Use of the complete basis set limit for computing highly accurate *ab initio* dipole moments

Eamon K. Conway 🆔, Iouli E. Gordon 🆔, Oleg L. Polyansky, and Jonathan Tennyson 🆔

View Online   **Export Citation**   **CrossMark**

ARTICLES YOU MAY BE INTERESTED IN

Use of the complete basis set limit for computing highly accurate *ab initio* dipole moments

View Online   Export Citation   CrossMark

Eamon K. Conway,[1,2,a] iD **Iouli E. Gordon,**[1] iD **Oleg L. Polyansky,**[2] **and Jonathan Tennyson**[2] iD

AFFILIATIONS

[1] Harvard & Smithsonian | Center for Astrophysics, Atomic and Molecular Physics Division, Cambridge, Massachusetts 02138, USA

[2] Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom

[a] Author to whom correspondence should be addressed: eamon.conway@cfa.harvard.edu

ABSTRACT

Calculating dipole moments with high-order basis sets is generally only possible for the light molecules, such as water. A simple, yet highly effective strategy of obtaining high-order dipoles with small, computationally less expensive basis sets is described. Using the finite field method for computing dipoles, energies calculated with small basis sets can be extrapolated to produce dipoles that are comparable to those obtained in high order calculations. The method reduces computational resources by approximately 50% (allowing the calculation of reliable dipole moments for larger molecules) and simultaneously improves the agreement with experimentally measured infrared transition intensities. For atmospherically important molecules, which are typically too large to consider the use of large basis sets, this procedure will provide the necessary means of improving calculated spectral intensities by several percent.

Published under license by AIP Publishing. https://doi.org/10.1063/1.5135931

I. INTRODUCTION

The ability of atmospheric science missions to accurately detect trace amounts of molecules in our atmosphere is placing significant demands on spectroscopy,[1–5] both theoretical and experimental. For instance, spectrometers on board of satellites including GOSAT[6] and OCO-2[7] now aim to detect carbon dioxide in the terrestrial atmosphere to an accuracy better than 0.3%. Achieving this precision requires all line parameters to be known to very high accuracy. Atmospheric missions rely on the spectroscopic data in the high-resolution transmission molecular absorption (HITRAN) database,[8] which typically combines experimental and theoretical data in order to achieve both completeness and accuracy of reference line lists. Every four years, the HITRAN parameters are being updated and extended to incorporate state of the art data.

Modern experimental methods, especially those that employ frequency combs (see, for instance, Refs. 9–11), can measure line positions with errors on the order of 10^{-5} cm^{-1}, if not smaller, which is several orders of magnitude better than the accuracy that can be achieved by calculations using the best semiempirical potential energy surfaces available.[12–14] For example, for water vapor, the most accurate potential energy surface available[14] predicts energy

levels below 15 000 cm^{-1} with a standard deviation $\sigma = 0.011$ cm^{-1}. These high-accuracy semiempirical potentials are achieved via a process known as refinement.[15] When experimental data are available, the parameters used in fitting the underlying electronic structure calculations are adjusted to improve the agreement between theory and the observed data. This procedure, when done correctly, allows energy levels to be predicted to less than 0.03 cm^{-1}; hence, the accuracy is no longer proportional to the level of theory considered for the electronic structure calculations.[13,14,16,17] However, when refinement is not possible, then the level of theory becomes significantly more important. Extrapolating calculated energies to the complete basis set (CBS) is one such method that has been used with great success. Second order corrections such as relativistic, adiabatic, nonadiabatic, quantum electrodynamics, and spin orbit coupling can also become important.[18–21]

Cavity ring down spectroscopy can measure those transitions that have very weak intensities with high accuracy.[22] For well-studied molecules including H_2O and CO_2, transition intensities obtained from *ab initio* calculations and experiments are generally in excellent agreement with deviation on the sub-one-percent scale for many bands.[23–25]

For a given method, the basis set chosen for the electronic structure calculations largely determines the accuracy of

an *ab initio* dipole calculation. The Dunning[26–28] aug-cc-pCV6Z basis set is the largest conventional basis set yet considered for the calculation of a global dipole surface[29,30] for a triatomic system, in this case water. Each of these dipole calculations required almost two days worth of central processing unit (CPU) time to compute;[29,30] hence, for molecules with more electrons, such as CO_2, use of the aug-cc-pCV6Z basis is currently unfeasible.

Even with the extensive computer resources used in the creation of accurate dipole moment surfaces (DMS's) for water, over 100 years of CPU time in the case the CKAPTEN DMS,[29] there still remains issues with predicted infrared transition intensities. For example, the ν_1 fundamental has been shown to be very sensitive to the choice of *ab initio* calculations[25,29,31,32] and improvements are needed. Computing dipoles with a larger basis set is one solution, although not feasible on a global scale. Similar problems persist for other atmospherically important molecules such as CO_2[33] and ozone.[34,35]

This work aims to explore the use of complete basis set (CBS) extrapolation in the computation of high-accuracy *ab initio* spectral intensities both to give increased accuracy and to reduce the computational cost.

II. METHOD

A. Calculations

We have computed over 16 000 aug-cc-pCV6Z[26–28] finite-field dipoles for water using the electronic structure program MOLPRO.[36] With these, we created a DMS for water vapor, named CKAPTEN.[29] Spectra computed with this CKAPTEN surface have been shown to produce excellent transition intensities when compared to both experiment and observation.[37] From these 16 000 configurations, we create a smaller subgrid with 2540 data points. The points were chosen such that there is sufficient coverage of both stretching and bending coordinates, which would provide accurate transition intensities up to approximately 15 000 cm^{-1}. For calculating reliable spectra in the ultraviolet, it is necessary to limit the use of fitting parameters as done in CKAPTEN, which reduces the possibility of artificial oscillations occurring in highly energetic regions of the dipole surface. In this work, we are not interested in such energetic transitions and are not required to use a few parameter fit nor to use 16 000 points to better constrain the fit. The stretching coordinates in this subgrid are in the range of 1.4 a_o–4 a_o, while the angular selection lies between $30°$ and $178°$. With respect to our equilibrium configuration of 1.8141 a_o and $104.52°$, our dataset includes dipoles with energies up to 42 481 cm^{-1}. The original 16 000 finite-field dipoles were computed at the multireference configuration interaction (MRCI) level with the aug-cc-pCV6Z basis set, a Davidson correction (+Q), and Douglass-Kroll-Hess-Hamiltonian to order two (DKH2) and with an electric field strength of 5 $\times 10^{-5}$ a.u. We calculate finite-field dipoles on our subgrid of 2540 points that use smaller basis sets, notably the aug-cc-pCV(Q,5)Z sets, while still using the same formalism employed for the aug-cc-pCV6Z dipoles. Calculations were all performed on Legion, Grace, and Myriad systems at the University College London High Performance Computing Facilities, and CPU times given below are for these computers.

B. Extrapolation technique

We configure our water molecule such that the z-axis bisects the angle HOH, with the x-axis in the HOH plane and perpendicular to z. In the finite field approach, we apply a small electric field, λ, in each of the directions: $+\hat{x}$, $-\hat{x}$, $+\hat{z}$, and $-\hat{z}$. The energy of the molecule in each respective field will be termed E_{px}, E_{nx}, E_{pz}, and E_{nz}. p and n signify positively and negatively directed fields, respectively, while x and z are the components. Each dipole component is calculated as

$$\mu_{x(z)} = \frac{(E_{px(pz)} - E_{nx(nz)})}{2\lambda}. \tag{1}$$

We extrapolate the individual energies E_{px}, E_{nx}, E_{pz}, and E_{nz} with a standard expression[38]

$$E_x = E_{cbs} + be^{-x}, \tag{2}$$

where E_{cbs} and b are fitted parameters, while x represents the level of theory. This will provide us with E_{px}^{cbs}, E_{nx}^{cbs}, E_{pz}^{cbs}, and E_{nz}^{cbs}. This formula (2) has successfully been used in the past to extrapolate *ab initio* energies for potential energy surfaces.[13,21,39] Instead of one single energy extrapolation to do, we have five: the zero-field calculation and one for each of our four dipole components within an electric field.

We are interested in applying this technique to produce extrapolated dipoles, $\mu_{x(z)}^{cbs_{Q5}}$ and $\mu_{x(z)}^{cbs_{56}}$. We expect the $\mu_{x(z)}^{cbs_{Q5}}$ dipoles to behave as the aug-cc-pCV6Z dipoles and the $\mu_{x(z)}^{cbs_{56}}$ dipoles to behave as aug-cc-pCV7Z dipoles.

To create $\mu_{x(z)}^{cbs_{Q5}}$ dipoles, we require $E_{px}^{cbs_{Q5}}$, $E_{nx}^{cbs_{Q5}}$, $E_{pz}^{cbs_{Q5}}$, and $E_{nz}^{cbs_{Q5}}$, which represent the individually extrapolated energies. Likewise, to calculate $\mu_{x(z)}^{cbs_{56}}$, we need $E_{px}^{cbs_{56}}$, $E_{nx}^{cbs_{56}}$, $E_{pz}^{cbs_{56}}$, and $E_{nz}^{cbs_{56}}$. Combining these values, we can now calculate our extrapolated dipoles,

$$\mu_{x(z)}^{cbs} = \frac{(E_{px(pz)}^{cbs} - E_{nx(nz)}^{cbs})}{2\lambda}. \tag{3}$$

In total, we possess five sets of dipoles, one for each of the aug-cc-pCV(Q,5,6)Z basis sets and then two sets of extrapolated sets, CBS^{Q5} and CBS^{56}.

C. Fitting

We fit each dipole set to a similar functional form used in the creation of CKAPTEN,[29]

$$\mu_z(r_1, r_2, \theta) = (\pi - \theta) \sum_{i,j,k} C_{ijk}^{(z)} \zeta_1^i \zeta_2^j \zeta_3^k, \tag{4}$$

$$\mu_x(r_1, r_2, \theta) = \sum_{i,j,k} C_{ijk}^{(x)} \zeta_1^i \zeta_2^j \zeta_3^k, \tag{5}$$

where $\zeta_1 = \frac{(r_1 + r_2)}{2} - r_e$, $\zeta_2 = (r_2 - r_1)$, and $\zeta_3 = \theta/\theta_e$. To physically model the dipole surface correctly, we have several conditions that must be adhered to the following requirements:

(i) $\mu_z(r_1, r_2, \theta = \pi) = 0$,
(ii) $\mu_z(r_1, r_2, \theta) = \mu_z(r_2, r_1, \theta)$, and
(iii) $\mu_x(r_1, r_2, \theta) = -\mu_x(r_2, r_1, \theta)$.

TABLE I. The weighted root-mean-square (rms) deviation of each fitted dipole component for each surface (see text for details).

	μ_z (a.u.)	μ_x (a.u.)
Q-zeta	7.29×10^{-5}	1.10×10^{-4}
5-zeta	5.55×10^{-5}	9.95×10^{-5}
CBS(45)	8.59×10^{-5}	1.20×10^{-4}
6-zeta	5.35×10^{-5}	9.96×10^{-5}
CBS(56)	7.77×10^{-5}	1.10×10^{-4}

These requirements mean that only even (odd) powers of j are used in the expression ζ^j for the $\hat{z}$ ($\hat{x}$) components.

The dipoles in each set are weighted[31,32] (w_i) as a function of their energy,

$$s_i = \tanh(-0.006(E_i - 30\,000) + 1.002\,002\,002)/2.002\,002\,002$$
$$w_i = 25\,000(s_i)/\max(E_i, 25\,000). \tag{6}$$

To facilitate an equal comparison between the different levels of theory, each dipole set must be fit to the same functional form, requiring the same number of parameters in each fit. We fit the $\hat{z}$ components with 228 parameters and the $\hat{x}$ with 163 parameters.

Table I shows the average root-mean-square deviation of the dipole fitting procedure for each of the individual surfaces. For the parallel $\hat{z}$ component, each surface is fit to a root-mean-square (rms) under 10^{-4} a.u., while for the perpendicular $\hat{x}$ component, each surface possesses an rms of approximately 1.1×10^{-4} a.u. The perpendicular component is often the most difficult to fit and carries a larger rms than the respective parallel component of the dipole[29,30]

To compare the surfaces, we need to assess the resulting transition intensities. To do this, we require wavefunctions, which we calculate from the potential energy surface of Mizus *et al.*[14] This PES, called PES15K, is valid for energies that fall below 15 000 cm^{-1}, which is sufficient for our study. For states falling below this threshold, PES15K predicts energies to a rms of only 0.011 cm^{-1}. Using the DVR3D[42] suite of programs, we calculate spectra for each of the

five dipole surfaces with $J_{max} = 6$, $\nu_{max} = 15\,000$ cm^{-1}, $S_{if} \geq 10^{-30}$ cm molecule^{-1}, and $T = 296$ K.

III. RESULTS

Table II contains energies E_{px}, E_{nx}, E_{pz}, and E_{nz} with the corresponding dipoles μ_x and μ_z obtained using basis sets aug-cc-pCV(X=Q,5,6)Z on one molecular geometry: $R_1 = 1.7$ a$_o$, $R_2 = 1.98$ a$_o$, and $\theta = 163°$. Attempts to extrapolate dipoles $\mu_{z,x}$ for the aug-cc-pCV(X = Q,5)Z basis sets with formula $\mu_{x,z} = \mu_{x,z}^{cbs} + be^{-x}$ result in failure. The best fit to these dipoles is a linear function with χ^2 on the order of 10^{-12}, which is incorrect: energies can be extrapolated to their complete basis set limit and dipoles growing linearly with growth in basis set size cannot be extrapolated.

Many extrapolation schemes exist, and it is often suggested to extrapolate the Hartree-Fock energy component separately from the correlation contribution as they can converge at different rates.[44] However, for the data in Table II, the rate of convergence of their combined energy is described best by an exponential curve. Extrapolating the components independently with different schemes should not provide drastically different results on the dipoles; hence, we estimate the fit uncertainty on the transition intensities to be on the subpercent scale.

For the energy extrapolations, the fitted b variables are found to be equal for extrapolating E_{pz} and E_{nx} energies and also for E_{px} and E_{nz}. The $\mu_{z,x}$ dipoles for QZ, 5Z, and 6Z shown in Table II were computed within the program MOLPRO using energies that were converged to the tenth significant figure; hence, these dipoles are accurate to the sixth significant figure (dividing by 2λ is equivalent to multiplying by 10 000).The individual energies, E_{px}, E_{nx}, E_{pz} and E_{nz}, were written out to only eight significant figures; hence, the extrapolated dipoles that use these energies are instead accurate to the fourth significant figure.

Table III shows eight selected molecular configurations and presents the dipole values calculated for each of the five theoretical methods with the average CPU time required per configuration. For the extrapolated surfaces, we combine the CPU time for each of the individual calculations required for the extrapolation to provide a total time. The CBS^{Q5} dipoles deviate by less than 10^{-4} a.u. from the CBS^{56}, which was not expected, yet these CBS^{Q5} dipoles only

TABLE II. Calculated dipole moment and four energies E_{px}, E_{nx}, E_{pz}, and E_{nz} with the molecular configuration consisting of $R_1 = 1.7$ a$_o$, $R_2 = 1.98$ a$_o$, and $\theta = 163°$ computed with basis sets aug-cc-pCV(X = Q,5,6)Z. The calculated dipoles carry more precision than the energies shown here. QZ and 5Z energies are extrapolated with $E = E_{cbs} + be^{-x}$, and the respective CBS^{Q5} dipoles are calculated.

|	QZ	5Z	6Z	CBS^{Q5}	b
E_{pz}	$-76.416\,957\,07$	$-76.425\,045\,17$	$-76.428\,138\,05$	$-76.429\,752\,26$	$0.698\,593\,47$
E_{nz}	$-76.416\,979\,09$	$-76.425\,067\,22$	$-76.428\,160\,13$	$-76.429\,774\,32$	$0.698\,596\,06$
μ_z	$0.220\,221$	$0.220\,487$	$0.220\,746$	$0.220\,675$	$\ldots$
E_{px}	$-76.416\,975\,14$	$-76.425\,063\,27$	$-76.428\,156\,17$	$-76.429\,770\,37$	$0.698\,596\,06$
E_{nx}	$-76.416\,961\,02$	$-76.425\,049\,12$	$-76.428\,142\,01$	$-76.429\,756\,20$	$0.698\,593\,47$
μ_x	$-0.141\,258$	$-0.141\,482$	$-0.141\,600$	$-0.141\,675$	$\ldots$

TABLE III. (Upper) Water dipoles, in a.u., for *ab initio* calculations with different basis sets/basis set extrapolations for a range of molecular configurations. Energies (E) are given relative to the equilibrium geometry, point 1. The average CPU time required to calculate a single point is also provided. (Lower) Average weighted calculated to measured intensity ratios for selected vibrational bands. The experimental data are due to Birk *et al.*[25,40,43] and Loos *et al.*[41]

No.	R_1	R_2	θ	E (cm^{-1})	QZ	5Z	CBS^{Q5}	6Z	CBS^{56}
1	1.814 1	1.814 1	104.52	0	0.726 7	0.728 2	0.729 1	0.728 7	0.729 0
2	1.68	1.98	74.00	6 190.43	0.887 0	0.889 0	0.890 2	0.889 9	0.890 4
3	1.94	1.94	59.00	13 166.24	0.943 2	0.945 4	0.946 7	0.946 2	0.946 7
4	2.10	2.18	166.00	22 472.55	0.179 3	0.179 8	0.180 1	0.180 0	0.180 1
5	2.10	3.00	105.00	28 438.03	0.686 4	0.689 3	0.691 0	0.690 4	0.691 0
6	1.60	3.50	95.00	37 444.38	0.676 8	0.679 2	0.680 6	0.680 2	0.680 8
7	2.00	4.00	80.00	39 816.63	0.669 0	0.669 1	0.670 4	0.670 1	0.670 7
8	2.80	1.85	30.00	41 035.46	0.990 0	0.992 4	0.993 8	0.993 4	0.994 0
		CPU TIME (s)			15 190	27 684	42 874	84 715	112 399
	$(\nu_1, \nu_2, \nu_3)'$-$(\nu_1, \nu_2, \nu_3)''$								
	(100)-(000)				8.42	4.74	1.99	2.83	1.48
	(010)-(000)				0.67	0.95	1.16	1.11	1.23
	(001)-(000)				2.44	1.53	0.98	0.89	0.52
	(020)-(000)				2.17	1.32	0.70	0.73	0.44
	(121)-(000)				1.25	1.71	1.78	0.80	0.62

require less than 50% of the CPU time to compute than the aug-cc-CV6Z calculations. The marginal difference between the CBS^{Q5} and CBS^{56} dipoles highlights the correlation between these basis sets. This marginal deviation also holds true for those dipoles computed for geometries with a large proportion of stretch and/or bend (see Table III), which means that the technique holds true for all regions of the DMS. This result implies that highly accurate, global dipole surfaces can now be calculated with lower levels of theory at a fraction of the CPU cost, with marginal loss of precision. This should prove to be important for molecules that are too large to compute large grids of energies with an aug-cc-pCV6Z basis set.

In Table III, the CBS^{56} dipoles are consistently larger than those calculated with the aug-cc-pCV6Z basis set and should be of a comparable magnitude to those calculations expected with a aug-cc-pCV7Z basis set, if the smooth exponential growth of the dipole remains consistent. Obtaining these CBS^{56} dipoles will be computationally expensive for all molecules, as calculations will be required at both the aug-cc-pCV5Z and aug-cc-pCV6Z levels of theory. This will be limited to the lighter molecules.

Birk *et al.*[25] analyzed transition intensities of H_2O in the IR[40,41,43] and found the most recent (at the time) *ab initio* line list to deviate from the experimental observations[41] in the ν_1 fundamental
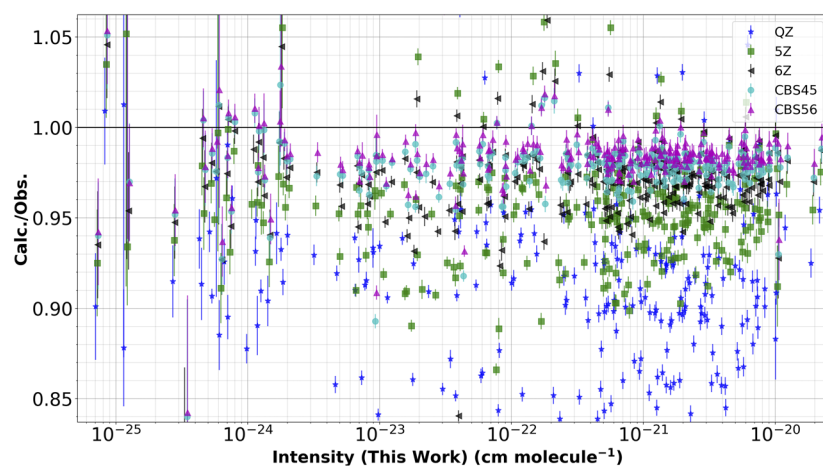
FIG. 1. Ratios of calculated ν_1 transition intensities obtained from each dipole moment surface against experimental measurements of Birk *et al.*[40] and Loos *et al.*[41] The error bars are experimental.
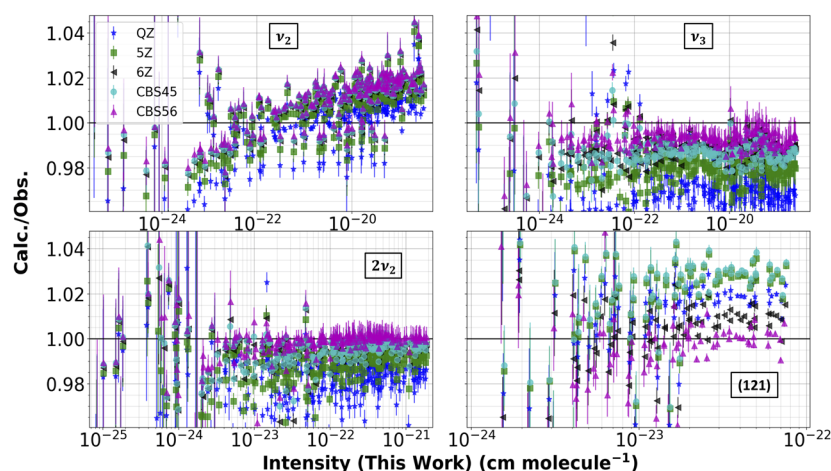
FIG. 2. Ratios of intensities calculated using DMS from Table I to the experimental measurements of Birk et al.[25,40,43] and Loos et al.,[41] shown for ν_3,[40,41] ν_2,[43] $2\nu_2$,[41] and (121).[25]

by 3%–15%. This *ab initio* line list was calculated with the LTP2011S DMS of Lodi et al.[30] Despite this large residual, all other bands in the 3000–4400 cm^{-1} region showed good agreement with experiment, suggesting the presence of an issue in the theoretical model. We recently calculated a new line list with the CKAPTEN DMS, and there were a large number of improvements over the recent POKAZATEL line list,[17] particularly for those transitions with wavelengths below 1 μm ($\omega > 10\,000$ cm^{-1}). The same conclusion was made for the isotopologs. The POKAZATEL line list was computed with a variation of the LTP2011 DMS dipole surface from Lodi et al., termed LTP2011S, which utilized fewer parameters for improved stability in highly energetic regions. However, our calculated intensities in the ν_1 fundamental showed no signs of improvement over POKAZATEL nor what was computed in 2011. Over 16 000 aug-cc-pCV6Z dipoles underlay this CKAPTEN surface; hence, discrepancies in this band are not due to the number of points fit nor is it a fitting issue as different functional forms were used for each of these models. These line lists used rotation-vibration wavefunctions that were also calculated with a variety of potential energy

surfaces; thus, these are also not the source of the deviation, although any energy dependence in the residuals will be due to the potential.[30]

Figure 1 plots the ratios of our calculated transition intensities in ν_1 to those measurements from Loos et al.[41] As the level of theory increases from QZ through to CBS56, the deviation in the ν_1 fundamental reduces from 2.83% with the 6Z DMS to only 1.48% with the new CBS56 DMS (see Table III). Indeed, the spectral intensities calculated with the CBS^{Q5} DMS are closer to the experimental values of Loos et al.[41] than the aug-cc-pCV6Z calculations. This suggests that better results can potentially be obtained by extrapolating the computationally less expensive calculations. To verify that the extrapolation technique works on a global scale, we need to also investigate other infrared bands measured by Loos et al. and Birk et al.[40,43] Ratios of transition intensities in ν_3, ν_2, and $2\nu_2$ are presented in Fig. 2. In addition, Table III presents a general overview of several other bands, which we do not present in our figures.

For the ν_3 and $2\nu_2$ bands, shown in Fig. 2, the computed CBS^{56} spectrum is again closer to the experimental intensities than any of
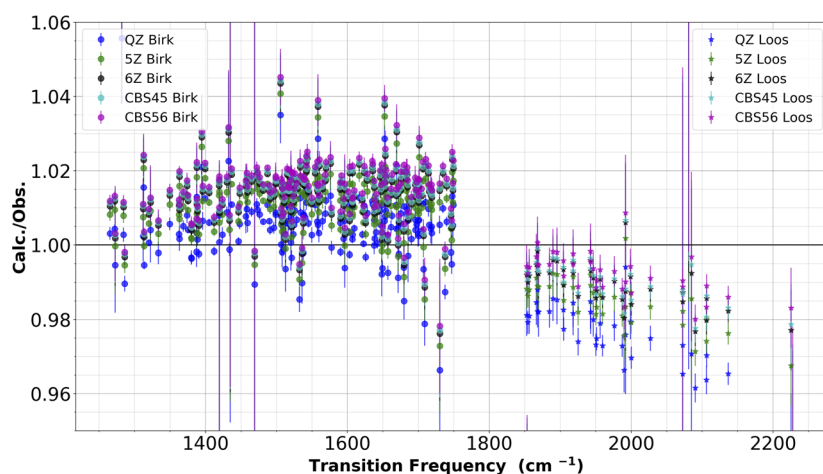

FIG. 3. Ratios of intensities calculated using the DMS from Table I to the experimental measurements of Birk et al.[40] and Loos et al.[41] for ν_2 transitions plotted as a function of frequency.

the other datasets. However, the ν_2 ratios suggest that the analysis of the experimental spectrum of Loos et al.[41] differs from that of Birk et al.[40] In Fig. 3, we plot intensity ratios for ν_2 transitions as a function of frequency and separate the datasets. Birk analyzed the spectrum with a speed-dependent Voigt profile, while Loos used a quadratic speed dependent hard collision profile with Rosenkranz line-mixing. Our results show that the choice of profile is very important for high-accuracy intensity measurements.

For more energetic transitions, wavelengths shorter than 1 μm, the CBS[56] intensities show excellent agreement with the measurements of Birk et al.[25] (see Fig. 2). For the same band, the calculated aug-cc-pCVQZ spectrum is closer to the experimental data than the CBS^{Q5} and aug-cc-pCV5Z spectra, which is counterintuitive. We also expect the CBS^{Q5} spectrum to be closer to experiment than the aug-cc-pCV5Z spectrum, which is also not the case. The results indicate there may be an underlying issue regarding the aug-cc-pCVQZ dataset in this band.

IV. CONCLUSION

We compute *ab initio* dipoles with the MRCI procedure embedded within the quantum chemistry package MOLPRO[36] with the aug-cc-pCV(X = Q,5,6)Z basis sets[26–28,36] for water vapor. Using these results, we similarly created extrapolated CBS^{Q5} and CBS^{56} dipoles, each obtained by extrapolating energies from the aug-cc-pCVQZ and aug-cc-pCV5Z calculations and likewise from the aug-cc-pCV5Z and aug-cc-pCV6Z levels of theory. The same functional form was fit to these dipole sets to create five individual dipole moment surfaces.

For a select number of global geometries studied, the extrapolated CBS^{Q5} and CBS^{56} dipoles are within 10^{-4} a.u. of each other. This finding does reflect the correlation in the aug-cc-pCV (X = Q,5,6) basis sets. The greatest achievement surrounds the computational time required for calculating the CBS^{Q5} dipoles: they require 50% less CPU time compared to the aug-cc-pCV6Z dipoles.

Using the potential energy surface of Mizus et al., known as PES15K, which predicts energy levels below 15 000 cm^{-1} to a root mean square of 0.011 cm^{-1}, we create wavefunctions up to total quantum number $J = 6$. With these, we calculate spectra for each of the dipole surfaces with an upper threshold of 15 000 cm^{-1} and where possible, we replace calculated energy levels with those in the MARVEL database.

The transition intensities obtained from each dataset are compared against the experimental measurements of Birk et al. and Loos et al., which are both in the IR. Overall, the CBS^{56} DMS, despite being fit with dipoles that are only accurate to the fourth significant figure, resulting intensities exhibit the best agreement with the measurements, and for several bands, the deviation is in the sub-one-percent range. Also, comparing the CBS^{56} and aug-cc-pCV6Z results indicates that it may be preferable, in cases, to consider using the CBS^{56} dipoles over the aug-cc-pCV6Z dipoles.

For those strong transitions in the IR, calculated intensities obtained using the CBS^{Q5} and CBS^{56} DMS's suggest that the limit of dipole precision required for obtaining subpercent accuracy against high-quality experiments could be on the order of 10^{-4} a.u.

We show that the two point formula works very well for extrapolating dipoles, reducing the deviation from experimental measurements by approximately one percent in certain bands, and thus, we

estimate that the application of a three-point formula for dipole extrapolation could potentially improve the accuracy by another 0.5%.

Coupled-cluster methods are, in general, computationally less expensive than MRCI but, nonetheless, provide accurate results for energies and hence finite field dipoles for geometries close to equilibrium.[45–47] The extrapolation technique is successful not because of the underlying method of calculation, in this case MRCI, but because of the nature of the aug-cc-pCV(X = Q,5,6)Z basis sets. We, therefore, expect the extrapolation technique to work for coupled-cluster calculations, providing the calculations remain converged. Applying this extrapolation technique to high-level coupled-cluster energies may prove to be a cheaper alternative of obtaining mid-/lower-level MRCI results.

REFERENCES

[1] P. Zoogman, X. Liu, R. M. Suleiman, W. F. Pennington, D. E. Flittner, J. A. Al-Saadi, B. B. Hilton, D. K. Nicks, M. J. Newchurch, J. L. Carr et al., J. Quant. Spectrosc. Radiat. Transfer **186**, 17 (2017).

[2] K. Weigel, A. Rozanov, F. Azam, K. Bramstedt, R. Damadeo, K.-U. Eichmann, C. Gebhardt, D. Hurst, M. Kraemer, S. Lossow et al., Atmos. Meas. Tech. **9**, 133 (2016).

[3] P. F. Levelt, J. Joiner, J. Tamminen, J. P. Veefkind, P. K. Bhartia, D. C. Stein Zweers, B. N. Duncan, D. G. Streets, H. Eskes, R. van der A et al., Atmos. Chem. Phys. **18**, 5699 (2018).

[4] S. T. Gulde, M. G. Kolm, D. J. Smith, R. Maurer, G. B. Courrèges-Lacoste, M. Sallusti, and G. Bagnasco, Proc. SPIE **10563**, 1056341 (2017).

[5] S. Noël, M. Buchwitz, H. Bovensmann, R. Hoogen, and J. P. Burrows, Geophys. Res. Lett. **26**, 1841, https://doi.org/10.1029/1999gl900437 (1999).

[6] A. Kuze, H. Suto, M. Nakajima, and T. Hamazaki, Appl. Opt. **48**, 6716 (2009).

[7] D. Crisp and C. Johnson, Acta Astronaut. **56**, 193 (2005).

[8] I. E. Gordon et al., J. Quant. Spectrosc. Radiat. Transfer **203**, 3 (2017).

[9] S. Vasilchenko, H. Tran, D. Mondelain, S. Kassi, and A. Campargue, J. Quant. Spectrosc. Radiat. Transfer **235**, 332 (2019).

[10] S. Mikhailenko, D. Mondelain, E. Karlovets, S. Kassi, and A. Campargue, J. Quant. Spectrosc. Radiat. Transfer **206**, 163 (2018).

[11] D. Mondelain, S. Mikhailenko, E. Karlovets, S. Bacguier, S. Kassi, and A. Campargue, J. Quant. Spectrosc. Radiat. Transfer **203**, 206 (2017).

[12] X. Huang, D. W. Schwenke, S. A. Tashkun, and T. J. Lee, J. Chem. Phys. **136**, 124311 (2012).

[13] I. I. Bubukina, O. L. Polyansky, N. F. Zobov, and S. N. Yurchenko, Opt. Spectrosc. **110**, 160 (2011).

[14] I. I. Mizus, A. A. Kyuberis, N. F. Zobov, V. Y. Makhnev, O. L. Polyansky, and J. Tennyson, Philos. Trans. R. Soc., A **376**, 20170149 (2018).

[15] S. N. Yurchenko, M. Carvajal, P. Jensen, F. Herregodts, and T. R. Huet, Chem. Phys. **290**, 59 (2003).

[16] M. J. Down, J. Tennyson, M. Hara, Y. Hatano, and K. Kobayashi, J. Mol. Spectrosc. **289**, 35 (2013).

[17] O. L. Polyansky, A. A. Kyuberis, N. F. Zobov, J. Tennyson, S. N. Yurchenko, and L. Lodi, Mon. Not. R. Astron. Soc. **480**, 2597 (2018).

[18] D. W. Schwenke, J. Phys. Chem. A **105**, 2352 (2001).

[19] O. L. Polyansky, A. G. Császár, S. V. Shirin, N. F. Zobov, P. Barletta, J. Tennyson, D. W. Schwenke, and P. J. Knowles, Science **299**, 539 (2003).

[20] H. M. Quiney, P. Barletta, G. Tarczay, A. G. Császár, O. L. Polyansky, and J. Tennyson, Chem. Phys. Lett. **344**, 413 (2001).

[21] O. L. Polyansky, R. I. Ovsyannikov, A. A. Kyuberis, L. Lodi, J. Tennyson, and N. F. Zobov, J. Phys. Chem. A **117**, 9633–9643 (2013).

[22] S. Mikhailenko, E. Karlovets, S. Vasilchenko, D. Mondelain, S. Kassi, and A. Campargue, J. Quant. Spectrosc. Radiat. Transfer **236**, 106574 (2019).

[23] O. L. Polyansky, K. Bielska, M. Ghysels, L. Lodi, N. F. Zobov, J. T. Hodges, and J. Tennyson, Phys. Rev. Lett. **114**, 243001 (2015).

[24] E. J. Zak, J. Tennyson, O. L. Polyansky, L. Lodi, S. A. Tashkun, and V. I. Perevalov, J. Quant. Spectrosc. Radiat. Transfer **177**, 31 (2016).

[25] M. Birk, G. Wagner, J. Loos, L. Lodi, O. L. Polyansky, A. A. Kyuberis, N. F. Zobov, and J. Tennyson, J. Quant. Spectrosc. Radiat. Transfer **203**, 88 (2017).

[26] T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

[27] D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. **103**, 4572 (1995).

[28] K. A. Peterson and T. H. Dunning, J. Chem. Phys. **117**, 10548 (2002).

[29] E. K. Conway, A. A. Kyuberis, O. L. Polyansky, J. Tennyson, and N. Zobov, J. Chem. Phys. **149**, 084307 (2018).

[30] L. Lodi, J. Tennyson, and O. L. Polyansky, J. Chem. Phys. **135**, 034113 (2011).

[31] H. Partridge and D. W. Schwenke, J. Chem. Phys. **106**, 4618 (1997).

[32] D. W. Schwenke and H. Partridge, J. Chem. Phys. **113**, 6592 (2000).

[33] X. Huang, D. W. Schwenke, and T. J. Lee, J. Quant. Spectrosc. Radiat. Transfer **230**, 222 (2019).

[34] O. L. Polyansky, N. F. Zobov, I. I. Mizus, A. A. Kyuberis, L. Lodi, and J. Tennyson, J. Quant. Spectrosc. Radiat. Transfer **210**, 127 (2018).

[35] V. G. Tyuterev, A. Barbe, D. Jacquemart, C. Janssen, S. N. Mikhailenko, and E. N. Starikova, J. Chem. Phys. **150**, 184303 (2019).

[36] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 242 (2012).

[37] E. K. Conway, I. E. Gordon, A. A. Kyuberis, O. L. Polyansky, J. Tennyson, and N. F. Zobov, J. Quant. Spectrosc. Radiat. Transfer **241**, 106711 (2019).

[38] D. Feller, J. Chem. Phys. **98**, 7059 (1993).

[39] P. Barletta, S. V. Shirin, N. F. Zobov, O. L. Polyansky, J. Tennyson, E. F. Valeev, and A. G. Császár, J. Chem. Phys. **125**, 204307 (2006).

[40] M. Birk, J. Loos, G. Wagner, D. Mondelain, S. Kassi, A. Campargue, F. Hase, J. Orphal, A. Perrin, L. Coudert *et al.*, in Proceedings of the ESA Living Planet Symposium, Prague, Czech Republic, 2016.

[41] J. Loos, M. Birk, and G. Wagner, J. Quant. Spectrosc. Radiat. Transfer **203**, 103 (2017).

[42] J. Tennyson, M. A. Kostin, P. Barletta, G. J. Harris, O. L. Polyansky, J. Ramanlal, and N. F. Zobov, Comput. Phys. Commun. **163**, 85 (2004).

[43] M. Birk and G. Wagner, J. Quant. Spectrosc. Radiat. Transfer **170**, 159 (2016).

[44] T. Helgaker, W. Klopper, H. Koch, and J. Noga, J. Chem. Phys. **106**, 9639 (1997).

[45] K. L. Chubb, A. Yachmenev, J. Tennyson, and S. N. Yurchenko, J. Chem. Phys. **149**, 014101 (2018).

[46] P. A. Coles, S. N. Yurchenko, R. P. Kovacich, J. Hobby, and J. Tennyson, Phys. Chem. Chem. Phys. **21**, 3264 (2019).

[47] A. Owens, A. Yachmenev, W. Thiel, A. Fateev, J. Tennyson, and S. N. Yurchenko, Mon. Not. R. Astron. Soc. **479**, 3002 (2018).