

# Comparison of the performance of local, gradient-corrected, and hybrid density functional models in predicting infrared intensities

Mathew D. Halls and H. Bernhard Schlegel<sup>a)</sup>

Department of Chemistry, Wayne State University, Detroit, Michigan 48202-3489

(Received 22 June 1998; accepted 18 September 1998)

Density functional theory has been used to calculate infrared (IR) intensities for a series of molecules (HF, CO, H<sub>2</sub>O, HCN, CO<sub>2</sub>, C<sub>2</sub>H<sub>2</sub>, H<sub>2</sub>CO, NH<sub>3</sub>, C<sub>2</sub>H<sub>4</sub>, CH<sub>2</sub>F<sub>2</sub>, CH<sub>2</sub>Cl<sub>2</sub>, and CH<sub>2</sub>Br<sub>2</sub>) in an effort to evaluate relative performance of different functionals. The methods employed in this study comprise most of the popular local, gradient-corrected, and hybrid functionals, namely, S-VWN, S-PL, B-LYP, B-P86, B-PW91, B3-LYP, B3-P86, and B3-PW91. Calculations were carried out using various qualities of split valence basis sets augmented with diffuse and polarization functions, both to determine basis set dependence and to evaluate the limit performance. Computed intensities were compared with results from conventional correlated *ab initio* methods (MP2 and QCISD). Hybrid functionals give results in closest agreement with QCISD over the other methods surveyed. Local and gradient-corrected methods performed remarkably alike, both are comparable to MP2, and outperform Hartree–Fock. Hartree–Fock intensities can be dramatically improved by scaling, making them similar to MP2 results. © 1998 American Institute of Physics. [S0021-9606(98)30748-5]

## I. INTRODUCTION

Vibrational spectroscopy is one of the most powerful and widely used tools in modern materials research. Quantum chemical calculations predicting harmonic frequencies and spectral intensities are essential when interpreting experimental spectra, particularly for large molecules where the high density of states results in spectral complexity in the region below 1600 cm<sup>-1</sup>.<sup>1,2</sup> For aromatic systems, the upper part of the region generally contains the in-plane molecular vibrations and the mid- to lower part contains the out-of-plane vibrations. Accurate vibrational assignments for such systems are necessary for materials characterization purposes; consider, for example, the postdeposition characterization of molecular films where orientation data is extracted from band intensities.<sup>3</sup> Assignments for complex systems can be proposed on the basis of frequency agreement between the computed harmonics and the observed fundamentals; however, more reliable assignments can be made by making use of predicted intensities as well, which can definitively resolve the assignments of closely spaced bands.<sup>4</sup> Other areas in which successful prediction of infrared intensities would benefit are well known, such as the determination of analyte species in difficult regimes. Thus, researchers would like computational models to provide accurate spectral intensities in addition to vibrational frequencies.

Density functional theory (DFT) is becoming increasingly popular among experimentalists and theoreticians in the chemical literature.<sup>5,6</sup> Numerous reports have been made citing the successes of DFT compared to conventional *ab initio* methods, in computing molecular and chemical properties such as geometries, harmonic frequencies, and

energies.<sup>7–10</sup> Density functional theory is superior to the conventional methods, such as HF (Hartree–Fock) and MP2 (second-order Møller–Plesset perturbation theory), for the calculation of polyatomic vibrational frequencies, with empirical scaling factors approaching unity.<sup>8,9</sup> Experience with DFT models has shown that DFT satisfactorily recovers electron correlation, predicting numerous properties in similar or better agreement with experimental data than conventional correlated *ab initio* methods, such as MP2. The cost effectiveness of DFT over conventional methods is an additional attractive feature.

The present paper concerns the performance of DFT in predicting infrared absorption intensities. Reports have been made in the literature on the performance of DFT in predicting dipole moment derivatives and infrared intensities in the double-harmonic approximation.<sup>10–14</sup> These studies have usually examined the results for a small number of model molecules, using a few levels of theory (exchange functional, correlation functional, and a fixed quality of basis set). In the present work DFT infrared intensities are examined in an effort to present an overview of the relative performance of various commonly used local, gradient-corrected, and hybrid density functionals, namely, S-VWN, S-PL, B-LYP, B-P86, B-PW91, B3-LYP, B3-P86, and B3-PW91. These DFT models are used to compute infrared intensities for a set of test molecules (HF, CO, H<sub>2</sub>O, HCN, CO<sub>2</sub>, C<sub>2</sub>H<sub>2</sub>, H<sub>2</sub>CO, NH<sub>3</sub>, C<sub>2</sub>H<sub>4</sub>, CH<sub>2</sub>F<sub>2</sub>, CH<sub>2</sub>Cl<sub>2</sub>, and CH<sub>2</sub>Br<sub>2</sub>) and the results are compared with those obtained from Hartree–Fock (HF) and two conventional correlated *ab initio* methods: second-order Møller–Plesset theory (MP2) and quadratic configuration interaction including single and double substitutions (QCISD). Experimental absolute infrared intensities are notoriously difficult to measure and in the literature there are often significant differences between values reported by different researchers. To facilitate evaluation in this work, we have

<sup>a)</sup> Author to whom all correspondence should be made. Electronic mail: hbs@chem.wayne.edu

adopted the QCISD level results as our benchmark for accuracy, and restricted our calculations to the double harmonic approximation.

The prediction of infrared intensities is difficult because of the need to describe the tail region of the wave function,<sup>15–17</sup> which requires large basis sets, and because of the need to include electron correlation. Studies on the effect of electron correlation on computed infrared intensities have appeared in the literature.<sup>18–20</sup> Electron correlation was found to be essential in obtaining quantitative results in best agreement with experimental intensities. The agreement between predicted and observed intensities improved significantly over the self-consistent field (SCF) results upon inclusion of electron correlation at even minimal levels. The general trend concerning correlation in predicting infrared intensities is with increasing electron correlation the theoretical intensities converge toward experimental values.<sup>20</sup> Quadratic configuration interaction methods represent treatment of electron correlation beyond the many-body perturbation approach (MPn).

Comparisons are made at the standard spectroscopic basis set level (6-31G\*) and using a large basis set [6-311+G(3df,3pd)] to assess limit performance. Basis set dependence is investigated for local and gradient-corrected DFT. This report presents an overview of the performance of these methods. Complete tables of the numeric results are provided as supplementary material.<sup>21</sup>

## II. METHODS

The calculations were performed using the GAUSSIAN 94 suite of programs.<sup>22</sup> Following full geometry optimizations, harmonic vibrational frequencies and dipole moment derivatives were calculated analytically for the Hartree–Fock, second-order Moller–Plesset theory (MP2, frozen core<sup>23</sup>), and DFT methods while numerical differentiation was used for the quadratic configuration interaction with single and double substitution (QCISD, frozen core<sup>24</sup>) method. The infrared intensities were calculated from the dipole moment derivatives in the double harmonic approximation.<sup>25</sup> The vibrational properties were computed for the test molecules in their ground state enforcing their respective point group symmetry. The functionals employed in this work consist of the following local functionals: S-VWN and S-PL; corresponding to the Slater–Dirac exchange functional (S)<sup>26</sup> with the Vosko–Wilk–Nusair fit for the correlation functional (VWN)<sup>27</sup> and with the Perdew local fit for the correlation functional (PL).<sup>28</sup> The gradient-corrected functionals used were B-LYP, B-P86, and B-PW91, corresponding to Becke’s gradient-corrected exchange functional (B)<sup>29</sup> with the Lee–Yang–Parr gradient-corrected correlation functional (LYP),<sup>30</sup> with Perdew’s gradient-corrected correlation functional (P86),<sup>31</sup> and with Perdew and Wang’s gradient-corrected correlation functional (PW91).<sup>32</sup> The following hybrid functionals were employed: B3-LYP, B3-P86, and B3-PW91; corresponding to Becke’s three-parameter exchange functional (B3)<sup>33</sup> with the Lee–Yang–Parr gradient-corrected correlation functional (LYP), with Perdew’s gradient-corrected correlation functional (P86), and with Perdew and Wang’s gradient-corrected correlation functional

(PW91). These methods are assessed with two basis sets, the standard spectroscopic basis set to evaluate working performance which is split-valence augmented with one set of polarization functions on heavy atoms, 6-31G\* [also designated 6-31G(d)],<sup>34,35</sup> and a large basis set to evaluate limit performance using a split valence basis of triple zeta quality, 6-311G<sup>36,37</sup> augmented with one set of diffuse functions on heavy atoms, three additional *d*-type functions and one *f*-type function on heavy atoms, and three *p*-type functions and one *d*-type on the hydrogens, 6-311+G(3df,3pd). The results for each of the two basis sets are compared against QCISD calculations with the same basis set. To assess the basis set dependence for DFT methods in predicting infrared intensities, two archetype functionals of the local and gradient-corrected type, S-VWN and B-LYP, were tested by systematically adding diffuse<sup>38</sup> and polarization functions<sup>39</sup> to a split valence double zeta basis set, 6-31G. These results are compared with QCISD intensities obtained for the large basis set used in this study.

## III. RESULTS AND DISCUSSION

### A. Comparison of QCISD with experiment

Experimental and theoretical infrared intensities have been compared in a number of previous studies.<sup>10–13,16,18–20</sup> Difficulties such as overlapping bands, resonances, intensity sharing, etc. typically limit the accuracy of experimental absolute infrared intensity measurements to  $\pm 10\%$ .<sup>20,40</sup> Theoretical calculations of infrared intensities are usually carried out in the double harmonic approximation, i.e., ignoring cubic and higher force constants and omitting second and higher dipole moment derivatives. This can affect bond stretching modes and large-amplitude bending modes, adding perhaps  $\pm 5\%$  to  $\pm 10\%$  to the uncertainty of the calculated intensities, even in the absence of resonances. It has been known for some time that fairly large basis sets are needed for a proper description of IR intensities.<sup>16,19</sup> Schaefer and co-workers<sup>20</sup> have shown that as the level of electron correlation is increased, the predicted IR intensities converge toward the experimental values. For a subset of the molecules considered in this study, they found that coupled cluster methods with polarized triple zeta basis sets CCSD/TZ2P and CCSD(T)/TZ(2df,2pd) provide a balanced treatment of the vibrational frequencies and IR intensities. With the exception of the asymmetric stretch in CO<sub>2</sub>, the contributions from triple excitations are small, averaging ca.  $\pm 3.1$  km mol<sup>-1</sup>.<sup>41</sup> Bartlett and co-workers<sup>42</sup> have demonstrated that contributions from core electron correlation are also small.

The best calculations in the present work use the QCISD method, which is comparable to CCSD, and employ a basis set that is slightly larger than those used in most previous studies. In Fig. 1, the infrared intensities calculated at the QCISD/6-311+G(3df,3pd) and CCSD(T)/TZ(2df,2pd) levels are compared with experimental data. For QCISD versus experiment, the average difference, average absolute difference and standard deviation are 6.21, 12.73, and 30.65 km mol<sup>-1</sup>; the corresponding values for CCSD(T) versus experiment are 8.42, 10.83, and 20.30 km mol<sup>-1</sup>.

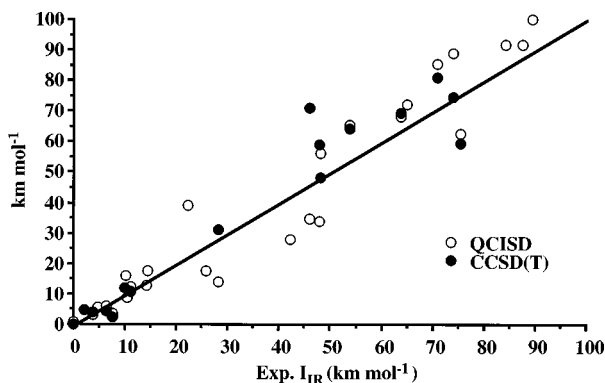


FIG. 1. Comparison of QCISD/6-311+G(3df,3pd) and CCSD(T)/TZ(2df,2pd) infrared intensities with experimental values. CCSD(T) intensities taken from Ref. 20.

Comparison of the QCISD and CCSD(T) results yields an average difference of  $0.88 \text{ km mol}^{-1}$  and an average absolute difference of  $4.03 \text{ km mol}^{-1}$  (omitting the problematic  $\text{CO}_2$  asymmetric stretch). The two different high-level theoretical methods agree quite well with each other, but both show larger discrepancies in comparisons with experiment. This suggests that a large fraction of the difference between the calculated and the observed intensities may be attributed to the double harmonic approximation and to the experimental uncertainties. To avoid these ambiguities in the assessment of the DFT methods, we will compare the DFT calculations to our highest level *ab initio* results, since the DFT and *ab initio* methods should ultimately converge to the same limit within the double harmonic approximation.

## B. Standard spectroscopic basis set performance

Previous studies have shown that a split valence basis set of double-zeta quality plus one set of polarization functions on heavy atoms (6-31G\*) represents an economical choice yielding vibrational harmonic frequencies which can be easily scaled for comparison with observed fundamentals,<sup>8,9</sup> and infrared intensities which are qualitatively acceptable.<sup>16</sup> The need for polarization functions when predicting infrared intensities is attributed to the added flexibility they lend to the electronic wave function, allowing it to deform in the presence of an electric field.<sup>43-45</sup> The 6-31G\* basis has become the most often used basis set for predicting these quantities. The performance of DFT methods in predicting vibrational frequencies with this basis set has been studied by Wong<sup>8</sup> and Scott and Radom.<sup>9</sup> Their work has shown that DFT outperforms conventional *ab initio* methods (HF, MP2, and QCISD) in predicting frequencies in agreement with experiment, with homogeneous empirical scale factors approaching unity.<sup>46</sup>

Figure 2 shows the average absolute difference from QCISD results using the same basis set, and the standard deviation for each of the methods considered here. Hartree-Fock, not surprisingly, falls furthest from the QCISD results and also has the largest standard deviation. This is to be expected due to the absence of electron correlation. The two local methods perform very similarly to each other, with average absolute differences of ca.  $10 \text{ km mol}^{-1}$ . The gradient-

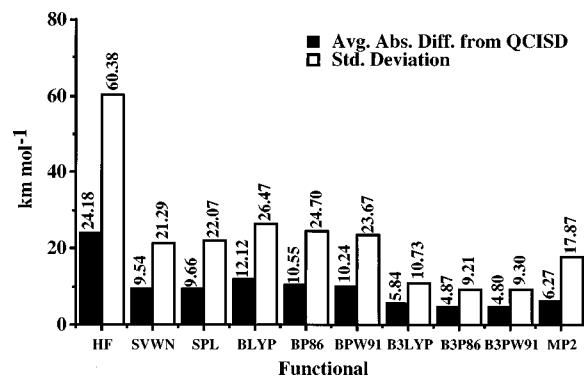


FIG. 2. Average absolute difference from QCISD and standard deviation with the 6-31G(d) basis set.

corrected functionals perform a bit worse than the local functionals having average absolute differences and standard deviations slightly greater than both local methods. Of the gradient-corrected functionals, the most often used combination, B-LYP, predicts intensities in somewhat greater difference from QCISD, than the other two gradient-corrected functionals. The best performers are the hybrid functionals having average absolute differences approximately half that of the local and gradient-corrected methods. The hybrid functionals also have the lowest standard deviation of all methods considered here. In the hybrid functional data, the trend is observed again that use of the Lee-Yang-Parr correlation functional slightly increases both the average absolute difference and the standard deviation, compared with the other two hybrid functionals, although the effect is small. The performance of MP2 is between the local and hybrid functionals, having an average absolute difference slightly larger than the hybrid methods but a standard deviation a little smaller than the local methods. This is an important note since the quality of MP2 results is often overestimated in the chemical literature and its results come at a much greater cost in comparison to DFT.

The normal mode intensity for which all methods considered here predicted large deviations from the QCISD calculations is the asymmetric stretch in  $\text{CO}_2$ . This is probably due to the neglect of triple excitations in the QCISD calculations.<sup>41</sup> Other vibrations that posed problems are the asymmetric CH stretch of  $\text{H}_2\text{CO}$ , and the CX stretches of  $\text{CH}_2\text{Cl}_2$  and  $\text{CH}_2\text{Br}_2$ . These vibrations are unique in that the Hartree-Fock and MP2 results for these vibrations are in better agreement with QCISD than the DFT intensities, including the hybrid methods. Analysis of the gradient-corrected and hybrid DFT intensities shows that use of the Lee-Yang-Parr correlation functional tends to predict intensities that are either in close agreement with, or somewhat greater than, those computed by the methods employing the P86 or PW91 correlation functionals.

Table I gives two measures of deviation from QCISD results: the average difference  $\sum\{I - I_{\text{QCISD}}\}/n$  and the average absolute difference  $\sum|I - I_{\text{QCISD}}|/n$ , where  $n$  is the number of computed intensities. The data there clearly shows that Hartree-Fock almost always overestimates infrared intensities, with the average difference from QCISD and the aver-

TABLE I. Average difference and average absolute difference from QCISD infrared intensities with the 6-31G(*d*) and 6-311+G(3*df*,3*pd*) basis set in km mol<sup>-1</sup>.

	HF	HF Scaled <sup>b</sup>	SVWN	SPL	BLYP	BP86	BPW91	B3LYP	B3P86	B3PW91	MP2
6-31G( <i>d</i> )											
Average difference	23.84	-3.42	-0.56	-0.81	-1.91	-0.61	-0.59	2.18	3.50	3.35	-1.87
Average absolute difference	24.18	11.05	9.54	9.66	12.12	10.55	10.24	5.84	4.87	4.80	6.27
6-311+G(3 <i>df</i> ,3 <i>pd</i> )											
Average difference	26.88	-3.16	-0.97	-0.97	-3.70	-3.95	-4.17	2.54	2.18	2.02	-3.86
Average absolute difference	27.14	9.35	10.90	10.88	10.57	11.12	11.10	5.04	5.38	5.28	8.49
Relative CPU time <sup>a</sup>	1.4	1.4	1.0	1.0	1.7	1.7	1.7	1.9	1.9	1.9	4.4

<sup>a</sup>Relative CPU time for a frequency job for CH<sub>2</sub>F<sub>2</sub> with the 6-311+G(3*df*,3*pd*) basis set.

<sup>b</sup>HF/6-31(*d*) intensities scaled by 0.63 and HF/6-311+G(3*df*,3*pd*) intensities scaled by 0.68.

age absolute difference from QCISD being equal within 1 km mol<sup>-1</sup>. From the table the methods resolve themselves into groups with similar performance. Hartree-Fock gives results in worst agreement with QCISD. The local DFT methods and the gradient-corrected DFT methods perform remarkably similar, perhaps countering the common expectation that the gradient-corrected results should be significantly better than the local results. The use of the Becke three-parameter hybrid functional with any of the gradient-corrected correlation functionals achieves results which are in excellent agreement with QCISD. By these measures MP2 performs between the local/gradient-corrected DFT level and the hybrid DFT level. With the 6-31G\* basis set the order of best agreement with QCISD is hybrid DFT > MP2 > local DFT ≈ gradient-corrected DFT > HF. Compared to experiment, the hybrid DFT methods have an average absolute difference of ca. 16 km mol<sup>-1</sup>.

### C. Basis set effects for local and gradient-corrected DFT

The quality of basis set needed to predict converged chemical properties is of great interest to researchers seeking to maximize the quality of results for the computational effort expended. Since polarized split valence basis sets of double-zeta quality (e.g., 6-31G\*) represent the middle ground in terms of cost and quality, we were interested in investigating the convergence of predicted infrared intensities at this level. The benchmark we chose here was the QCISD method employing a large basis set of triple zeta quality augmented with one set of diffuse functions, three sets of *d*-type functions, and one set of *f*-type functions on the heavy atoms, and three sets of *p*-type functions and one set of *d*-type functions on hydrogens [6-311+G(3*df*,3*pd*)], to ensure results approaching the QCISD limit. Here we present the results for the set of test molecules (minus CH<sub>2</sub>Cl<sub>2</sub> and CH<sub>2</sub>Br<sub>2</sub>)<sup>47</sup> of systematically increasing the quality of the double-zeta basis set by adding additional diffuse and polarization functions.

Figure 3 shows the average absolute difference and standard deviation from the QCISD results for the local DFT method S-VWN. Interesting to note is that the standard spectroscopic basis set shows good performance. The removal of a set of *d*-type functions on the heavy atoms from the 6-31G\* basis set increases the average absolute difference by a factor of 2 and increases the standard deviation by ca. 100

km mol<sup>-1</sup>. For the 6-31G\* basis set, addition of a set of diffuse functions on the heavy atoms decreases the average difference by ca. 1 km mol<sup>-1</sup>. Adding polarization functions to hydrogens decreases the average absolute difference by about another 1 km mol<sup>-1</sup>. This trend continues showing a small sensitivity to additional polarization functions on heavy atoms, dropping the difference by ca. 3 km mol<sup>-1</sup> each time, and an even lower sensitivity to additional polarization functions on hydrogens. In terms of convergence the average absolute difference seems settled with the 6-31+G(3*d*,2*p*) basis set at 12 km mol<sup>-1</sup> and a standard deviation of 30 km mol<sup>-1</sup>. The best agreement for S-VWN with an augmented double-zeta basis set seems to be an average absolute difference of 12 km mol<sup>-1</sup> with a standard deviation of 30 km mol<sup>-1</sup>. The standard working basis set, 6-31G\*, gives results with an average absolute difference of 20 km mol<sup>-1</sup>, less than 10 km mol<sup>-1</sup> from the converged limit and therefore seems to be an excellent compromise between accuracy and cost for infrared intensity calculations.

Figure 4 shows average absolute difference data and standard deviation data for the gradient-corrected B-LYP method. With this method the 6-31G\* basis set performs less well. The results for B-LYP show increased sensitivity over S-VWN to the addition of diffuse and polarization functions on both heavy atoms and hydrogen. The addition of diffuse functions on heavy atoms to the 6-31G\* basis set decreases

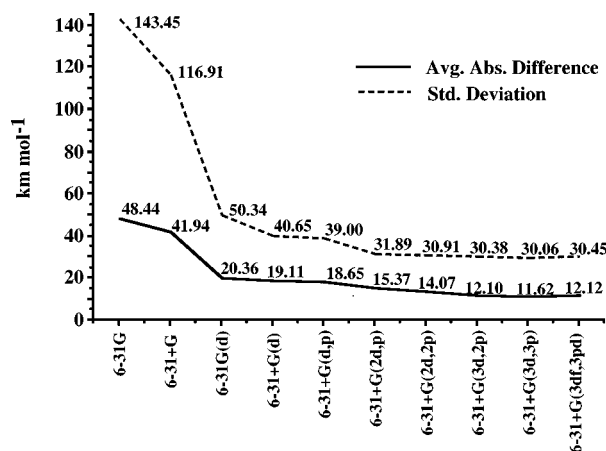


FIG. 3. Basis set dependence for S-VWN of the average absolute difference and standard deviation from QCISD/6-311+G(3*df*,3*pd*) infrared intensities.

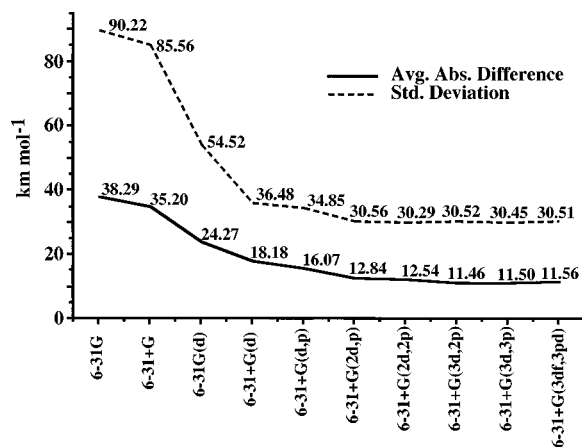


FIG. 4. Basis set dependence for B-LYP of the average absolute difference and standard deviation from QCISD/6-311+G(3df,3pd) infrared intensities.

the average absolute difference of ca. 6 km mol<sup>-1</sup>. Adding a set of *p*-type polarization functions to hydrogens improves the average agreement by 2 km mol<sup>-1</sup>. After the addition of the first sets of diffuse and polarization functions, the sensitivity to further additions is diminished. The addition of second and third sets of polarization functions on heavy atoms and hydrogens show similar insensitivity as with the S-VWN method. The converged average absolute difference for B-LYP is 11 km mol<sup>-1</sup> with a standard deviation of 30 km mol<sup>-1</sup>. In order to get within 10 km mol<sup>-1</sup> average absolute difference of the converged result, a larger basis set is required than needed with S-VWN. The B-LYP functional requires diffuse as well as *d* polarization on heavy atoms to meet this criterion (i.e., 6-31+G\*).

#### D. Large basis set performance

Model chemistries used to predict molecular properties are usually characterized by limitations in two dimensions: incomplete treatment of electron correlation and basis set truncation.<sup>48</sup> Using the largest feasible basis set allows better assessment of the method itself, assuming the appropriate convergence of the property of interest. In an effort to evaluate the performance of DFT in predicting infrared intensities at the large basis set limit, calculations for the set of test molecules (minus CH<sub>2</sub>Cl<sub>2</sub> and CH<sub>2</sub>Br<sub>2</sub>)<sup>47</sup> were carried out using the 6-311+G(3df,3pd) basis set.

The average absolute differences from QCISD at this basis level are shown in Fig. 5. As with the working spectroscopic basis set, 6-31G\*, Hartree–Fock has the largest absolute difference and standard deviation. The differences from QCISD for each DFT method are remarkably conserved from the lower basis set level, increasing in absolute difference by ca. 1 km mol<sup>-1</sup> lending support for the relative performance comparisons made here. For the methods using the LYP correlation functional, the average absolute difference decreased a small amount with the large basis set. MP2 shows an increase in absolute difference from QCISD with this basis set, compared to results with the 6-31G\* basis set. The best performers are the hybrid DFT methods, all having absolute differences ca. 5 km mol<sup>-1</sup> from QCISD. The local

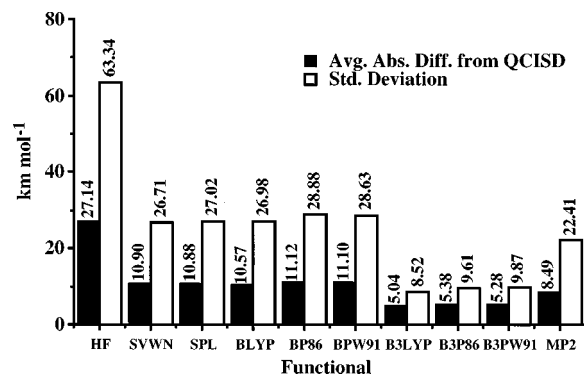


FIG. 5. Average absolute difference from QCISD and standard deviation with the 6-311+G(3df,3pd) basis set.

DFT methods perform very well with average absolute differences a little better than the gradient-corrected methods and lower standard deviations. MP2 performs very similarly to the local functionals having an average absolute difference of 8 km mol<sup>-1</sup> and a standard deviation of 22 km mol<sup>-1</sup>. Table I shows the difference between the QCISD values and those obtained for the other methods at the large basis set level. Hartree–Fock again consistently overestimates infrared intensities with the average difference from QCISD and the average absolute difference from QCISD being only 1 km mol<sup>-1</sup> apart. Also, comparing the average differences and the absolute average differences suggests that the DFT and MP2 methods predict intensities distributed around the QCISD results, with the hybrid methods predicting intensities in small excess of QCISD more often. Results with the large basis set suggests a grouping of the methods based on relative performance that is in agreement with the 6-31G\* basis set results. At this level the normal mode for which the methods give the worst agreement with QCISD is again the asymmetric stretch of CO<sub>2</sub> (probably due to the neglect of the triples contribution in the QCISD calculation).<sup>41</sup> The methods from best to worst with respect to agreement with QCISD are hybrid DFT > MP2 > local DFT ≈ gradient-corrected DFT > HF. Compared to experiment, the average absolute difference for the hybrid DFT methods, ca. 14 km mol<sup>-1</sup>, is very similar to the QCISD method, ca. 13 km mol<sup>-1</sup>.

#### E. Scaling of Hartree–Fock and DFT infrared intensities

The consistent overestimation of infrared intensities by Hartree–Fock makes it a target for homogeneous scaling in order to bring the computed intensities into better agreement with higher level theory such as QCISD. Although our set of test molecules is relatively small, we can obtain an empirical scale factor which may be refined by future studies. Least squares analysis of the Hartree–Fock and QCISD/6-311+G(3df,3pd) intensities suggests a scale factor of 0.68. Figure 6 shows a scatter plot of raw HF intensities and scaled HF intensities versus QCISD values. The scaling factor performs well in bringing the distribution of HF intensities into greater coincidence with the QCISD values. After scaling the average absolute difference from QCISD, decreases by ca.

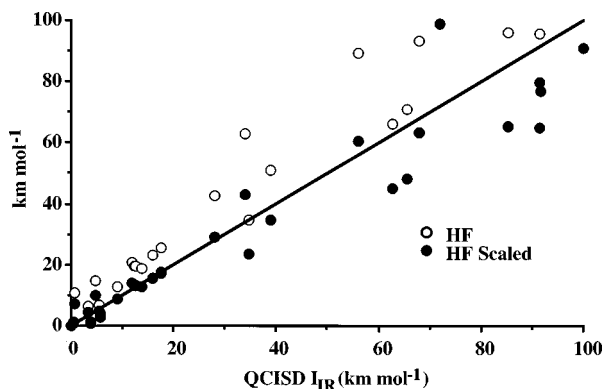


FIG. 6. Comparison of unscaled Hartree-Fock infrared intensities and the Hartree-Fock intensities scaled by a factor of 0.68 with the QCISD values using the 6-311+G(3df,3pd) basis set.

66% to 9.35 km mol<sup>-1</sup> with a standard deviation of 14.85 km mol<sup>-1</sup>. Similar treatment of the Hartree-Fock and QCISD results for the 6-31G\* basis set yields a least squares scale factor of 0.63. Use of this scaling factor decreases the average absolute difference by ca. 54% to 11.05 km mol<sup>-1</sup> with a standard deviation of 17.44 km mol<sup>-1</sup> for scaled Hartree-Fock results at that level.

Figures 7 and 8 show scatter plots for the local and gradient-corrected DFT and the hybrid DFT methods, respectively. All computed intensities coincide very closely for the functionals in a given category (local, gradient corrected, or hybrid). The distribution of DFT calculated intensities around the QCISD values makes the strategy of homogeneous scaling less applicable to the DFT intensities. The DFT computed intensities are sufficiently close to QCISD intensities that scaling is not necessary. The hybrid DFT scatter plot shows the decrease in average absolute difference from QCISD very clearly compared to the local and gradient-corrected plot with a visibly tighter distribution about the QCISD line. On the hybrid DFT plot, MP2 values are also included for comparison. MP2 clearly does a poorer job in predicting reliable infrared intensities in agreement with QCISD than the DFT methods, having several extreme outliers.

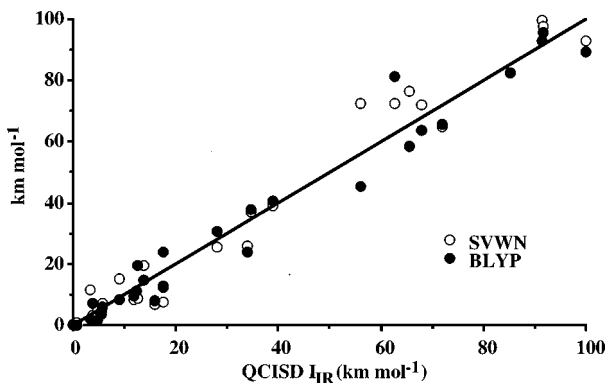


FIG. 7. Comparison of local DFT (SVWN) and gradient corrected DFT (BLYP) infrared intensities with QCISD values using the 6-311+G(3df,3pd) basis set.

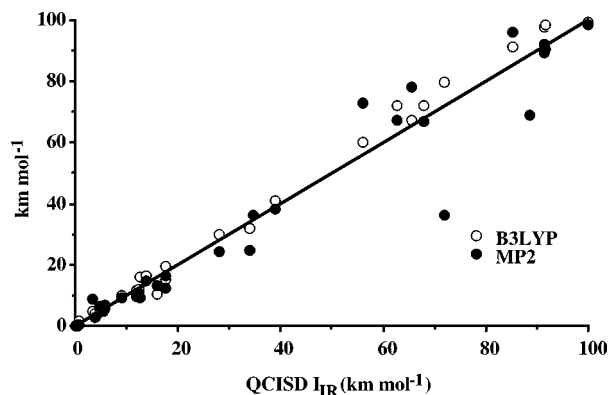


FIG. 8. Comparison of hybrid DFT (B3LYP) and MP2 infrared intensities with QCISD values using the 6-311+G(3df,3pd) basis set.

In summary, the order of agreement in predicted infrared intensities between QCISD and the methods surveyed here is: hybrid DFT > MP2 > local DFT  $\approx$  gradient-corrected DFT > HF. These results are significant in light of the substantial savings in CPU time DFT affords the researcher. For comparison, the relative CPU times for the frequency calculation from an optimized geometry for CH<sub>2</sub>F<sub>2</sub> using the 6-311+G(3df,3pd) basis set with the methods discussed in this report are given in Table I. All the DFT methods are very cost effective in computing infrared intensities. The local DFT methods take less time than Hartree-Fock and give remarkably better intensities. The gradient-corrected DFT methods take almost twice as long as the local DFT methods; however, their performance is the same or a little worse than the local methods in comparison to QCISD. The hybrid DFT functionals are the best performers of the methods considered here; although they are a little more expensive than the gradient-corrected methods, they yield far superior results. MP2 requires greater than four times the CPU time the local DFT methods require with no substantial gain in performance.

#### IV. CONCLUSIONS

Density functional theory offers excellent quantitative performance in the prediction of infrared intensities in good agreement with high-level *ab initio* correlated methods (QCISD) at a small fraction of the cost. Intensities computed at the Hartree-Fock level are in much worse agreement with QCISD than DFT; however, there is evidence that an empirical scale factor can improve the agreement of the intensities, making them comparable to MP2. Based on this work we propose a scale factor of 0.63 for the 6-31G\* basis set which may be refined in future studies. The local and gradient-corrected DFT methods predict intensities in good agreement with QCISD and each other, although the gradient-corrected methods take twice as much time as the local methods. The hybrid DFT methods are the best performers in computing intensities in closest agreement with QCISD, whereas MP2 shows performance intermediate between local DFT gradient-corrected DFT and hybrid DFT, but with several severe outliers and poor scaling with system size. For cost-effective prediction of vibrational properties (harmonic fre-

quencies and infrared intensities), local and hybrid DFT methods offer both good performance/cost and scaling with size.

## ACKNOWLEDGMENTS

We gratefully acknowledge support from the National Science Foundation (Grant No. CHE9400678) and a grant for computing resources from the Pittsburgh Supercomputing Center.

- <sup>1</sup>P. Pulay in *Modern Electronic Structure Theory*, edited by D. Yarkony (World Scientific, Singapore, 1995), p. 1191.
- <sup>2</sup>C. W. Bauschlicher and S. R. Langhoff, *Chem. Rev.* **91**, 701 (1991).
- <sup>3</sup>A. Kam, R. Aroca, J. Duff, and C. P. Tripp, *Chem. Mater.* **10**, 172 (1998).
- <sup>4</sup>B. A. Hess, L. J. Schaad, and P. L. Polavarapu, *J. Am. Chem. Soc.* **106**, 4348 (1984).
- <sup>5</sup>W. Kohn, A. D. Becke, and R. G. Parr, *J. Phys. Chem.* **100**, 12974 (1996).
- <sup>6</sup>R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
- <sup>7</sup>B. G. Johnson, P. M. W. Gill, and J. A. Pople, *J. Chem. Phys.* **98**, 5612 (1993).
- <sup>8</sup>M. W. Wong, *Chem. Phys. Lett.* **256**, 391 (1996).
- <sup>9</sup>A. P. Scott and L. Radom, *J. Phys. Chem.* **100**, 16502 (1996).
- <sup>10</sup>F. De Prof, J. M. L. Martin, and P. Geerlings, *Chem. Phys. Lett.* **250**, 393 (1996).
- <sup>11</sup>D. Porezag and M. R. Pederson, *Phys. Rev. B* **54**, 7830 (1996).
- <sup>12</sup>C. H. Choi and M. Kertesz, *J. Phys. Chem.* **100**, 16530 (1996).
- <sup>13</sup>K. D. Dobbs and D. A. Dixon, *J. Phys. Chem.* **98**, 4498 (1994).
- <sup>14</sup>D. Papousek, Z. Papouškova, and D. P. Chong, *J. Phys. Chem.* **99**, 15387 (1995).
- <sup>15</sup>C. E. Dykstra, *Ab Initio Calculation of the Structures and Properties of Molecules* (Elsevier Science, Netherlands, 1988).
- <sup>16</sup>Y. Yamaguchi, M. Frisch, J. Gaw, H. F. Schaefer, and J. S. Binkley, *J. Chem. Phys.* **84**, 2262 (1986).
- <sup>17</sup>K. Ohno and T. Ishida, *Int. J. Quantum Chem.* **29**, 677 (1986).
- <sup>18</sup>J. F. Stanton, W. N. Lipcomb, D. H. Magers, and R. J. Bartlett, *J. Chem. Phys.* **90**, 3241 (1989).
- <sup>19</sup>M. D. Miller, F. Jensen, O. L. Chapman, and K. N. Houk, *J. Phys. Chem.* **93**, 4495 (1989).
- <sup>20</sup>J. R. Thomas, B. J. DeLeeuw, G. Vacek, T. D. Crawford, Y. Yamaguchi, and H. F. Schaefer, *J. Chem. Phys.* **99**, 403 (1993).
- <sup>21</sup>See AIP Document No. E-PAPS: E-JCPA6-109-307848 for 5 tables of numerical results for the 6-31G\* and 6-311+G(3df,3pd) basis sets, S-VWN and B-LYP basis set dependence data and comparisons with experiment. E-PAPS document files may be retrieved free of charge from our FTP server (<http://www.aip.org/epaps/epaps.html>) or from <ftp.aip.org> in the directory /epaps/. For further information, e-mail: [paps@aip.org](mailto:paps@aip.org) or fax: 516-576-2223.
- <sup>22</sup>GAUSSIAN 94, Revision C.3, M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople (Gaussian, Inc., Pittsburgh, PA, 1995).
- <sup>23</sup>C. Moller and M. S. Plesset, *Phys. Rev.* **46**, 618 (1934).
- <sup>24</sup>J. A. Pople, M. Head-Gordon, and K. Raghavachari, *J. Chem. Phys.* **76**, 5968 (1987).
- <sup>25</sup>J. Overend, in *Infrared Spectroscopy and Molecular Structure*, edited by M. Davies (Elsevier Scientific, Amsterdam, 1963), p. 345.
- <sup>26</sup>J. C. Slater, *Quantum Theory of Molecules and Solids*, Vol. 4. (McGraw-Hill, New York, 1974).
- <sup>27</sup>S. J. Vosko, L. Wilk, and M. Nusair, *Can. J. Phys.* **58**, 1200 (1980).
- <sup>28</sup>J. P. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).
- <sup>29</sup>A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).
- <sup>30</sup>C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
- <sup>31</sup>J. P. Perdew, *Phys. Rev. B* **33**, 8822 (1986).
- <sup>32</sup>J. P. Perdew, in *Electronic Structure of Solids*, edited by P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991).
- <sup>33</sup>A. D. Becke, *Can. J. Phys.* **98**, 5648 (1993).
- <sup>34</sup>P. C. Hariharan and J. A. Pople, *Theor. Chim. Acta* **28**, 213 (1973).
- <sup>35</sup>J. A. Pople, H. B. Schlegel, R. Krishnan, D. J. Defrees, J. S. Binkley, M. J. Frisch, R. A. Whiteside, R. F. Hout, and W. J. Hehre, *Int. J. Quantum Chem., Quantum Chem. Symp.* **15**, 269 (1981).
- <sup>36</sup>R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, *J. Chem. Phys.* **72**, 650 (1980).
- <sup>37</sup>A. D. McLean and G. S. Chandler, *J. Chem. Phys.* **72**, 5639 (1980).
- <sup>38</sup>T. Clark, J. Chandrasekhar, G. W. Spitznagel, and P. v. R. Schleyer, *J. Comput. Chem.* **4**, 294 (1983).
- <sup>39</sup>M. J. Frisch, J. A. Pople, and J. S. Binkley, *J. Chem. Phys.* **80**, 3265 (1984).
- <sup>40</sup>J. Overend, in *Vibrational Intensities in Infrared and Raman Spectroscopy*, edited by w. B. Person and G. Zerbi (Elsevier Scientific, Amsterdam, 1982), p. 14.
- <sup>41</sup>The asymmetric stretch of CO<sub>2</sub> was the only mode found to have a large triples correction [115 km mol<sup>-1</sup> for CCSD(T)-CCSD with the TZ(2df,2pd) basis set (Ref. 20)].
- <sup>42</sup>K. K. Baek, J. D. Watts, and R. J. Bartlett, *J. Chem. Phys.* **107**, 3853 (1997).
- <sup>43</sup>K. Raghavachari and J. A. Pople, *Int. J. Quantum Chem.* **20**, 167 (1981).
- <sup>44</sup>H. J. Werner and W. Meyer, *Mol. Phys.* **31**, 855 (1976).
- <sup>45</sup>G. H. F. Dierksen and A. J. Sadlej, *Theor. Chim. Acta* **63**, 69 (1983).
- <sup>46</sup>Note, however, that the calculations yield harmonic frequencies, whereas the observed values are anharmonic fundamentals which are typically 5% lower than the experimental harmonic frequencies.
- <sup>47</sup>Omitted due to the high computational cost of obtaining QCISD/6-311+G(3df,3pd) results.
- <sup>48</sup>W. J. Hehre, L. Random, P. v. R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory* (Wiley, New York 1986).