

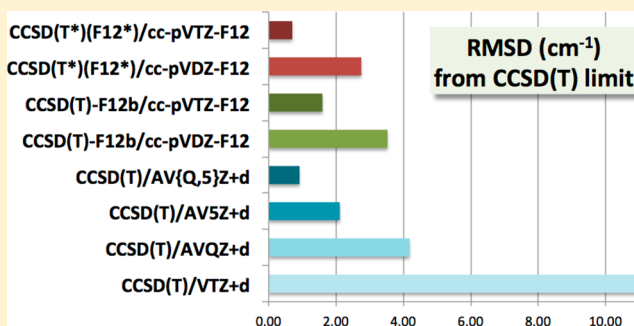
# Assessment of CCSD(T)-F12 Approximations and Basis Sets for Harmonic Vibrational Frequencies

Jan M. L. Martin\* and Manoj K. Kesharwani

Department of Organic Chemistry, Weizmann Institute of Science, 76100 Rehovot, Israel

## Supporting Information

**ABSTRACT:** We consider basis set convergence and the effect of various approximations to CCSD(T)-F12 for a representative sample of harmonic frequencies (the HFREQ2014 set). CCSD(T\*)(F12\*)/cc-pVDZ-F12 offers a particularly favorable compromise between accuracy and computational cost: its RMSD  $<3\text{ cm}^{-1}$  from the valence CCSD(T) limit is actually less than the remaining discrepancy with the experimental value at the valence CCSD(T) limit (about  $5\text{ cm}^{-1}$  RMSD). CCSD(T)-F12a and CCSD(T)-F12b appear to benefit from error compensation between CCSD and (T).



## INTRODUCTION

The importance of vibrational spectra in chemistry, biochemistry, and chemical biophysics hardly requires elaboration. In recent years, CCSD(T)<sup>1</sup> anharmonic force fields have proven a powerful predictive spectroscopic tool.<sup>2,3</sup> Basis set convergence of CCSD(T) vibrational frequencies is fairly slow,<sup>4–7</sup> and hence, such force fields are computationally very expensive beyond very small molecules, although this is somewhat mitigated by the recent availability of CCSD(T) analytical second derivatives<sup>8</sup> in a parallel implementation.<sup>9</sup> One recent approach has been to combine CCSD(T) harmonic frequencies with a DFT anharmonic force field.<sup>10,11</sup> Another tack has been to carry out pointwise basis set extrapolations, as seen, e.g., in the work of Huang et al.<sup>12–14</sup> and in ref 15.

Yet another tactic would be to speed up basis set convergence of the CCSD(T) part by means of explicitly correlated (F12) approaches:<sup>16,17</sup> for many applications, these offer quintuple-zeta or better results with triple- $\zeta$  sized basis sets, or more generally a gain by 2–3 “zeta steps.” Huang et al.<sup>14</sup> compared extrapolation and F12 techniques for quartic force fields of H<sub>2</sub>O, N<sub>2</sub>H<sup>+</sup>, NH<sub>2</sub><sup>+</sup>, and C<sub>2</sub>H<sub>2</sub> and concluded that F12 is competitive and in some cases necessary. Rauhut et al.<sup>18</sup> carried out comparative CCSD(T) and CCSD(T)-F12{a,b} basis set convergence studies for harmonic frequencies of H<sub>2</sub>O, HCN, CO<sub>2</sub>, C<sub>2</sub>H<sub>2</sub>, and H<sub>2</sub>CO, as well as compared conventional and explicitly correlated anharmonic frequencies with the aug-cc-pVTZ basis set<sup>19</sup> for the above molecules plus H<sub>2</sub>O<sub>2</sub>, CH<sub>2</sub>NH, and C<sub>2</sub>H<sub>2</sub>O: they conclude that the “two zeta gain rule” applies here as well and unreservedly advocate for explicitly correlated methods.

Since, however, the full CCSD-F12 approach is both algebraically and computationally very involved,<sup>20</sup> and no practical implementation of CCSD(T)-F12 without any approximations is available, one now needs to pay heed not

only to the choice of the basis sets but also to the choice of the specific approximation made.

Werner and co-workers proposed<sup>21,22</sup> the CCSD-F12a and CCSD-F12b approximations, which they implemented in the popular MOLPRO package;<sup>23</sup> more recently, Hättig et al. proposed<sup>24</sup> the CCSD(F12\*) approximation (a.k.a., CCSD-F12c) which offers CCSD(F12)<sup>25,26</sup> quality at a cost comparable to that of CCSD-F12b. It should be noted here that the CCSD-F12b approximation includes a subset of terms from CCSD(F12\*) and CCSD-F12a likewise from CCSD-F12b; thus, CCSD-F12a, CCSD-F12b, and CCSD(F12\*) can be considered a hierarchy of increasingly rigorous approximations to CCSD-F12.

In the present work, we will consider basis set convergence of harmonic frequencies using different approximations to CCSD(T)-F12 for a medium-sized benchmark set of small molecules (HFREQ2014) selected and expanded from our earlier HFREQ27 set.<sup>27</sup> We will show that CCSD(T\*)(F12\*)/cc-pVTZ-F12 and especially CCSD(T\*)(F12\*)/cc-pVDZ-F12 offer particularly cost-effective means of calculating reliable vibrational spectra.

## COMPUTATIONAL DETAILS

Most calculations were carried out using MOLPRO 2012.1<sup>23</sup> running on the Faculty of Chemistry HPC cluster at the Weizmann Institute of Science. Some additional calculations were performed using Turbomole 6.5<sup>28</sup> running on the same platform.

In most calculations, we employed the cc-pVnZ-F12 correlation consistent basis sets ( $n = \text{D, T, Q}$ ) optimized for F12 calculations<sup>29</sup> and the auxiliary basis sets<sup>30</sup> and CABS

Received: February 27, 2014

Published: April 8, 2014



Table 1. Difference Statistics for CCSD(T) Harmonic Frequencies ( $\text{cm}^{-1}$ ) Using Various Basis Sets and F12 Approximations<sup>a</sup>

|             | VQZ-F12<br>(F12*) | VQZ-F12<br>F12b | AV5Z<br>(F12*) <sup>a</sup> | VTZ-F12<br>(F12*) | VTZ-F12<br>F12b | VTZ-F12<br>F12a | VDZ-F12<br>(F12*) | VDZ-F12<br>F12b | VDZ-F12<br>F12a |
|-------------|-------------------|-----------------|-----------------------------|-------------------|-----------------|-----------------|-------------------|-----------------|-----------------|
| (T*) scaled |                   |                 |                             |                   |                 |                 |                   |                 |                 |
| MSDiff      | REF               | -1.08           | -0.50                       | -0.33             | -2.71           | -3.29           | -0.38             | -4.11           | -5.91           |
| RMSDiff     | REF               | 1.33            | 0.91                        | 0.69              | 3.33            | 4.04            | 2.72              | 6.36            | 8.07            |
| Max+Dev     | REF               | 0.14            | 0.36                        | 1.48              | 0.15            | -0.27           | 9.31              | 7.99            | 5.27            |
| Max-Dev     | REF               | -4.75           | -3.75                       | -2.27             | -12.85          | -14.30          | -9.39             | -21.82          | -26.40          |
| T unscaled  |                   |                 |                             |                   |                 |                 |                   |                 |                 |
| MSDiff      | 0.79              | -0.34           | 0.22 <sup>b</sup>           | 1.17              | -1.12           | -1.71           | 3.34              | -0.18           | -1.97           |
| RMSDiff     | 0.99              | 0.57            | 0.74                        | 1.61              | 1.59            | 2.21            | 4.43              | 3.52            | 4.24            |
| Max+Dev     | 3.87              | 0.53            | 2.37                        | 5.52              | 1.39            | 0.40            | 14.53             | 13.26           | 10.56           |
| Max-Dev     | -1.28             | -1.74           | -0.78                       | -1.44             | -5.31           | -6.74           | -6.39             | -12.22          | -15.05          |

<sup>a</sup>The total number of nonunique frequencies in the HFREQ2014 set is 124. Retaining degenerate duplicates is equivalent to having weights equal to the degeneracies in the statistics. Geminal exponents used for the F12 calculations are as recommended in ref 59: 0.9 for cc-pVDZ-F12, 1.0 for cc-pVTZ-F12, 1.1 for cc-pVQZ-F12, and 1.4 for aug-cc-pV5Z. VnZ-F12 is the abbreviation for cc-pVnZ-F12. Color shading is on a spectrum from green for the lowest to red for the highest RMSD, with yellow representing the median value. <sup>a</sup>, subset of 33 frequencies. <sup>b</sup>, subset of 24 frequencies.

Table 2. Difference Statistics, Relative to CCSD(T\*)-F12(\*)/cc-pVQZ-F12, for Conventional CCSD(T) Harmonic Frequencies ( $\text{cm}^{-1}$ ) Using Various Basis Sets<sup>a</sup>

|         | AV{Q,S}Z <sup>b</sup> | cc-pV(T+d)Z | cc-pV(Q+d)Z | aug-cc-pV(Q+d)Z | aug-cc-pV(5+d)Z <sup>b</sup> |
|---------|-----------------------|-------------|-------------|-----------------|------------------------------|
| MSDiff  | -0.33                 | -2.80       | -0.88       | -3.09           | -1.46                        |
| RMSDiff | 0.86                  | 11.25       | 5.45        | 4.17            | 2.09                         |
| Max+Dev | 2.21                  | 52.36       | 27.17       | 3.68            | 0.88                         |
| Max-Dev | -2.36                 | -42.05      | -24.45      | -12.69          | -6.70                        |

<sup>a</sup>AV{Q,S}Z refers to pointwise energy extrapolation from CCSD(T)/aug-cc-pV(Q+d)Z and CCSD(T)/aug-cc-pV(5+d)Z calculations using the Martin-Karton formula<sup>66</sup> for the SCF component and the Schwenke formula<sup>67</sup> for the correlation energy. The statistics for AV{Q,S}Z and aug-cc-pV(5+d)Z are exclusive of CH<sub>3</sub>OH. <sup>b</sup>Subset of 112 frequencies.

(complementary auxiliary basis sets)<sup>31</sup> developed for use with them. Slater-type geminals are used throughout: the recommended values<sup>29</sup> of the geminal exponent were used for them, namely, 0.9 for cc-pVDZ-F12, 1.0 for cc-pVTZ-F12, 1.1 for cc-pVQZ-F12, and 1.4 for aug-cc-pV5Z. The SCF component was improved through the “CABS correction” (refs 21 and 32 and references therein).

In conventional calculations, we employed the correlation consistent basis sets cc-pV(T+d)Z, cc-pV(Q+d)Z, aug-cc-pV(Q+d)Z, and aug-cc-pV(5+d)Z.<sup>33</sup> The “+d” notation refers to the addition of a high-d exponent to second row atoms, which is essential<sup>34–36</sup> for accurate results in molecules (such as SO<sub>2</sub>) where the second-row atom is in a high oxidation state and can receive back-donation into the 3d orbital.<sup>37</sup> (Such basis functions are already included<sup>29</sup> in cc-pVnZ-F12.)

Experimental harmonic frequencies for the diatomics were taken from Huber and Herzberg<sup>38</sup> (for ClF, a significantly revised remeasurement<sup>39,40</sup> was substituted), while (quasi-)experimental data for polyatomics were compiled from various studies. Some of these are from experimental studies (C<sub>2</sub>H<sub>2</sub>,<sup>41</sup> ClCN,<sup>42</sup> CO<sub>2</sub>,<sup>43</sup> CS<sub>2</sub>,<sup>44</sup> H<sub>2</sub>O,<sup>45</sup> H<sub>2</sub>S,<sup>46</sup> HCN,<sup>47</sup> N<sub>2</sub>O,<sup>48</sup> OCS,<sup>49</sup> and SO<sub>2</sub><sup>50</sup>) and others from computed force fields that were adjusted in variational or perturbational calculations for best agreement with experiment (NH<sub>3</sub>,<sup>51</sup> C<sub>2</sub>H<sub>4</sub>,<sup>52</sup> CCl<sub>2</sub>,<sup>53</sup> CF<sub>2</sub>,<sup>54</sup> CH<sub>4</sub>,<sup>55</sup> H<sub>2</sub>CO,<sup>56</sup> H<sub>2</sub>CS,<sup>57</sup> PH<sub>3</sub>,<sup>58</sup> and CH<sub>3</sub>OH<sup>59</sup>), while in two cases (HOCl<sup>60</sup> and BH<sub>3</sub><sup>61</sup>), outright *ab initio* calculations of near-spectroscopic quality were used.

Internal options in the CCSD-F12 code were set to ortho\_cabs=1,thrcabs=1d-9,thrcabs\_rel=1d-9. Atomic masses for the most abundant isotopes were used throughout

(overriding the MOLPRO default). Total energies were converged to 1 picohartree where feasible (cc-pVDZ-F12, cc-pVTZ-F12, and most cc-pVQZ-F12), but for the aug-cc-pV5Z and some second-row cc-pVQZ-F12 calculations, convergence criteria had to be relaxed to 10 or even 100 picohartree. In cases where 1 or 10 picohartree was feasible, a step size of 0.005 au was used in the numerical differentiation in order to reduce discretization error; the default stepsize of 0.01 au was employed where only 100 picohartree convergence was possible. For a few molecules and the conventional CCSD(T) method, we considered various step sizes for the double-numerical frequency calculations and compared them to the fully analytical values obtained using CFOUR.<sup>62</sup> These experiments suggest that our computed vibrational frequencies should have a numerical precision, on average, of about 0.25  $\text{cm}^{-1}$  or better.

We considered the CCSD-F12a, CCSD-F12b,<sup>21,22</sup> and CCSD(F12\*)<sup>24</sup> (a.k.a., CCSD-F12c) approximations. F12 approaches as presently practiced do not directly affect the connected quasiperturbative triples, so the basis set convergence behavior of the (T) contribution is effectively that of an orbital-based calculation. In an attempt to speed up their basis set convergence, Marchetti and Werner<sup>63</sup> proposed scaling the (T) contribution by the MP2-F12/MP2 correlation energy ratio. Marchetti-Werner scaling will be indicated by the notation (T\*) instead of (T) in the present article. (We note in passing that a genuine explicitly correlated (T) correction<sup>64</sup> has recently been proposed.)

For a number of the diatomic molecules, we also considered the full CCSD(F12) method by means of the implementation

Table 3. Basis Set Convergence of CCSD Harmonic Frequencies ( $\text{cm}^{-1}$ ) Using Various Basis Sets and F12 Approximations<sup>a</sup>

|         | VQZ-<br>F12 | VQZ-<br>F12 | AV5Z   | AV5Z  | VTZ-<br>F12 | VTZ-<br>F12 | VTZ-<br>F12 | VDZ-<br>F12 | VDZ-<br>F12 | VDZ-<br>F12 |
|---------|-------------|-------------|--------|-------|-------------|-------------|-------------|-------------|-------------|-------------|
|         | (F12*)      | F12b        | (F12*) | F12b  | (F12*)      | F12b        | F12a        | (F12*)      | F12b        | F12a        |
| MSDiff  | REF         | -0.59       | -0.22  | -0.67 | -0.26       | -2.07       | -2.61       | -0.55       | -3.09       | -4.88       |
| RMSDiff | REF         | 1.08        | 0.55   | 0.76  | 0.64        | 2.60        | 3.25        | 2.24        | 4.97        | 6.63        |
| Max+Dev | REF         | 3.40        | 0.50   | 0.25  | 1.41        | 1.58        | 0.97        | 5.77        | 5.12        | 2.27        |
| Max-Dev | REF         | -3.04       | -2.42  | -1.74 | -2.00       | -8.21       | -9.42       | -5.82       | -15.38      | -19.92      |

<sup>a</sup>Results with the AV5Z basis set are for a 32-frequency subset.

Table 4. RMS Differences (Angstrom) in Calculated Bond Lengths for the HFREQ2014 Molecules

| explicitly correlated             |             |                 |        |                 |         |                       |         |                            |                     |         |
|-----------------------------------|-------------|-----------------|--------|-----------------|---------|-----------------------|---------|----------------------------|---------------------|---------|
|                                   | (F12*)      | F12b            | (F12*) | F12b            | (F12*)  | F12b                  | F12a    | (F12*)                     | F12b                | F12a    |
|                                   | VQZ-F12     | VQZ-F12         | AV5Z   | AV5Z            | VTZ-F12 | VTZ-F12               | VTZ-F12 | VDZ-F12                    | VDZ-F12             | VDZ-F12 |
| CCSD(T*)                          | REF         | 0.0004          | 0.0001 |                 | 0.0004  | 0.0011                | 0.0014  | 0.0013                     | 0.0021              | 0.0027  |
| CCSD(T) <sup>a</sup>              | 0.0003      | 0.0001          | 0.0002 |                 | 0.0003  | 0.0006                | 0.0008  | 0.0005                     | 0.0010              | 0.0015  |
| CCSD only                         | REF         | 0.0003          | 0.0001 | 0.0001          | 0.0003  | 0.0009                | 0.0011  | 0.0013                     | 0.0018              | 0.0023  |
| conventional CCSD(T) <sup>a</sup> |             |                 |        |                 |         |                       |         | CCSD(T*) (F12*) incl. core |                     |         |
| cc-pV(T+d)Z                       | cc-pV(Q+d)Z | aug-cc-pV(Q+d)Z |        | aug-cc-pV(5+d)Z |         | AV(Q,5)Z <sup>b</sup> |         | CVTZ-F12                   | CVQZ-F12            |         |
| 0.0056                            | 0.0021      | 0.0023          |        | 0.0010          |         | 0.0002                |         | 0.0025 <sup>c</sup>        | 0.0026 <sup>c</sup> |         |

<sup>a</sup>Relative to valence CCSD(T\*)-F12(\*)/cc-pVQZ-F12. <sup>b</sup>Pointwise energy extrapolation. <sup>c</sup>This number reflects the effect of core–valence correlation.

in Turbomole 6.5.<sup>28</sup> Geometries and harmonic frequencies were obtained by calculating energies at 11 points spaced at 0.01 Å around the CCSD-F12b equilibrium distance, fitting the highest-order polynomial in  $(r - r_e)/r_e$  that it is still statistically significant (typically order 5–7), and performing a Dunham analysis.<sup>65</sup> For the species considered, CCSD(F12) and CCSD(F12\*) spectroscopic parameters were found to agree to within the uncertainties of the fits. It was thus concluded that, while CCSD(F12) and CCSD(F12\*) show discrepancies in total energies, the two models are functionally equivalent for the purpose of calculating vibrational frequencies.

## RESULTS AND DISCUSSION

Full harmonic frequency data for the molecules considered are given in the Supporting Information. As we have no “exact” CCSD(T) reference values per se, we have arbitrarily chosen CCSD(T\*) (F12\*)/cc-pVQZ-F12 as our vantage point and calculated difference statistics from that for the various levels considered. These are given in Table 1, while for perspective, we are including statistics for orbital-based CCSD(T) in Table 2.

In addition, in order to separate out the (T) contribution, we are presenting the same statistics for various CCSD-F12 approximations in Table 3.

Basis set convergence of CCSD(T)-F12a and CCSD(T)-F12b for a small subset of HFREQ2014 was recently considered by Rauhut et al.<sup>18</sup> We verified, by recalculation, that the small discrepancies between their and our data result from their use of  $\beta = 1.0$  and of isotope-averaged masses versus ours of most-abundant-isotope masses and of basis set-optimized  $\beta$  values.

First of all, we note that CCSD(F12\*)/cc-pVQZ-F12 and CCSD(F12\*)/aug-cc-pV5Z agree to about 0.6  $\text{cm}^{-1}$  RMSD, basically to within the numerical precision of the calculations. Second, CCSD(F12\*)/cc-pVQZ-F12 and CCSD(F12\*)/cc-pVTZ-F12 agree to about 0.6  $\text{cm}^{-1}$  over the entire set of

molecules, suggesting that CCSD(F12\*)/cc-pVQZ-F12 is effectively converged with the basis set and can be used as a reference. Third, CCSD-F12b agrees to better than 1.1  $\text{cm}^{-1}$  with those values for both the cc-pVQZ-F12 and aug-cc-pV5Z basis sets. Fourth, CCSD-F12b/cc-pVTZ-F12 and CCSD-F12a/cc-pVTZ-F12 deviate by about 2.6 and 3.3  $\text{cm}^{-1}$  RMS, which actually exceeds the 2.2  $\text{cm}^{-1}$  obtained at the CCSD(F12\*) level with the small cc-pVDZ-F12 basis set. With said basis set, CCSD-F12b and CCSD-F12a have RMSDs of 5 and nearly 7  $\text{cm}^{-1}$ . Inspection of the mean signed deviations reveals that CCSD-F12b and CCSD-F12a both systematically underestimate frequencies with smaller basis sets, while CCSD(F12\*) does not appreciably do so even with cc-pVDZ-F12.

It would appear to be obvious that CCSD(F12\*) basis set convergence is considerably faster than that of CCSD-F12b. In fact, the difference would appear to exceed the advantage of CCSD-F12b over CCSD-F12a. The same trends are seen in the computed bond distances (Table 4).

Let us turn now toward CCSD(T). Here, we have an additional factor to consider: whether or not to include Marchetti-Werner scaling for the (T) contribution. As expected, this appears to matter little for the cc-pVQZ-F12 and especially aug-cc-pV5Z basis sets, but the gap widens for cc-pVTZ-F12 and especially cc-pVDZ-F12. For the entire set of molecules, CCSD(T\*) (F12\*)/cc-pVTZ-F12 agrees with CCSD(T\*) (F12\*)/cc-pVQZ-F12 to just 0.7  $\text{cm}^{-1}$  and 0.0004 Å RMSD. In fact, conventional CCSD(T) calculations with pointwise aug-cc-pV(*n*+d)Z basis set extrapolation (*n* = Q,5) agree to better than 1  $\text{cm}^{-1}$  and 0.0002 Å RMSD.

As (T) systematically lowers calculated harmonic frequencies (by an average of about 27  $\text{cm}^{-1}$  for the molecules considered, reaching about 72, 80, and 95  $\text{cm}^{-1}$  for CO, N<sub>2</sub>, and F<sub>2</sub>), the decision not to scale the triples causes the frequencies to shift upward on average. For CCSD(T)-F12b with the cc-pVTZ-F12 basis set, this appears to compensate for the downward bias in



CCSD-F12b/cc-pVTZ-F12, leading to the low RMSD  $\approx 1.6$   $\text{cm}^{-1}$  compared to  $3.3$   $\text{cm}^{-1}$  for CCSD(T\*)-F12b/cc-pVTZ-F12. Note that, in the absence of Marchetti-Werner scaling, the average errors of CCSD(T)(F12\*)/cc-pVTZ-F12 and CCSD(T)-F12b/cc-pVTZ-F12 have opposite signs. The surprisingly low RMSD for CCSD(T)-F12b/cc-pVTZ-F12 is likewise due to an error compensation.

Perhaps the single most useful observation in this table is the low RMSD  $\approx 2.7$   $\text{cm}^{-1}$  of the inexpensive CCSD(T\*)(F12\*)/cc-pVDZ-F12 level of theory, which is actually superior to  $4.2$   $\text{cm}^{-1}$  at the CCSD(T)/aug-cc-pV(Q+d)Z level and approaches  $2.1$   $\text{cm}^{-1}$  at the CCSD(T)/aug-cc-pV(5+d)Z level. In other words, CCSD(T\*)(F12\*) gains almost three “zeta’s” over a conventional orbital-based CCSD(T) frequency calculation.

For some perspective (Table 5), the RMSD between (quasi)experimental harmonics and CCSD(T\*)(F12\*) is itself

**Table 5. Comparison of CCSD(T\*)(F12\*) Harmonic Frequencies to (Quasi-)Experimental Harmonic Frequencies ( $\text{cm}^{-1}$ )<sup>a</sup>**

|         | cc-pVTZ-F12<br>frozen<br>core | cc-pVQZ-F12<br>frozen<br>core | cc-pCVTZ-F12<br>all electron,<br>$\beta = 1.4$ | cc-pCVQZ-F12<br>all electron,<br>$\beta = 1.5$ |
|---------|-------------------------------|-------------------------------|--|--|
| MSDiff  | 0.22                          | 0.55                          | 4.60   | 4.31   |
| RMSDiff | 4.69                          | 4.68                          | 7.01   | 7.02   |
| Max+Dev | 13.66                         | 15.03                         | 19.67  | 19.92  |
| Max-Dev | -12.86                        | -12.26                        | -11.49   | -12.16   |

<sup>a</sup>cc-pCVQZ-F12 results are for a 120-frequency subset.

$4.7$   $\text{cm}^{-1}$  near the basis set limit, which actually goes up to about  $7.0$   $\text{cm}^{-1}$  if core–valence correlation is considered using the cc-pCVTZ-F12 and cc-pCVQZ-F12 basis sets.<sup>68</sup> This deterioration is mostly due to stretching frequencies, where valence CCSD(T) benefits from an error compensation<sup>2,5,6,15</sup> between core–valence correlation, which blue-shifts frequencies, and various factors that red-shift them, most importantly post-CCSD(T) correlation contributions, followed by scalar relativistic corrections (especially the more ionic bond stretches<sup>69</sup>), and finally diagonal Born–Oppenheimer corrections (which red-shift A–H stretches by small amounts). As can be seen in, e.g., Table 2 of ref 15, detailed analysis of higher-order correlation contributions in diatomics reveals that connected quadruple excitations universally lower harmonic frequencies (by as much as  $16$  and  $20$   $\text{cm}^{-1}$  for  $\text{F}_2$  and  $\text{N}_2$ , respectively) and even connected quintuple excitations (by as much as  $3$  and  $4$   $\text{cm}^{-1}$  for  $\text{N}_2$  and  $\text{O}_2$ , respectively); higher-order triple excitations tend to red-shift in cases dominated by a single reference configuration but to blue-shift in cases with significant nondynamical correlation.

A remark is due about CPU times. While timing comparisons are approximate due to clock speed adjustment to the thermal profile (“TurboBoost”) in Intel Nehalem CPUs, elapsed times on 8 cores for the numerical Hessian of ethylene are illustrative. Conventional CCSD(T)/cc-pVnZ calculations ( $n = \text{T}, \text{Q}, \text{5}$ ) require about 8 min, 2 h, and (on 16 cores) 8 h, respectively, while the explicitly correlated CCSD(T\*)(F12\*)/cc-pVnZ-F12 ( $n = \text{D}, \text{T}, \text{Q}$ ) required approximately 8 min, 45 min, and 5 h, respectively. In other words, a CCSD(T\*)(F12\*)/cc-pVDZ-F12 calculation that yields frequencies comparable in quality to a CCSD(T)/cc-pV5Z calculation is 2 orders of magnitude less expensive computationally. CCSD(T\*)-F12b/cc-pVnZ-F12

represents just a modest additional time savings, at 5 min for  $n = \text{D}$  and 4 h for  $n = \text{Q}$ .

## CONCLUSIONS

From the results presented in our article, it would seem clear that basis set convergence of vibrational frequencies computed at the CCSD(T\*)(F12\*) level is considerably faster than that of other approximate CCSD(T)-F12 methods and in particular that CCSD(T\*)(F12\*)/cc-pVDZ-F12 represents a particularly cost-effective option, with an RMSD from the valence CCSD(T) limit that approaches that of CCSD(T)/aug-cc-pV(5+d)Z calculations. In fact, the residual basis set incompleteness error of less than  $3$   $\text{cm}^{-1}$  RMS is less than the  $\sim 5$   $\text{cm}^{-1}$  RMS difference between the valence CCSD(T) basis set limit and (quasi-)experimental values, and any attempt to further reduce this number needs to address both core–valence correlation and post-CCSD(T) correlation effects (as these tend to affect the frequencies in opposite directions). We also find that CCSD(T)-F12a and CCSD(T)-F12b benefit from error compensation between the CCSD and (T) components.

## ASSOCIATED CONTENT

### Supporting Information

(Quasi-)experimental harmonic frequencies and detailed sourcing for the HFREQ2014 molecules; and computed harmonic frequencies for the HFREQ2014 molecules at the levels of theory considered in this article. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*Fax: +972 8 9344142. E-mail: [gershom@weizmann.ac.il](mailto:gershom@weizmann.ac.il).

### Funding

This research was supported in part by the Lise Meitner–Minerva Center for Computational Quantum Chemistry and by the Helen and Martin Kimmel Center for Molecular Design.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank Drs. Andrey Yachmenev and Walter Thiel for communicating the harmonic frequencies implicit in their quasi-spectroscopic force field<sup>56</sup> for  $\text{H}_2\text{CO}$ , and Drs. Xiaogang Wang and Tucker Carrington for communicating the same for their newly fitted force field of methane. Dr. Brina Brauer is thanked for helping track down some experimental references.

## REFERENCES

- (1) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A Fifth-Order Perturbation Comparison of Electron Correlation Theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483.
- (2) Császár, A. G. Anharmonic Molecular Force Fields. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 273–289.
- (3) Martin, J. M. L.; Lee, T. J.; Taylor, P. R.; François, J.-P. The Anharmonic Force Field of Ethylene,  $\text{C}_2\text{H}_4$ , by Means of Accurate ab Initio Calculations. *J. Chem. Phys.* **1995**, *103*, 2589.
- (4) Martin, J. M. L. On the Performance of Correlation Consistent Basis Sets for the Calculation of Total Atomization Energies, Geometries, and Harmonic Frequencies. *J. Chem. Phys.* **1994**, *100*, 8186.
- (5) Ruden, T. A.; Helgaker, T.; Jørgensen, P.; Olsen, J. Coupled-Cluster Connected Quadruples and Quintuples Corrections to the

Harmonic Vibrational Frequencies and Equilibrium Bond Distances of HF, N<sub>2</sub>, F<sub>2</sub>, and CO. *J. Chem. Phys.* **2004**, *121*, 5874–5884.

(6) Martin, J. M. L. Spectroscopic Quality ab Initio Potential Curves for CH, NH, OH and HF. A Convergence Study. *Chem. Phys. Lett.* **1998**, *292*, 411–420.

(7) Tew, D. P.; Klopper, W.; Heckert, M.; Gauss, J. Basis Set Limit CCSD(T) Harmonic Vibrational Frequencies. *J. Phys. Chem. A* **2007**, *111*, 11242–11248.

(8) Gauss, J.; Stanton, J. F. Analytic CCSD(T) Second Derivatives. *Chem. Phys. Lett.* **1997**, *276*, 70–77.

(9) Harding, M. E.; Metzroth, T.; Gauss, J.; Auer, A. A. Parallel Calculation of CCSD and CCSD(T) Analytic First and Second Derivatives. *J. Chem. Theory Comput.* **2008**, *4*, 64–74.

(10) Boese, A. D.; Klopper, W.; Martin, J. M. L. Anharmonic Force Fields and Thermodynamic Functions Using Density Functional Theory. *Mol. Phys.* **2005**, *103*, 863–876.

(11) Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini, C. Accurate Structure, Thermodynamic and Spectroscopic Parameters from CC and CC/DFT Schemes: The Challenge of the Conformational Equilibrium in Glycine. *Phys. Chem. Chem. Phys.* **2013**, *15*, 10094–10111.

(12) Huang, X.; Lee, T. J. A Procedure for Computing Accurate ab Initio Quartic Force Fields: Application to HO<sub>2</sub><sup>+</sup> and H<sub>2</sub>O. *J. Chem. Phys.* **2008**, *129*, 44312.

(13) Huang, X.; Lee, T. J. Accurate ab Initio Quartic Force Fields for NH<sub>2</sub><sup>−</sup> and CCH<sup>−</sup> and Rovibrational Spectroscopic Constants for Their Isotopologs. *J. Chem. Phys.* **2009**, *131*, 104301.

(14) Huang, X.; Valeev, E. F.; Lee, T. J. Comparison of One-Particle Basis Set Extrapolation to Explicitly Correlated Methods for the Calculation of Accurate Quartic Force Fields, Vibrational Frequencies, and Spectroscopic Constants: Application to H<sub>2</sub>O, N<sub>2</sub>H<sup>+</sup>, NO<sub>2</sub><sup>+</sup>, and C<sub>2</sub>H<sub>2</sub>. *J. Chem. Phys.* **2010**, *133*, 244108.

(15) Karton, A.; Martin, J. M. L. Performance of W4 Theory for Spectroscopic Constants and Electrical Properties of Small Molecules. *J. Chem. Phys.* **2010**, *133*, 144102.

(16) Kong, L.; Bischoff, F. A.; Valeev, E. F. Explicitly Correlated R12/F12 Methods for Electronic Structure. *Chem. Rev.* **2012**, *112*, 75–107.

(17) Hättig, C.; Klopper, W.; Köhn, A.; Tew, D. P. Explicitly Correlated Electrons in Molecules. *Chem. Rev.* **2012**, *112*, 4–74.

(18) Rauhut, G.; Knizia, G.; Werner, H.-J. Accurate Calculation of Vibrational Frequencies Using Explicitly Correlated Coupled-Cluster Theory. *J. Chem. Phys.* **2009**, *130*, 054105.

(19) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796.

(20) Köhn, A.; Richings, G. W.; Tew, D. P. Implementation of the Full Explicitly Correlated Coupled-Cluster Singles and Doubles Model CCSD-F12 with Optimally Reduced Auxiliary Basis Dependence. *J. Chem. Phys.* **2008**, *129*, 201103.

(21) Adler, T. B.; Knizia, G.; Werner, H.-J. A Simple and Efficient CCSD(T)-F12 Approximation. *J. Chem. Phys.* **2007**, *127*, 221106.

(22) Knizia, G.; Adler, T. B.; Werner, H.-J. Simplified CCSD(T)-F12 Methods: Theory and Benchmarks. *J. Chem. Phys.* **2009**, *130*, 054104.

(23) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; O'Neill, D. P.; Palmieri, P.; Peng, D.; Pflüger, K.; Pitzer, R. M.; Reiher, M.; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M. *MOLPRO*, A Package of ab Initio Programs, version 2012.1; University College Cardiff Consultants Ltd.: Wales, U.K., 2012.

(24) Hättig, C.; Tew, D. P.; Köhn, A. Communications: Accurate and Efficient Approximations to Explicitly Correlated Coupled-Cluster Singles and Doubles, CCSD-F12. *J. Chem. Phys.* **2010**, *132*, 231102.

(25) Tew, D. P.; Klopper, W.; Neiss, C.; Hättig, C. Quintuple-Zeta Quality Coupled-Cluster Correlation Energies with Triple-Zeta Basis Sets. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1921–1930.

(26) Tew, D. P.; Klopper, W.; Neiss, C.; Hättig, C. Comment on: “Quintuple-Zeta Quality Coupled-Cluster Correlation Energies with Triple-Zeta Basis Sets” by Tew, D. P.; Klopper, W.; Neiss, C.; C. Hättig [erratum]. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6325.

(27) Kozuch, S.; Gruzman, D.; Martin, J. M. L. DSD-BLYP: A General Purpose Double Hybrid Density Functional Including Spin Component Scaling and Dispersion Correction. *J. Phys. Chem. C* **2010**, *114*, 20801–20808.

(28) Furche, F.; Ahlrichs, R.; Hättig, C.; Klopper, W.; Sierka, M.; Weigend, F. *Turbomole*. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013**, *4*, 91–100.

(29) Peterson, K. A.; Adler, T. B.; Werner, H.-J. Systematically Convergent Basis Sets for Explicitly Correlated Wavefunctions: The Atoms H, He, B-Ne, and Al-Ar. *J. Chem. Phys.* **2008**, *128*, 084102.

(30) Yousaf, K. E.; Peterson, K. A. Optimized Auxiliary Basis Sets for Explicitly Correlated Methods. *J. Chem. Phys.* **2008**, *129*, 184108.

(31) Yousaf, K. E.; Peterson, K. A. Optimized Complementary Auxiliary Basis Sets for Explicitly Correlated Methods: Aug-Cc-pVnZ Orbital Basis Sets. *Chem. Phys. Lett.* **2009**, *476*, 303–307.

(32) Noga, J.; Šimunek, J. On the One-Particle Basis Set Relaxation in R12 Based Theories. *Chem. Phys.* **2009**, *356*, 1–6.

(33) Dunning, T. H.; Peterson, K. A.; Wilson, A. K. Gaussian Basis Sets for Use in Correlated Molecular Calculations. X. The Atoms Aluminum through Argon Revisited. *J. Chem. Phys.* **2001**, *114*, 9244.

(34) Bauschlicher, C. W.; Partridge, H. The Sensitivity of B3LYP Atomization Energies to the Basis Set and a Comparison of Basis Set Requirements for CCSD(T) and B3LYP. *Chem. Phys. Lett.* **1995**, *240*, 533–540.

(35) Martin, J. M. L. Basis Set Convergence Study of the Atomization Energy, Geometry, and Anharmonic Force Field of SO<sub>2</sub>: The Importance of Inner Polarization Functions. *J. Chem. Phys.* **1998**, *108*, 2791.

(36) Bell, R. D.; Wilson, A. K. SO<sub>3</sub> Revisited: Impact of Tight *d* Augmented Correlation Consistent Basis Sets on Atomization Energy and Structure. *Chem. Phys. Lett.* **2004**, *394*, 105–109.

(37) Martin, J. M. L. Heats of Formation of Perchloric Acid, HClO<sub>4</sub>, and Perchloric Anhydride, Cl<sub>2</sub>O<sub>7</sub>. Probing the Limits of W1 and W2 Theory. *J. Mol. Struct. Theochem* **2006**, *771*, 19–26.

(38) Huber, K. P.; Herzberg, G. *Constants of Diatomic Molecules*; Van Nostrand Reinhold: New York, 1979.

(39) Bürger, H.; Schulz, P.; Jacob, E.; Fahnle, M. The Fourier-Transform IR-Spectra of ClF and BrF. *Z. Naturforsch., A: Phys. Sci.* **1986**, *41*, 1015–1020.

(40) Irikura, K. K. Experimental Vibrational Zero-Point Energies: Diatomic Molecules. *J. Phys. Chem. Ref. Data* **2007**, *36*, 389.

(41) Strey, G.; Mills, I. M. Anharmonic Force Field of Acetylene. *J. Mol. Spectrosc.* **1976**, *59*, 103–115.

(42) Saouli, A.; Bredohl, H.; Dubois, I.; Fayt, A. FT Infrared Spectra of ClCN between 1200 and 5000 cm<sup>−1</sup> and Global Rovibrational Analysis of the Main Isotopomers. *J. Mol. Spectrosc.* **1995**, *174*, 20–50.

(43) Teffo, J.-L.; Sulakshina, O. N.; Perevalov, V. I. Effective Hamiltonian for Rovibrational Energies and Line Intensities of Carbon Dioxide. *J. Mol. Spectrosc.* **1992**, *156*, 48–64.

(44) Liou, H. T.; Huang, K. L. Fermi-Resonances, Potential Energy Function, and Cluster-like Spectra of the Ground Electronic State of CS<sub>2</sub>. *Chem. Phys.* **1999**, *246*, 391–431.

(45) Benedict, W. S.; Gailar, N.; Plyler, E. K. Rotation-Vibration Spectra of Deuterated Water Vapor. *J. Chem. Phys.* **1956**, *24*, 1139.

(46) Allen, H. C.; Plyler, E. K. Infrared Spectrum of Hydrogen Sulfide. *J. Chem. Phys.* **1956**, *25*, 1132.

(47) Smith, A. M.; Coy, S. L.; Klemperer, W.; Lehmann, K. K. Fourier Transform Spectra of Overtone Bands of HCN from 5400 to 15100 cm<sup>−1</sup>. *J. Mol. Spectrosc.* **1989**, *134*, 134–153.

(48) Teffo, J.-L.; Chédin, A. Internuclear Potential and Equilibrium Structure of the Nitrous Oxide Molecule from Rovibrational Data. *J. Mol. Spectrosc.* **1989**, *135*, 389–409.

- (49) Lahaye, J.-G.; Vandenhaute, R.; Fayt, A. CO<sub>2</sub> Laser Saturation Stark Spectra and Global Rovibrational Analysis of the Main Isotopic Species of Carbonyl Sulfide (OC<sup>34</sup>S, O<sup>13</sup>CS, and <sup>18</sup>OCS). *J. Mol. Spectrosc.* **1987**, *123*, 48–83.
- (50) Lafferty, W. J.; Pine, A. S.; Flaud, J. M.; Camy-Peyret, C. The 2ν<sub>3</sub> Band of <sup>32</sup>S<sup>16</sup>O<sub>2</sub>: Line Positions and Intensities. *J. Mol. Spectrosc.* **1993**, *157*, 499–511.
- (51) Huang, X.; Schwenke, D. W.; Lee, T. J. Rovibrational Spectra of Ammonia. I. Unprecedented Accuracy of a Potential Energy Surface Used with Nonadiabatic Corrections. *J. Chem. Phys.* **2011**, *134*, 044320.
- (52) Martin, J. M. L.; Taylor, P. R. The Geometry, Vibrational Frequencies, and Total Atomization Energy of Ethylene. A Calibration Study. *Chem. Phys. Lett.* **1996**, *248*, 336–344.
- (53) Tarczay, G.; Miller, T. A.; Czako, G.; Császár, A. G. Accurate ab Initio Determination of Spectroscopic and Thermochemical Properties of Mono- and Dichlorocarbenes. *Phys. Chem. Chem. Phys.* **2005**, *7*, 2881–2893.
- (54) Demaison, J.; Margulès, L.; Martin, J. M. L.; Boggs, J. E. Anharmonic Force Field, Structure, and Thermochemistry of CF<sub>2</sub> and CCl<sub>2</sub>. *Phys. Chem. Chem. Phys.* **2002**, *4*, 3282–3288.
- (55) Wang, X.-G.; Carrington, T. An Accurate Potential Energy Surface for Methane. In *68th International Symposium on Molecular Spectroscopy*, Department of Chemistry, Ohio State University, Columbus, Ohio, 2013.
- (56) Yachmenev, A.; Yurchenko, S. N.; Jensen, P.; Thiel, W. A New “Spectroscopic” Potential Energy Surface for Formaldehyde in Its Ground Electronic State. *J. Chem. Phys.* **2011**, *134*, 244307.
- (57) Yachmenev, A.; Yurchenko, S. N.; Ribeyre, T.; Thiel, W. High-Level ab Initio Potential Energy Surfaces and Vibrational Energies of H<sub>2</sub>CS. *J. Chem. Phys.* **2011**, *135*, 074302.
- (58) Wang, D.; Shi, Q.; Zhu, Q.-S. An ab Initio Quartic Force Field of PH<sub>3</sub>. *J. Chem. Phys.* **2000**, *112*, 9624.
- (59) Hänninen, V.; Halonen, L. Calculation of Spectroscopic Parameters and Vibrational Overtones of Methanol. *Mol. Phys.* **2003**, *101*, 2907–2916.
- (60) Skokov, S.; Peterson, K. A.; Bowman, J. M. An Accurate Ab Initio HOCl Potential Energy Surface, Vibrational and Rotational Calculations, and Comparison with Experiment. *J. Chem. Phys.* **1998**, *109*, 2662.
- (61) Schuurman, M. S.; Allen, W. D.; Schaefer, H. F. The ab Initio Limit Quartic Force Field of BH<sub>3</sub>. *J. Comput. Chem.* **2005**, *26*, 1106–1112.
- (62) Stanton, J. F.; Gauss, J.; Harding, M. E.; Szalay, P. G.; Auer, A. A.; Bartlett, R. J.; Benedikt, U.; Berger, C.; Bernholdt, D. E.; Bomble, Y. J.; Cheng, L.; Christiansen, O.; Heckert, M.; Heun, O.; Huber, C.; Jagau, T.-C.; Jonsson, D.; Juselius, J.; Klein, K.; Lauderdale, W. J.; Matthews, D. A.; Metzroth, T.; Mück, L. A.; O’Neill, D. P.; Price, D. R.; Prochnow, E.; Puzzarini, C.; Ruud, K.; Schiffman, F.; Simmons, C.; Schwalbach, W.; Stopkiewicz, S.; Tajti, A.; Vázquez, J.; Wang, F.; Watts, J. D.; Almlöf, J.; Taylor, P. R.; Helgaker, T.; Jensen, H. J. A.; Jørgensen, P.; Olsen, J.; Mitin, A. V.; van Wüllen, C. *CFOUR*, Department of Chemistry, University of Mainz, Germany, 2010.
- (63) Marchetti, O.; Werner, H.-J. Accurate Calculations of Intermolecular Interaction Energies Using Explicitly Correlated Wave Functions. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3400–3409.
- (64) Köhn, A. Explicitly Correlated Connected Triple Excitations in Coupled-Cluster Theory. *J. Chem. Phys.* **2009**, *130*, 131101.
- (65) Dunham, J. The Energy Levels of a Rotating Vibrator. *Phys. Rev.* **1932**, *41*, 721–731.
- (66) Karton, A.; Martin, J. M. L. Comment on: “Estimating the Hartree–Fock Limit from Finite Basis Set Calculations” [Jensen, F. (2005) *Theor. Chem. Acc.* 113:267]. *Theor. Chem. Acc.* **2005**, *115*, 330–333.
- (67) Schwenke, D. W. The Extrapolation of One-Electron Basis Sets in Electronic Structure Calculations: How It Should Work and How It Can Be Made to Work. *J. Chem. Phys.* **2005**, *122*, 14107.
- (68) Hill, J. G.; Mazumder, S.; Peterson, K. A. Correlation Consistent Basis Sets for Molecular Core-Valence Effects with Explicitly Correlated Wave Functions: The Atoms B–Ne and Al–Ar. *J. Chem. Phys.* **2010**, *132*, 054108.
- (69) Martin, J. M. L. A Fully Ab Initio Potential Curve of near-Spectroscopic Quality for OH<sup>−</sup> Ion: Importance of Connected Quadruple Excitations and Scalar Relativistic Effects. *Spectrochim. Acta, Part A* **2001**, *57*, 875–885.