarizability anisotropy, these are defined as, $F^{\gamma}_{v',v}(J) = \frac{\langle\gamma\rangle^2_{v',J+2;v,J}}{\langle\gamma\rangle^2_{v',J-2;v,J}}$

Sec 5. Error in wavenumber calibration

In the present wavenumber calibration scheme, both the x-values (pixels) and the y-values (reference Ramanshift/ cm^{-1}) have errors. However, the reference Ramanshift values have very small error compared to the uncertainty in band positions in pixel scale (see Table S8). Hence, the source of error is the uncertainty in band positions (in pixels obtained from band fitting). (For example, see Fig. 2 in the main document, and Fig. S14 in this Supplementary material.)

To transform the error in pixel positions to wavenumbers, cm^{-1}, three set of x-vectors (in pixels) were setup, first without any error, $x_0 = 0..1599$ (corresponding to 1600 pixels); second with error (δx) subtracted, $x_{-\delta} = 0 - \delta x$, 1-$\delta x$,.., 1599-$\delta x$, and third with error added, $x_{+\delta} = 0 + \delta x$, 1+$\delta x$,.., 1599+$\delta x$. Corresponding y-vectors $(y_0, y_{-\delta}, y_{+\delta})$ were computed for the three x-vectors as $y = f(x)$ where, $f(x) = c_0 + c_1x + c_2x^2 + c_3x^3$; c_0, c_1, c_2 and c_3 are the polynomial coefficients from the least-squares fit. The error was taken as the maximum difference between y_0 and $y_{-\delta}$ or $y_{+\delta}$ over all pixel points. This error originating from the uncertainty in pixel positions was combined with the residual in y, using error propagation, to obtain the net error in wavenumber calibration. See following table for a summary of error estimation in wavenumber calibration for parallel and perpendicular polarizations.

TABLE S8: Table listing sources of error and the net error (**Net δ**, cm^{-1}) in wavenumber calibration for parallel and pependicular polarization measurements.

Parallel polarization				
	Source			**Net δ[a]**
	From residual in x, pixel		From residual in y, cm^{-1}	
Measure	**δpixel**	**δcm^{-1} [b]**	**δcm^{-1}**	
3σ[c]	0.108	0.24	0.004956	0.24

Perpendicular polarization				
	Source			**Net δ[a]**
	From residual in x, pixel		From residual in y, cm^{-1}	
Measure	**δpixel**	**δcm^{-1} [b]**	**δcm^{-1}**	
3σ[c]	0.166	0.34	0.002458	0.34

a : Defined as $\sqrt{(\text{error from x, in } cm^{-1})^2 + (\text{error from y, in } cm^{-1})^2}$
b : After transforming to wavenumbers, as mentioned in the above text.
c : 3σ std. deviation covers 99.7% of the probability.

Error in the transition wavenumbers of select vibrational bands from 5 organic liquids (in Table 3 of the main article) includes the net error mentioned above and the uncertainty in band position obtained while fitting.

Sec 6. Obtained correction curves for intensity calibration

Correction curves for pixel-to-pixel variation in the sensitivity, obtained by fitting broadband whitelight assuming black body emission in photons/wavenumber.These are specific to our experimental setup, its alignment and the specific optical components used including the lamp.

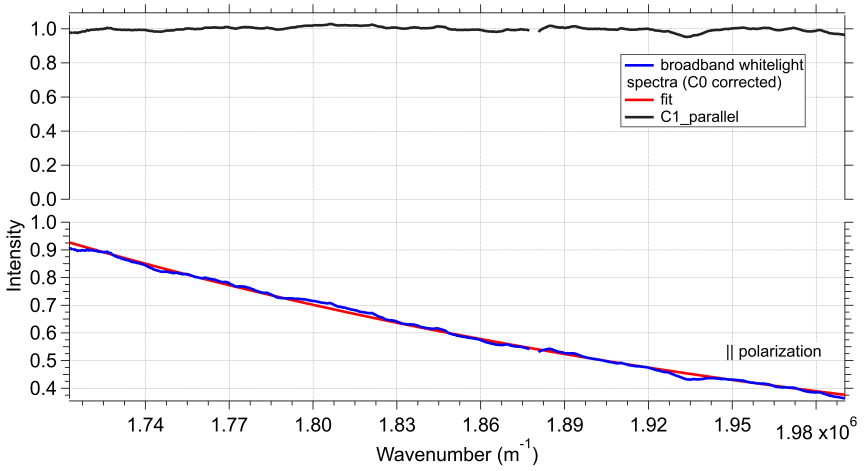


FIG. S9: Correction for parallel polarized spectra, for pixel-to-pixel change in sensitivity (shown as black color and labelled C_parallel). This correction was obtained by ratio of the observed broadband white light spectra (in blue color) to the fit, assuming black body emission (in red color).

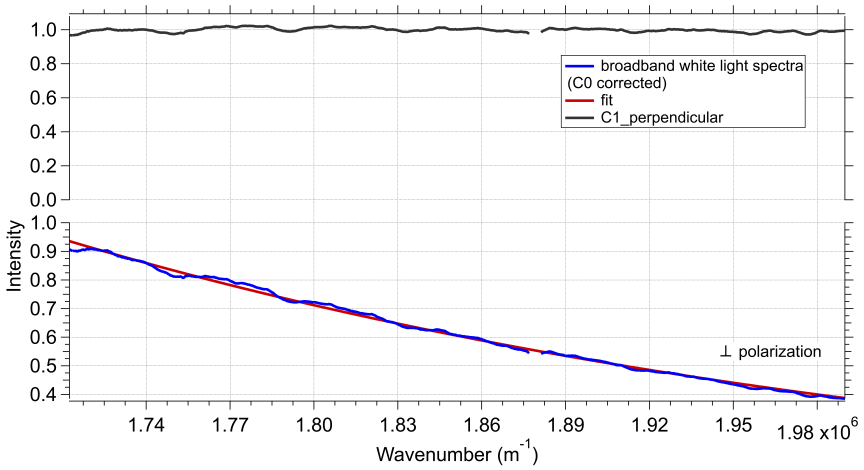


FIG. S10: Second correction for perpendicularly polarized spectra, for pixel-to-pixel change in sensitivity (shown as black color and labelled C_perpendicular). This correction was obtained by ratio of the observed broadband white light spectra (in blue color) to the fit, assuming black body emission (in red color).

Following correction curves were obtained as the final correction from experimental rotational Raman band intensities assuming quadratic polynomial for data from parallel polarized spectra, and cubic polynomial for data from perpendicularly polarized spectra.

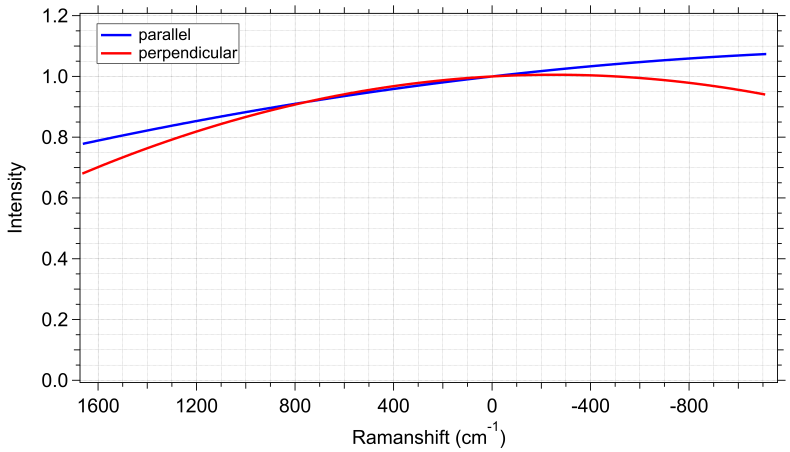


FIG. S11: Correction curves for parallel and perpendicularly polarized spectra. These are specific to our experimental setup, its alignment and the specific optical components used.

TABLE S12: Residuals from fit for obtaining final correction for intensities from intensity ratios of rotational Raman bands. Change of residuals over the number of unknown variables in the fit is shown as Fig. S12 below. It is observed that residuals are almost constant if quadratic or higher polynomial are used for representing the wavelength dependent sensitivity.

Parallel polarization		
Variable(s)	**Model**	**Residual**
1	Linear	0.028180
2	**Quadratic**	**0.027637**
3	Cubic	0.027637
4	Quartic	0.027637

Perpendicular polarization		
Variable(s)	**Model**	**Residual**
1	Linear	0.019169
2	Quadratic	0.011554
3	**Cubic**	**0.011548**
4	Quartic	0.011547

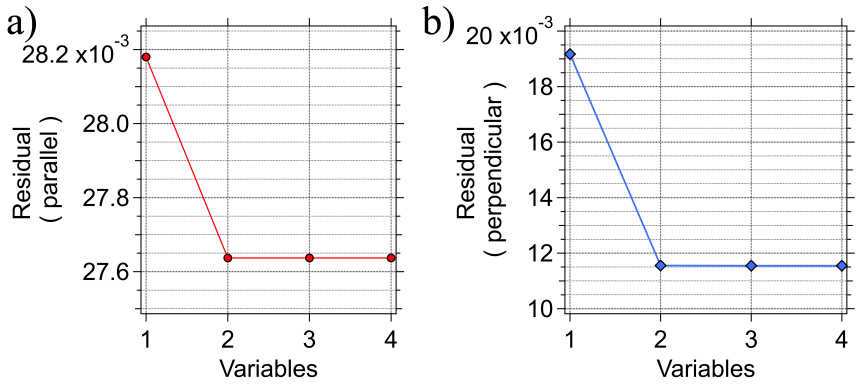


FIG. S13: Residuals from linear, quadratic, cubic and quartic fitting for data from a) parallel and b) perpendicularly polarized, spectra. The number of unknown variables in the fit are across the horizontal axis.

Sec 7. Results for perpendicular polarization

A. Wavenumber calibration

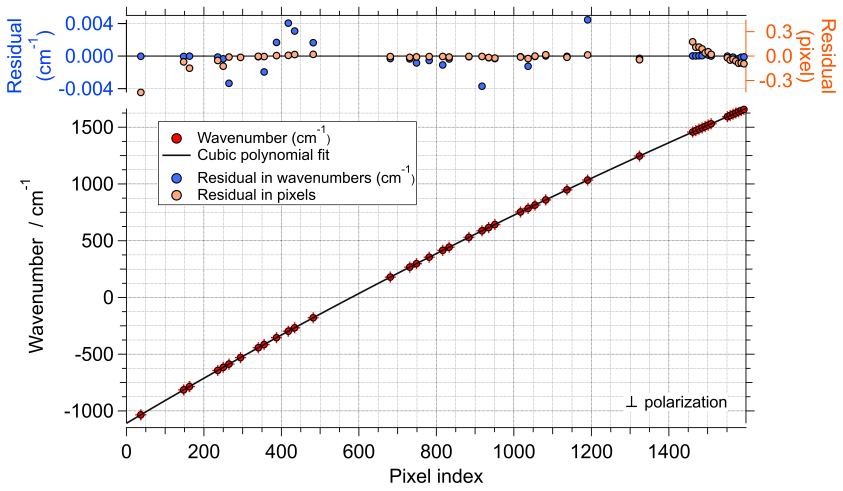


FIG. S14: Fit of reference wavenumber (cm^{-1}) over band position (in pixel) for wavenumber calibration using pure rotational Raman bands from H_2, HD and D_2 and vibration-rotation bands from O_2, for perpendicular polarization. Total 45 data points were used (31 coming from H_2, HD and D_2; 14 from O_2)

B. Fitting for sensitivity curve for perpendicular polarization

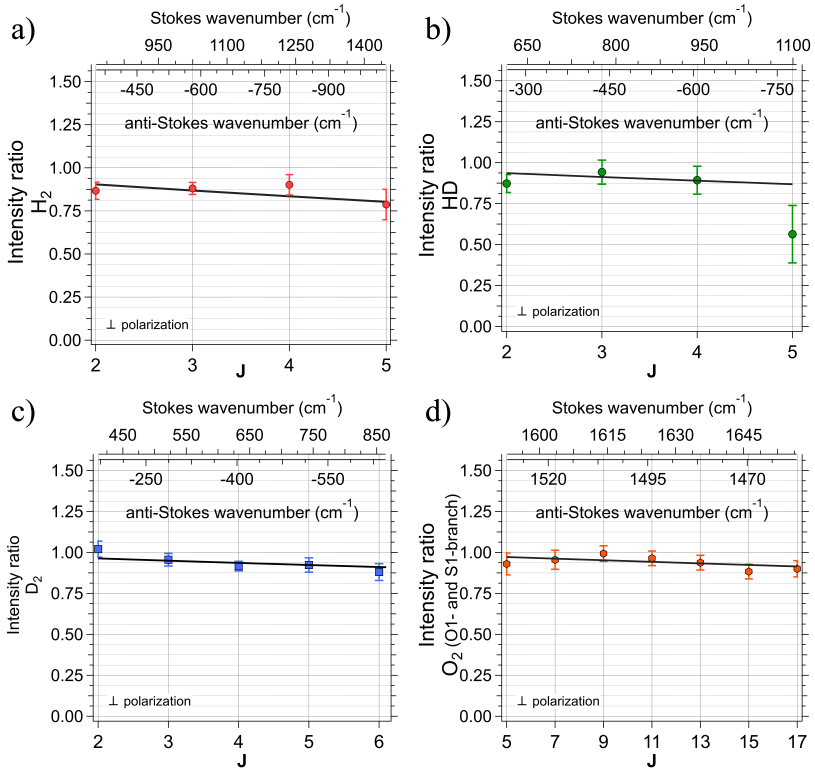


FIG. S15: Fit of intensity ratios for H_2, HD, D_2 and O_2, where intensity ratio refers to the ratio of experimental intensities (Stokes to anti-Stokes for H_2, HD, D_2; S1 to O1 for O_2) to corresponding theoretical intensities. The intensity ratio is the LHS of the Eqn. (6). The fit (in black trace) is the ratio of corresponding values of the polynomial at specific pairs of wavenumbers. The fit is the RHS of Eqn. (6) in the main article.

C. Temperature determination from water

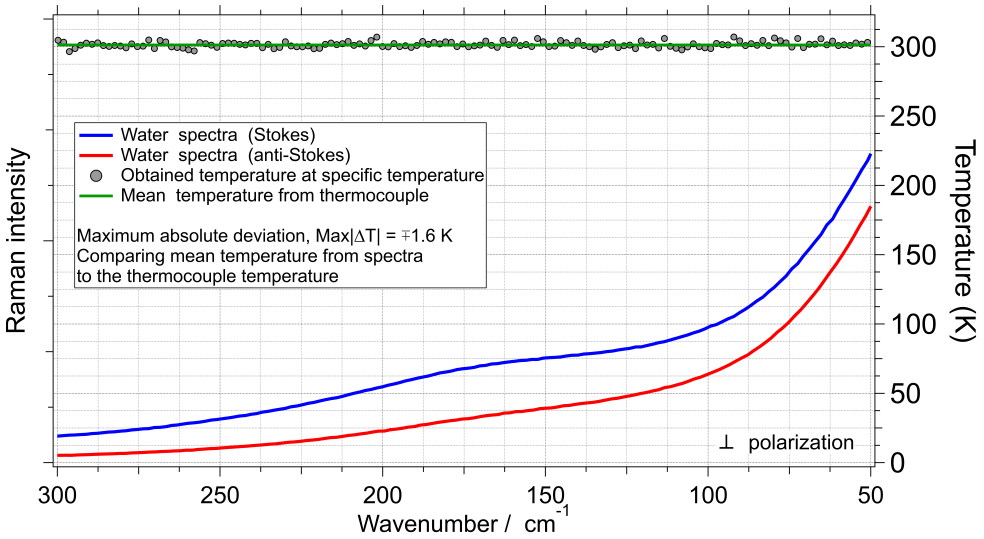
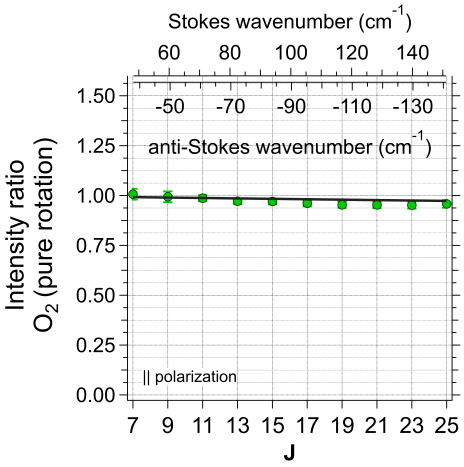


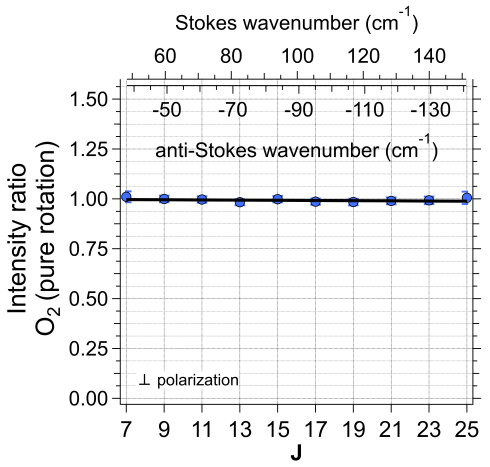
FIG. S16: Comparison of temperatures determined for liquid water using the anti-Stokes to Stokes ratio in the low wavenumber region of the Raman spectra water. The temperatures determined from spectra at specific wavenumbers are shown in grey circles. Mean temperature measured using a thermocouple during the spectra acquisition is shown as green horizontal line. Comparison of average temperature from all the spectral data points (computed from anti-Stokes to Stokes Raman intensities) and mean temperature from thermocouple was performed to obtain the deviation. Maximum deviation in three independent experiments was found to be ±1.6 K for perpendicular polarization.

Sec 8. Fitting result for pure rotation bands from O_2 for calibration

In our fitting analysis for obtaining the wavelength dependent sensitivity, data points from the pure rotation bands from O_2 were not included. When included, for wavenumber calibration, the change in the value of obtained fit coefficients was less than 0.2%. For intensity calibration, for obtaining wavelength dependent sensitivity, the resulting fit curves were found to change within 0.7%, in absolute value across the spectral range. Following figures show the obtained fit towards the wavelength dependent sensitivity **without** inclusion of these data points revealing good correspondence and indicating that the obtained fits are effectively valid for these data points as well.



(a) Fitting for parallel polarized spectra



(b) Fitting for perpendicularly polarized spectra

FIG. S17: Fit of intensity ratios for O_2, where intensity ratio refers to the ratio of experimental intensities (Stokes to anti-Stokes for O_2) to corresponding theoretical intensities. The intensity ratios (in green and blue data points) are the LHS of the Eqn. (6) in the main article. The fit (in black trace) is the ratio of corresponding values of the polynomial at specific pairs of wavenumbers. The fit is RHS of Eqn. (6) in the main article.

Sec 9. Standard Raman spectra of organic liquids

All Raman spectra plotted below are the intensity calibrated parallel and perpendicularly polarized spectra, measured using 532.2 nm laser excitation. Temperature of the liquid, having uncertainty of +/-0.8 K, measured during Raman spectra acquisition is mentioned in the captions. The Rayleigh has been artificially removed in all the spectra.

A. Carbon tetrachloride

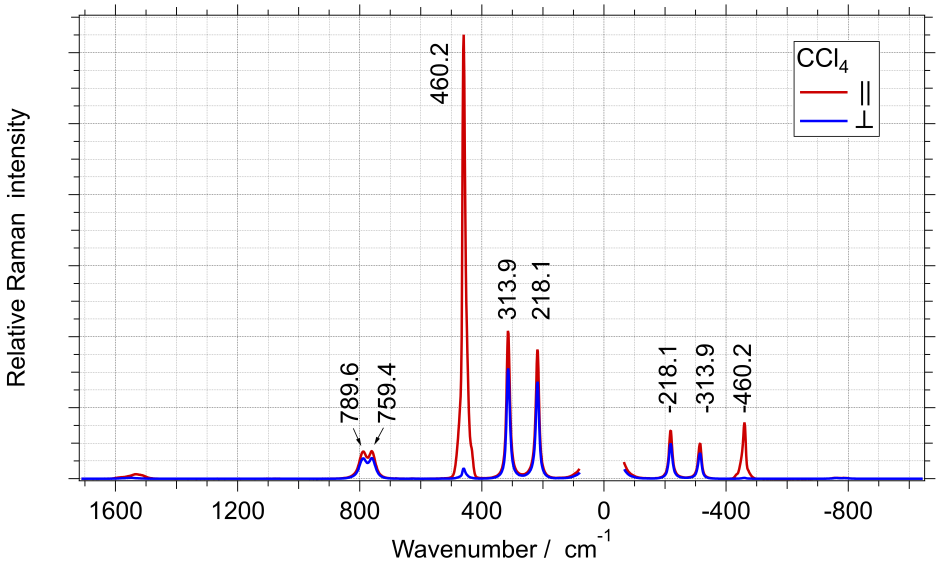


FIG. S18: Calibrated Raman spectra of liquid CCl_4 at 298.6 K

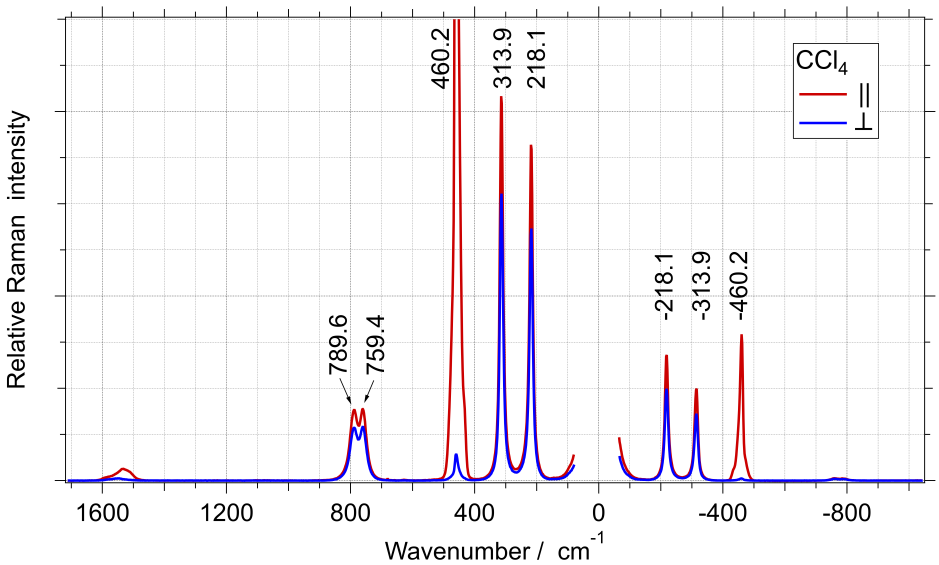


FIG. S19: Zoomed up of the above plot (liquid CCl_4 at 298.6 K)

B. Cyclohexane

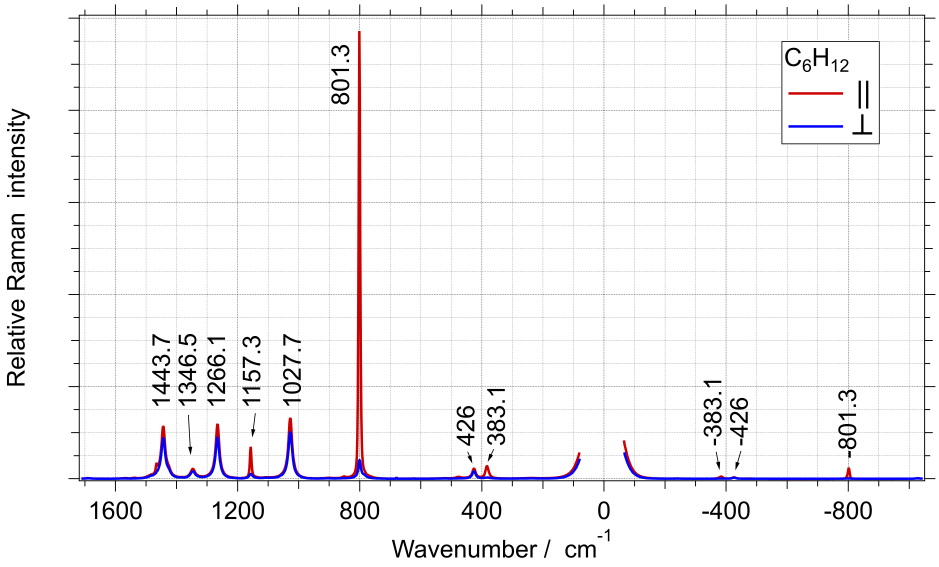


FIG. S20: Calibrated Raman spectra of liquid C_6H_{12} at 298.6 K

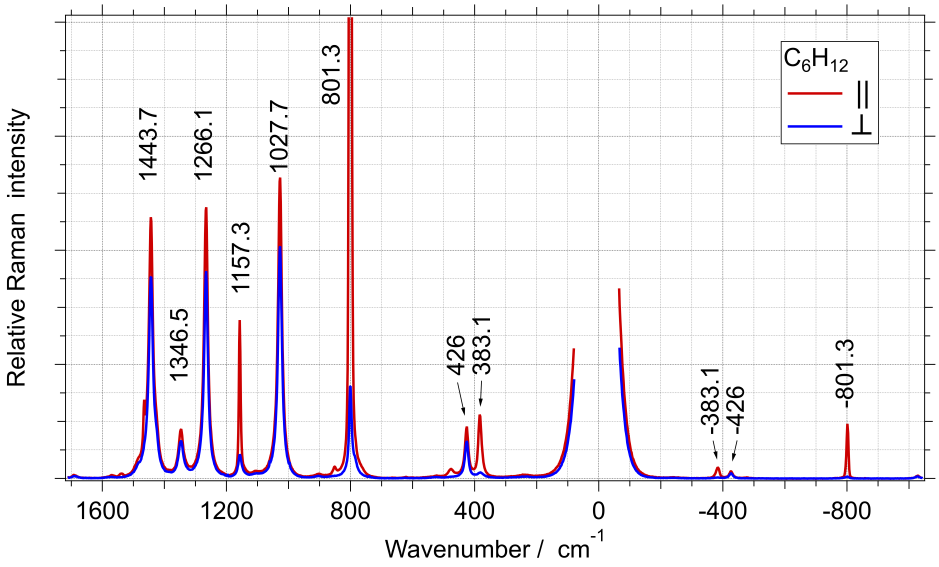


FIG. S21: Zoomed up of the above plot (liquid C_6H_{12} at 298.6 K)

C. Benzene

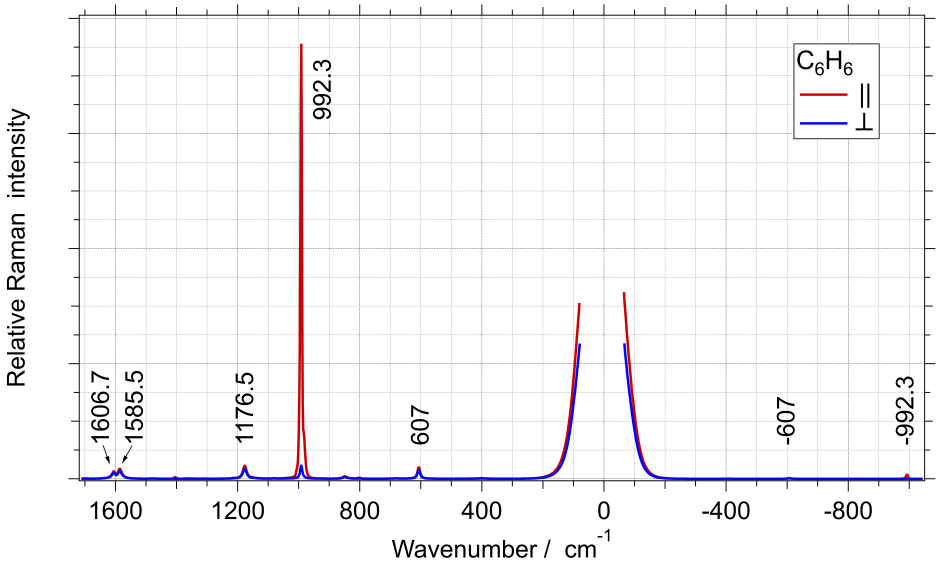


FIG. S22: Calibrated Raman spectra of liquid C_6H_6 at 298.6 K

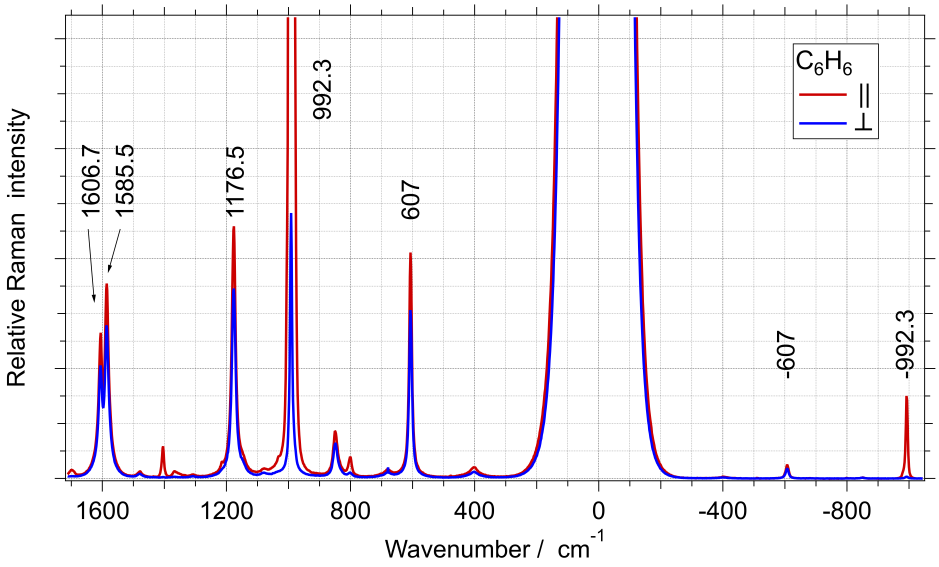


FIG. S23: Zoomed up of the above plot (liquid C_6H_6 at 298.6 K)

D. Toluene

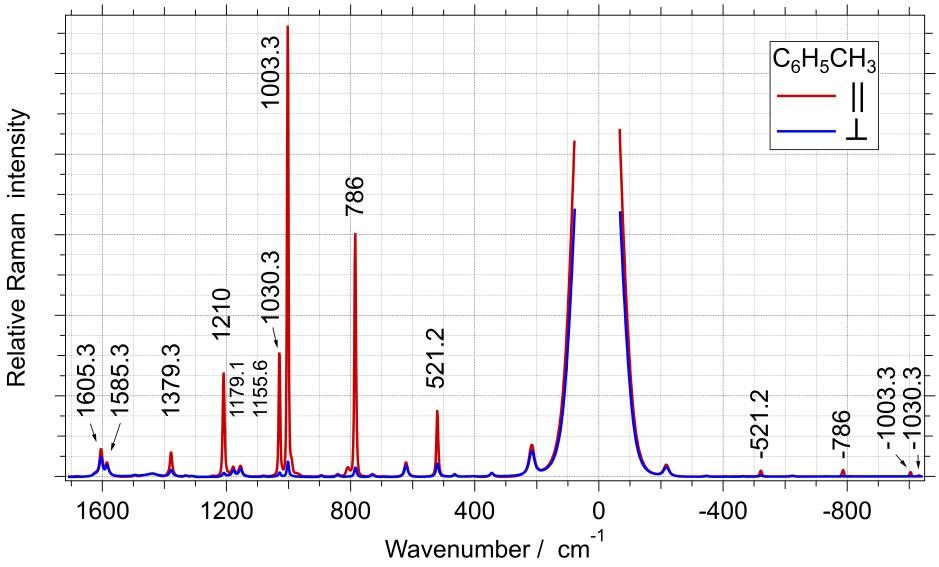


FIG. S24: Calibrated Raman spectra of liquid $C_6H_5CH_3$ at 299.2 K

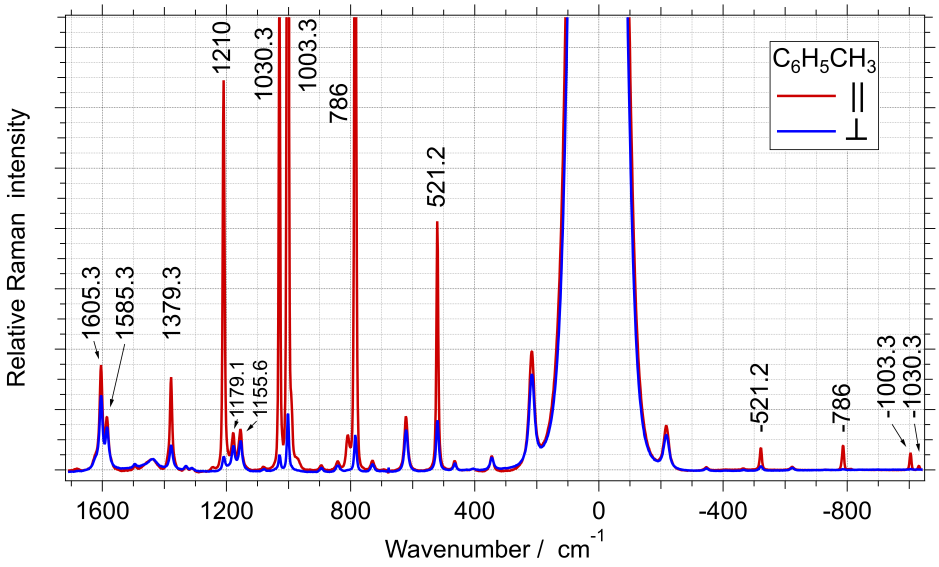


FIG. S25: Zoomed up of the above plot (liquid $C_6H_5CH_3$ at 299.2 K)

E. Benzonitrile

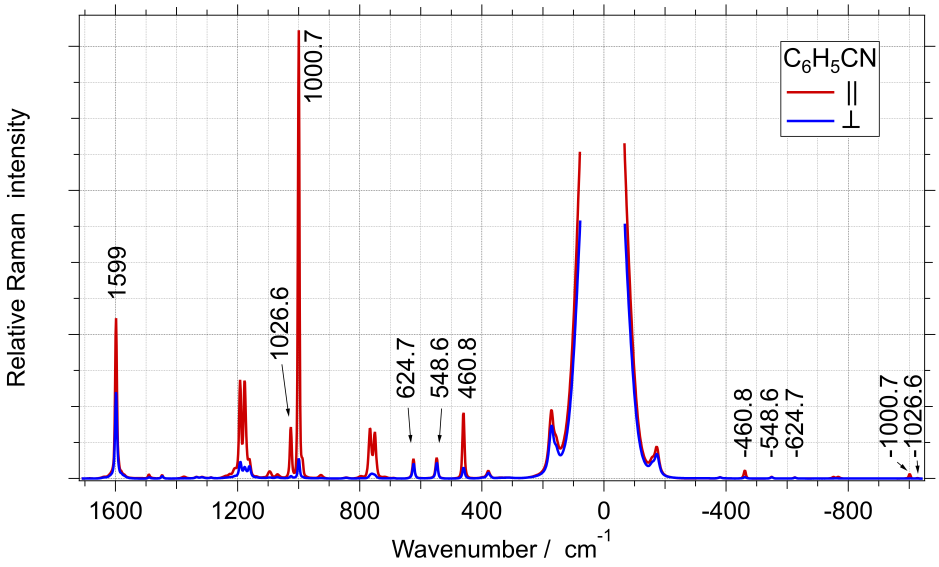


FIG. S26: Calibrated Raman spectra of liquid C_6H_5CN at 299.5 K

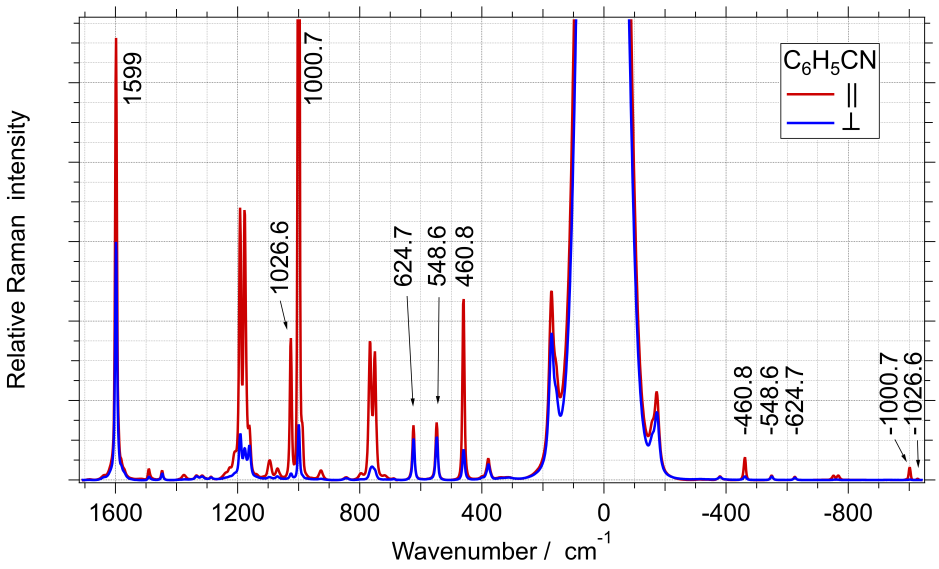


FIG. S27: Zoomed up of the above plot (liquid C_6H_5CN at 299.5 K)

F. Details about band fitting and analysis of the Raman spectra of organic liquids

TABLE S28: Details on band fitting of organic liquids

Molecule	Index	(ν cm^{-1})	Fit function used[a]	Comment[b]
CCl_4	1	-460.2	L	3 Lorentzians, no constraints. (3 Gaussians also work)
	2	-313.9	L	
	3	-218.1	L	
	4	218.1	L	
	5	313.9	L	
	6	460.2	L	Same as 1.
	7	759.4	V	
	8	789.6	V	
C_6H_{12}	9	-801.3	L	
	10	-426.0	L	
	11	-383.1	L	
	12	383.1	L	
	13	426.0	L	
	14	801.3	L	
	15	1027.7	L	
	16	1157.3	L	
	17	1266.1	L	
	18	1346.5	L	
	19	1443.7	L	Shoulder band on the higher wavenumber side (1465 cm^{-1}, polarized) was considered as a separate band while fitting and its band area was not included within the band area of 1443.7 cm^{-1} band.

[a] G : Gaussian, L : Lorentzian, V : Voigt.
[b] Band fitting has been performed in IgorPro using Multi-peak fit.

TABLE S29: Details on band fitting of organic liquids

Molecule	Index	(ν cm^{-1})	Fit function used[a]	Comment[b]
C_6H_6	20	-992.3	L	Shoulder band near -983 cm^{-1}, coming from natural ^{13}C substituted benzene species, was included in the fit and its band area was treated together in the analysis.
	21	-607.0	L	
	22	607.0	L	
	23	992.3	L	Same as 20, but for shoulder band at +983 cm^{-1}
	24	1176.5	L	
	25	1585.5	L	
	26	1606.7	L	
$C_6H_5CH_3$	27	-1030.3	G	
	28	-1003.3	L	Shoulder band near -983 cm^{-1} was included. Same reason as 20.
	29	-786.0	L	
	30	-521.2	L	
	31	521.2	L	
	32	786.0	L	
	33	1003.3	L	Same as 28 and 23.
	34	1030.3	G	
	35	1155.6	L	
	36	1179.1	L	
	37	1210.0	L	
	38	1379.3	L	
	39	1585.3	L	
	40	1605.3	L	
C_6H_5CN	41	-1026.6	G	
	42	-1000.7	G	Shoulder band is included as another G. Same as 20 and 28.
	43	-624.7	L	
	44	-548.6	L	
	45	-460.8	L	
	46	460.8	L	
	47	548.6	L	
	48	624.7	L	
	49	1000.7	G	Same as 42.
	50	1026.6	G	
	51	1599.0	L	

[a] G : Gaussian, L : Lorentzian, V : Voigt.
[b] Band fitting has been performed in IgorPro using Multi-peak fit.

Sec 10. Raman spectrum of indene for wavenumber calibration

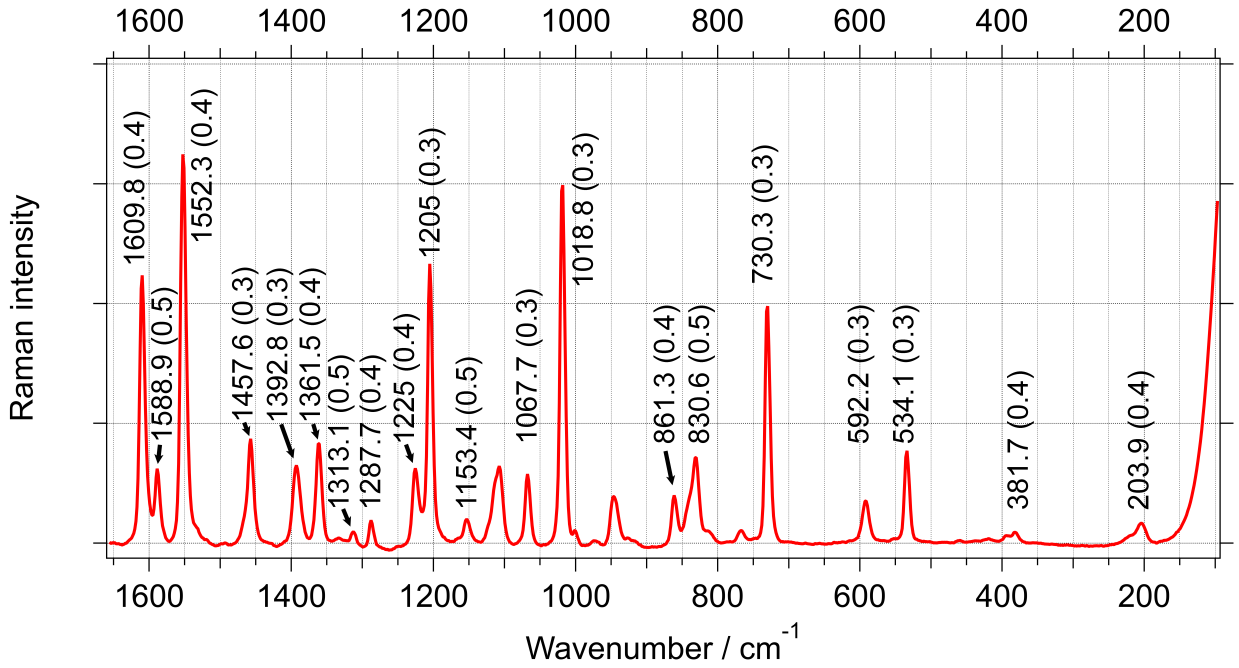


FIG. S30: Raman spectra of indene with band positions and the uncertainties. Raman intensities are not calibrated.

Bibliography

[1] B. Valeur, *Molecular Fluorescence: Principles and Applications.* (Wiley-VCH, Verlag GmbH, 2001).
[2] G. Angulo, G. Grampp, and A. Rosspeintner, Spectrochim. Acta A **65**, 727 (2006).
[3] J. Mooney and P. Kambhampati, J. Phys. Chem. Lett. **5**, 3497 (2014).
[4] J. Mooney and P. Kambhampati, J. Phys. Chem. Lett. **4**, 3316 (2013).
[5] Igor Pro, WaveMetrics, Lake Oswego, OR, USA, A scientific data analysis software with numerical computing environment and a programming language.
[6] E. Jones, T. Oliphant, P. Peterson, *et al.*, "SciPy: Open source scientific tools for Python," (2001–), accessed: 2019-09-02.
[7] F. Gao and L. Han, Computational Optimization and Applications **51**, 259 (2012).
[8] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes:The Art of Scientific Computing*, 3rd ed. (Cambridge University Press, New York, NY, USA, 2007).
[9] "A repository containing a Igor procedures and Python modules for wavenumber and intensity calibration using Raman spectra from H_2, HD, D_2 and O_2." `https://github.com/ankit7540/RamanSpecCalibration`, accessed: 2020-02-28.
[10] L. Bytautas, N. Matsunaga, and K. Ruedenberg, J. Chem. Phys. **132**, 074307 (2010).
[11] H. G. M. Edwards, D. A. Long, K. A. B. Najm, and M. Thomsen, J Raman Spectrosc. **10**, 60 (1981).
[12] M. A. Buldakov, V. N. Cherepanov, B. V. Korolev, and I. I. Matrosov, J. Mol. Spectrosc. **217**, 1 (2003).