What levels of coupled cluster theory are appropriate for transition metal systems? A study using near exact quantum chemical values for 3d transition metal binary compounds

Diptarka Hait,[*,†,‡] Norman M. Tubman,[†,¶] Daniel S. Levine,[†] K. Birgitta Whaley,[†] and Martin Head-Gordon[*,†,‡]

†*Kenneth S. Pitzer Center for Theoretical Chemistry, Department of Chemistry, University of California, Berkeley, California 94720, USA*

‡*Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

¶*Quantum Artificial Intelligence Lab. (QuAIL), Exploration Technology Directorate, NASA Ames Research Center, Moffett Field, CA 94035, USA*

E-mail: diptarka@berkeley.edu; mhg@cchem.berkeley.edu

Abstract

Transition metal compounds are traditionally considered to be challenging for standard quantum chemistry approximations like coupled cluster (CC) theory, which are usually employed to validate lower level methods like density functional theory (DFT). To explore this issue, we present a database of bond dissociation energies (BDEs) for 74 spin states of 69 diatomic species containing a 3d transition metal atom and a main group element, in the moderately sized def2-SVP basis. The presented BDEs appear to have an (estimated) 3σ error less than 1 kJ/mol relative to the exact solutions to the non-relativistic Born-Oppenheimer Hamiltonian. These benchmark values were used to assess the performance of a wide range of standard single reference CC models, as the results should be beneficial for understanding the limitations of these models for transition metal systems. We find that interactions between metals and monovalent ligands like hydride and fluoride are well described by CCSDT. Similarly, CCSDTQ appears to be adequate for bonds between metals and nominally divalent ligands like oxide and sulphide. However, interactions with polyvalent ligands like nitride and carbide are more challenging, with even CCSDTQ(P)_Λ yielding errors on the scale of a few kJ/mol. We also find that many perturbative and iterative approximations to higher order terms either yield disappointing results, or actually worsen the performance relative to the baseline low level CC method, indicating that complexity does not always guarantee accuracy.

1 Introduction

Single reference methods like Kohn-Sham density functional theory (KS-DFT)[1,2] and coupled cluster (CC)[3,4] theory constitute the backbone of main group quantum chemistry. These methods implicitly rely on the assumption that one particular electronic configuration (i.e. Slater determinant) is overwhelmingly dominant over all others. While exceptions to this single reference assumption are known in main group chemistry (especially when bond dissociations[3,5] and polyradicaloids[6] are involved), KS-DFT has been spectacularly successful in predicting energies and properties for species comprised of main group elements alone.[2,7–9] CC theory has arguably been even more successful, with coupled cluster singles and doubles with perturbative triples (CCSD(T))[10] being generally considered to be the ‘gold standard’ that is readily capable of achieving errors below the widely accepted ‘chemical accuracy’ limit of 4 kJ/mol (i.e about 1 kcal/mol),[4] for predicting reaction energies and barrier heights.

The situation is quite different for transition metal containing systems. While there are many transition metal compounds that are clearly well described by a single Slater determinant, a far larger proportion of cases have many nearly degenerate low lying electronic configurations that can contribute significantly to the ground state, resulting in a breakdown of the implicit single reference assumption. There exists no formal barrier against the applicability of KS-DFT or the CC hierarchy for multireference systems as these are in principle capable of being exact. In practice however, existing approximations to KS-DFT and affordable CC theory have a rather questionable record when it comes to transition metal compounds.[11–13] The natural fit for such species are therefore explicitly multireference methods like complete active space configuration interaction (CASCI), which treats all configurations within an active space of orbitals and electrons on a variationally equal footing. However, the computational cost of CASCI scales exponentially with the size of the active space, with the modern limit of exact CASCI being around 22 electrons in 22 orbitals ($\approx 5 \times 10^{11}$ determinants), which requires months of supercomputer CPU time.[14] Practical CAS calculations therefore typically orbital optimize CAS wave functions (CASSCF[15]) from relatively small active spaces to minimize variational energy tas much as possible. The variational CASSCF energy is often subsequently augmented with perturbative corrections (usually via the CASPT2[16,17] approach). Such approaches however would miss nondynamical correlation outside the small initial active space that a CAS calculation with a larger active space could capture.

There exist approximations to CASCI that are capable of treating larger active spaces without compromising the fundamentally multireference nature of the method (unlike standard CC theory). Well known examples of such methods are density matrix renormalization group (DMRG),[18] quantum Monte-Carlo (QMC)[19–21] and selected configuration interaction (SCI).[22–24] Multireference coupled cluster methods, like Mk-MRCC[25,26] and internally contracted MRCC[27] also show promise for tackling multireference systems. The computational cost for these methods however still scales far too sharply with system size, preventing application to very large systems like the metal-electrolyte interface in electrochemistry, a full metalloenzyme active site or large MOFs, for example. DFT remains the only viable deterministic *ab initio* avenue to such large systems, necessitating functionals that are as accurate for transition metal chemistry as they are for the main group elements.

The development of highly accurate main group functionals was assisted by availability of even more accurate benchmark data against which these functionals could be trained and assessed.[2,7] A target DFT functional RMSE of 4 kJ/mol would require benchmark data with RMSE of about 0.4 kJ/mol, in order to ensure DFT errors are precise enough for fitting/assessments to be meaningful. This level of accuracy for main group chemistry is typically achieved via composite schemes like the W4[28] or HEAT[29] approaches, which rely on the great accuracy of CC methods for main group chemistry. Similarly accurate *ab initio* data for transition metal chemistry is however extremely difficult to come by, especially on account of uncertainties surrounding applicability of CC theories for such systems.

In this work, we present a set of benchmark bond dissociation energies (BDEs) for a total of 74 spin states of 69 3d transition metal binary compounds using the extrapolated Adaptive Sampling Configuration Interaction (ASCI) method[23,30] in a double zeta def2-SVP basis.[31] Such a basis is too small for direct assessment of density functionals, but is large enough to be useful in assessing CC methods (in the same basis) to determine their reliability for transition metal containing systems. Our benchmark values appear to have an estimated 3σ (i.e. 3 standard deviation) error that is less than 1 kJ/mol for all cases, relative to the exact non-relativistic Born-Oppenheimer results in the same basis (as estimated from the extrapolation procedure). This indicates a 1σ error smaller than the target of 0.4 kJ/mol, which should be satisfactory as benchmarks for comparing against CC theory within the same basis. We then present an assessment of reasonably affordable CC methods against this benchmark to attempt to understand the limitations of CC models for transition metal systems.

2 Dataset Species and previous work on them

Table 1: Experimental ground electronic states for the species considered.[32–34] The quantities within parentheses are the best theoretical estimates[34] if they are not marked by *. Quantities marked with * are our best guesses regarding spin of the ground state, using isoelectronic/isovalent species as a guide.

	Sc	**Ti**	**V**	**Cr**	**Mn**	**Fe**	**Co**	**Ni**	**Cu**	**Zn**
H	$^{1}\Sigma+$	$^{4}\Phi$	$^{5}\Delta$	$^{6}\Sigma+$	$^{7}\Sigma+$	$^{4}\Delta$	$^{3}\Phi$	$^{2}\Delta$	$^{1}\Sigma+$	$^{2}\Sigma+$
F	$^{1}\Sigma+$	$^{4}\Phi/^{4}\Sigma_{-}$	$(^{5}\Pi/^{5}\Delta)$	$^{6}\Sigma+$	$^{7}\Sigma+$	$^{6}\Delta$	$^{3}\Phi$	$^{2}\Pi$	$^{1}\Sigma+$	$^{2}\Sigma+$
Cl	$^{1}\Sigma+$	$^{4}\Phi$	$(^{5}?)^*$	$^{6}\Sigma+$	$^{7}\Sigma+$	$^{6}\Delta$	$(^{3}?)^*$	$(^{2}?)^*$	$^{1}\Sigma+$	$^{2}\Sigma+$
O	$^{2}\Sigma+$	$^{3}\Delta$	$^{4}\Sigma-$	$^{5}\Pi$	$^{6}\Sigma+$	$^{5}\Delta$	$^{4}\Delta$	$^{3}\Sigma_{-}$	$^{2}\Pi$	$^{1}\Sigma+$
S	$^{2}\Sigma+$	$^{3}\Delta$	$^{4}\Sigma-$	$^{5}\Pi$	$^{6}\Sigma+$	$^{5}\Delta$	$^{4}\Delta$	$^{3}\Sigma_{-}$	$^{2}\Pi$	$^{1}\Sigma+$
C	$^{4}\Pi$	$^{3}\Sigma$	$(^{2}\Delta)$	$(^{3}\Sigma_{-})$	$(^{4}?)^*$	$^{3}\Delta$	$^{2}\Sigma+$	$^{1}\Sigma+$	$(^{2}?/^{4}?)^*$	$^{3}\Sigma-$
N	$^{1}\Sigma+$	$^{2}\Sigma+$	$^{3}\Delta$	$^{4}\Sigma-$	$^{5}\Pi$	$^{4}?$	$(^{5}?)^*$	$(^{2}\Pi)$	$^{3}\Sigma-$	$^{4}\Sigma-$

We have studied 70 3d transition metal binary compounds of the form MX (where M is the metal and X one of H,C,N,O,F,S and Cl). This is not representative of standard transition metal chemistry, as the common octahedral and tetrahedral ligand field motifs are entirely missing from the dataset. However, near-exact quantum chemistry is tractable for these simple linear species, making them an ideal test case for determining how standard methods perform. Indeed, a number of DFT studies have been carried out for such molecules[32,33,35] but correlated wave function studies have received greater attention in recent years. This was partly catalyzed by Xu et.al.,[11] who suggested that CC methods like CCSD(T) gave subpar performance relative to DFT methods for reproducing experimental BDEs for a subset of 20 compounds (the 3dMLBE20 database[11]). Subsequent work[36] however argued that standard CC theory was adequate for the same 20 compounds. In fact, another study[37] presented benchmark BDE data for the 3dMLBE20 database using the HEAT protocol,[29] which would be very accurate if the high level coupled cluster theory employed therein was adequate. Efforts have also been made to assess the importance of multireference contributions using multireference CC,[12] and very recently, an auxiliary-field QMC study of a subset of these compounds has been carried out,[13] though the reported error bars appear to be rather too large (on the scale of 4-8 kJ/mol).

In our opinion, a rather important problem with many of the aforementioned works is an excessive reliance on questionable experimental BDE data. Most of these papers do indeed note that much of the available data is unreliable, and at times advocate revising specific experimental values[12,37] or select one specific number out of competing ones.[33] Therefore, we define our objective to be the production of quantum chemistry benchmarks for these species, that approach the exact solutions to the Born-Oppenheimer non-relativistic Schrödinger equation in a finite basis set. Approximate CASCI methods like SCI and DMRG offer an obvious route to solving the Born-Oppenheimer non-relativistic Hamiltonian, as they are tractable for systems of this size (indeed, the FeS molecule has been studied with SCI recently[38]) while being systematically improvable to the exact result. We have therefore chosen to use an SCI method to generate a sufficiently accurate dataset that can stand on its own legs and have ignored experimental BDEs altogether for now. The only experimental quantity we have taken into account is the ground spin state of the dataset species (presented in Table 1), which we employed to determine what spin states to target.

3 Brief overview of Quantum Chemistry Approaches

3.1 Adaptive Sampling Configuration Interaction (ASCI)

SCI methods assume that a large fraction of determinants within the full CASCI Hilbert space have very small coefficients in the exact wave function, and a good approximation can be generated by simply selecting the (expected) top contributors to the true wave function, followed by diagonalization within that reduced subspace.[22,39] The energy of the SCI wave function can be further lowered via rotation to an approximate natural orbital basis,[3,23,39] or orbital optimization within the active space.[40,41] The remaining correlation energy could be estimated via other approaches like second order perturbation theory (PT2).[23,42] SCI+PT2 *absolute* energies in fact often readily converge to the chemical accuracy limit of 4 kJ/mol for even strongly correlated transition metal systems,[23,24] making them a useful option for transition metal quantum chemistry. SCI methods nonetheless cannot avoid the asymptotic exponential scaling behavior of CASCI, as the total number of important configurations grows exponentially with system size. The computational cost however grows much more slowly, and SCI+PT2 energies can be calculated for sizeable systems like iron porphyrin[40,41,43] or Fe_4S_8 ferrodoxin[44] that lie well beyond the reach of exact CASCI solvers.

Multiple flavors of SCI, utilizing different selection rules and search spaces, exist in the literature.[22–24,45] The ASCI approach is the specific flavor that we have developed (as described in detail elsewhere[23,30]), and will therefore use in our study. The ASCI energy is augmented with a deterministic Epstein-Nesbet PT2 correction.[46–48] Further accuracy gains are possible by extrapolating the ASCI+PT2 energy against the PT2 correction to the zero PT2 limit of the exact wave function. We observe nearly linear behavior of the ASCI+PT2 energy against the PT2 correction, which is consistent with behavior witnessed for other SCI methods.[49,50] Similar linear fits are also employed by DMRG to reach more accurate final estimates.[18] The standard deviation in the estimated intercept from the ASCI+PT2 linear fit also quite naturally supplies an error metric for the final estimated ASCI+PT2 energies. Further details about the extrapolation have been given in Appendix C.

3.2 Coupled Cluster Methods

A fairly large number of CC methods have been analyzed in this present work, and we therefore provide a very brief primer to these methods: see Ref 4 for further details. CC methods approximate the exact wave function by generating excitations out of a reference determinant with an exponential operator. Given a reference determinant $|\Phi_0\rangle$ with occupied spin orbitals $i, j \ldots$ and virtuals $a, b \ldots$, the CC wave function to n excitations is:

$$|\Psi\rangle = e^{\hat{T}} |\Phi_0\rangle \tag{1}$$

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 \ldots \hat{T}_n \tag{2}$$

$$\hat{T}_1 = \sum_{i,a} t_i^a a_a^\dagger a_i; \; \hat{T}_2 = \sum_{i<j,a<b} t_{ij}^{ab} a_a^\dagger a_b^\dagger a_j a_i; \ldots \tag{3}$$

Closed equations for the CC energy and the t amplitudes $t_{i\cdots}^{a\cdots}$ are obtained by projecting the Schrödinger equation with the CC wave function of Eqn 1 into the subspace of Hilbert space spanned by $|\Phi_0\rangle$ and all excitations out of it to order n. The best known member of the CC family is coupled cluster singles and doubles (CCSD)[3] where $\hat{T}$ is truncated at $\hat{T}_1 + \hat{T}_2$. More accurate is CCSDT[51] (where all triples are incorporated as well, via $\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3$) and and even more accurate is CCSDTQ[52] (where quadruples are also fully accounted for). In general, CCSDT... n (which we henceforth refer to as CC(n)) has a computational cost that scales as $O(N_o^n N_v^{n+2})$ (where N_o is number of occupied orbitals and N_v is the number of virtuals), which quickly becomes very large with n. CCSDTQ is the highest order exact CC method we have been able to afford in this present work.

3.2.1 Perturbative Corrections

The accuracy of CC methods can be augmented by incorporation of a perturbative correction for higher order excitations not present in the truncated exponential wave function of Eqn 1. The classic example is CCSD(T),[10] where the effect of connected triples is accounted for via incorporation of terms that are exact to fourth order in Moller Plesset perturbation theory. Other examples include CCSDT(Q)/A,[53] CCSDT(Q)/B,[53] CCSD[T][54] etc. Perturbative corrections to CC($n-1$) are noniterative and tend to scale as $O(N_o^n N_v^{n+1})$, which makes them more affordable than an exact CC(n) calculation.

A related class of perturbative corrections are the Λ amplitude based corrections, which employ the left hand solutions to the projected CC equations as well. These methods have the same scaling as the more traditional perturbative corrections, but are considerably more computationally expensive. Examples of this category include CCSD(T)$_\Lambda$,[55] CCSDT(Q)$_\Lambda$[56] etc. An important member of this class is CCSD(2),[57,58] which explicitly finds a PT2 correction to CCSD, and thus accounts for some quadruple excitations along with the more standard triples correction. CCSD(2) has a somewhat better track record than CCSD(T) for some strongly correlated systems,[8] due to renormalization of the one body terms in the former that mitigates some unphysical behavior that can be observed in the latter.

3.2.2 Iterative Approximations

It is also possible to preserve only certain terms coming from higher order excitations while solving CC equations, to generate iterative approximations to the full higher order CC calculation. Examples of such methods are CCSDT-1a,[59] CCSDT-1b,[54] CCSDT-3[54] (each being a better approximation to CCSDT than the preceding one and are consequently more computationally expensive), and their higher order generalizations CCSDTQ-1a,1b,3 etc. These iterative approximation for $\hat{T}_n$ therefore has a cost that is intermediate between the exact CC$(n-1)$ problem and the exact CC(n) problem, depending on the complexity of the surviving terms. We refer the interested reader to Refs 53 and 56 for greater discussion about approximate corrections to CC methods, beyond the bare bones picture presented here.

4 Computational Methods

All ASCI+PT2, CCSD, CCSD(T) and CCSD(2) calculations were carried out with Q-Chem 5.2[60] and all other CC calculations were performed with MRCC.[61,62] The geometries were obtained from the compilation in Ref 32 (experimental wherever possible, BP86 optimized otherwise) and are given in the Supporting Information. The spin states employed are listed in Table 1, although alternative spin states were also considered for some species (as noted in Tables 2 and 6). The def2-SVP basis[31] was used for all calculations. Restricted open-shell Hartree-Fock (ROHF) orbitals were used for all calculations, in order to minimize spin contamination in the CC wave function (and eliminate it in the energy). The 1s orbitals of C,N,O and F were held frozen (i.e. a He core), as were the 1s, 2s and 2p orbitals for S, Cl and the metal atoms (i.e. a Ne core). The error in the extrapolated ASCI+PT2 numbers is found from the standard deviation in the intercept of the linear fit of the ASCI energy against the PT2 correction, with the reported error being thrice this standard deviation (i.e.

3σ error, which should correspond to a 99.7% confidence interval under the not quite justified assumption of a normal distribution).

We note that the electronic structure of many of the species considered is quite challenging at the HF level, with multiple saddle points in orbital space and, in some cases, multiple local minima. This is a significant challenge as it is difficult to be certain if we had reached the self-consistent field (SCF) global minima. We ran stability analysis on all the Q-Chem SCF solutions to ensure that they were at least local minima, but this was not possible with the MRCC package. The lowest energy MRCC SCF solution was often above the energy of the best Q-Chem SCF solution despite our best efforts to coax MRCC into locating the same minimum as Q-Chem. The converse was not true-i.e., no MRCC SCF solution was below the corresponding Q-Chem one in energy. A comparison between ASCI and higher order CC is not straightforward for systems where the two codes operate on different references, as it is quite likely that the methods will converge to different eigenstates of the Hamiltonian. The cases with SCF differences were consequently eliminated from the statistical analysis. We have noticed similar local extremum problems at the SCF level with larger basis sets (def2-TZVPP[31]), and different SCF references in past assessments could easily account for a fair bit of variation in reported numbers. A full listing of Q-Chem SCF energies has been supplied in the Supporting Information, in order to guide any future assessments.

5 Results and Discussion

5.1 Bond Dissociation Energies

Table 2 lists all the extrapolated ASCI+PT2 bond dissociation energies that we have calculated to better than (estimated) 3σ error of 1 kJ/mol. The list of error estimates, along with geometry and spin state listings, can be found in the Supporting Information. We were unable to converge the energies of CoF, (triplet) CoCl and (quartet) FeN to the desired level of accuracy, and therefore refrained from reporting them in Table 2. We consequently have

Table 2: Bond Dissociation Energies (BDEs) predicted by extrapolated ASCI+PT2 for the def2-SVP basis, in kJ/mol. Multiple spin states were considered for some species, resulting in BDE values for high spin (HS) and low spin (LS) cases. All values presented have an (estimated) 3σ error under 1 kJ/mol (typically much less). IC stands for insufficiently converged (i.e. high error) cases that have consequently not been reported.
*MRCC could not find the lowest energy SCF state and so CC data is incomplete.

	Sc	**Ti**	**V**	**Cr**	**Mn**	**Fe**	**Co**	**Ni**	**Cu**	**Zn**
H	210.4	185.7*	199.6	202.3	131.9	137.7 (hextet) 134.8 (quartet)	162.7*	195.3	248.9	79.3
F	578.2	500.3*	469.6	431.2	387.3	399.3	IC (quintet)* IC (triplet)*	319.1 (quartet) 275.1 (doublet)	372.5	260.1
Cl	413.2	364.4*	349.2	331.0*	289.4	286.0	245.1 (quintet)* IC (triplet)*	226.0 (quartet) 270.5 (doublet)	333.1	180.3
O	615.9	605.3	562.6*	377.8	268.6	291.3	291.4	225.7	226.1	77.5
S	410.2	379.5	346.3*	248.2	173.1	223.3	225.6	224.9*	230.1	72.7
C	210.4	326.5	302.4	256.9*	218.4	282.4	275.2*	299.8*	207.7 (quartet) 182.1 (doublet)	92.2
N	275.0	415.5	404.2	300.7	77.7	134.2 (hextet)* IC (quartet)	122.9	93.1*	123.4	-15.0 (quartet) -132.2 (doublet)

BDEs for 74 spin states for 69 compounds, although we miss the likely ground spin states for CoCl and FeN. We also note that ZnN is unstable with respect to dissociation to the ground state Zn and N atoms, but it can be bound on excited state surfaces. Indeed, Zn seems to form significantly weaker bonds (relative to all other metals, for any given ligand) as a consequence of the full d shell. In contrast, the strongest bonds are typically formed by Sc, Ti and V, which have relatively more empty d orbitals.

A detailed performance analysis of CC methods is given in subsequent subsections. This is organized based on ligand valency, since we found that the monovalent H, F and Cl containing species have very similar behavior, as do the divalent O and S containing species. To a lesser extent, the polyvalent C and N ligand containing species also show similar features. We note that donation of ligand lone pairs into metal d orbitals is a possibility for all non-hydride species, resulting in a deviation from typical ligand valency, but we nonetheless find it useful to classify ligands based on nominal valency for the purpose of analyzing trends.

5.2 Monovalent Ligands

Table 3: Root Mean Squared Error (RMSE) and Mean Error (ME) in BDEs predicted by CC methods for monovalent ligands, against the ASCI benchmark (in kJ/mol). The CC methods have been listed roughly in order of their complexity. Not included: CrCl, and compounds of Co and Ti, due to lack of complete data.

Method	Hydride		Fluoride		Chloride	
	RMSE	ME	RMSE	ME	RMSE	ME
CCSD	8.9	-5.5	8.5	-7.1	9.2	-6.9
CCSD(T)	6.9	-4.1	4.2	-3.2	8.2	-4.6
CCSD(T)$_\Lambda$	1.6	-0.3	2.0	-1.1	1.9	-1.4
CCSD(2)	1.5	-0.9	1.4	-0.5	2.1	-1.6
CCSDT-1a	3.9	1.4	2.3	1.1	3.5	0.7
CCSDT-1b	3.0	0.5	2.3	0.9	2.6	0.2
CCSDT-3	1.4	-0.3	1.3	0.0	1.3	-0.8
CCSDT	0.5	-0.3	1.6	-1.5	1.3	-1.2
CCSDT(Q)/A	2.7	-1.6	2.2	-1.7	4.4	-2.4
CCSDT(Q)/B	2.1	-1.2	1.6	-1.1	3.0	-1.7
CCSDT(Q)$_\Lambda$	0.8	-0.5	0.6	-0.5	1.2	-0.9
CCSDTQ-1a	2.0	-1.1	2.0	-1.7	2.6	-1.8
CCSDTQ-1b	1.4	-0.8	0.8	-0.6	1.5	-1.1
CCSDTQ-3	0.5	-0.4	0.7	-0.6	1.0	-0.9
CCSDTQ	0.2	-0.1	0.5	-0.4	0.6	-0.6

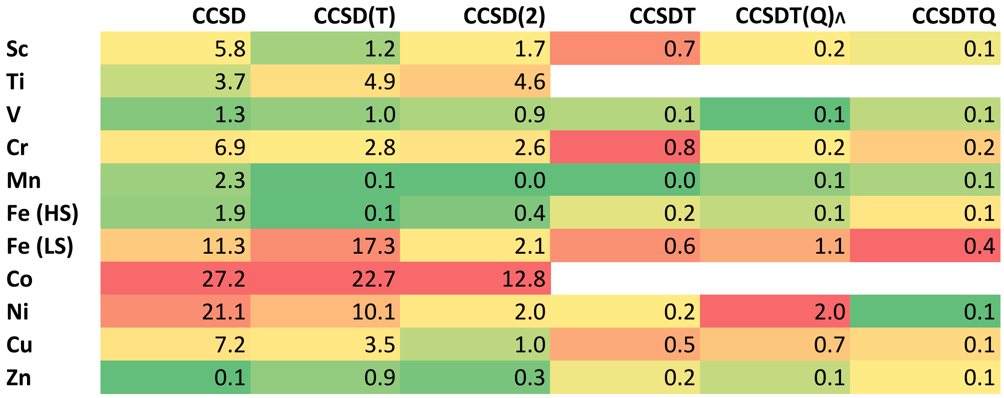
The statistical errors predicted by tested CC methods against the ASCI benchmark for monovalent ligand containing species is presented in Table 3, although data for some species (Co and Ti compounds, as well as CrCl) is missing from this analysis due to lack of complete CC data (as well as ASCI data, in the case of CoF and triplet CoCl). More detailed, species-specific analysis is provided in the next few subsections, but several general trends are immediately apparent from Table 3. As anticipated, CCSDTQ is the most accurate method, predicting very low RMSE that are well within our benchmark error bars for both hydrides and halides. In contrast, it is clear that CCSD (despite being exact for an ensemble of isolated two electron pairs) is nowhere near chemical accuracy for describing this single bond, mostly due to systematic underestimation of the metal-ligand interaction on account of missing correlations involving three or more electrons. CCSD(T) partially mitigates this, but not to a satisfactory extent as there is still a sizeable chunk of missing correlation energy. The more robust CCSD(2) and CCSD(T)$_\Lambda$ methods have significantly lower errors, suggesting that they are more appealing as perturbative corrections to bare CCSD. Local approximations to CCSD(2)/CCSD(T)$_\Lambda$ analogous to DLPNO-CCSD(T)[63] could therefore be of great utility for studying transition metal chemistry involving metal to ligand single bonds. A fair amount of the error nonetheless still originates from systematic underbinding. The full triples term in CCSDT appears to be sufficient for obtaining the correct BDE of hydrides, but systematically underbinds the halides by approximately 1.2-1.5 kJ/mol.

An interesting feature is that some complex methods statistically fail relative to simpler models. The $-1a$ and $-1b$ families of iterative corrections give quite disappointing results, with the performance of these iterative quadruples being worse than bare triples alone in many cases! Similarly, the (Q)/A and (Q)/B perturbative corrections either worsen predictive power relative to bare CCSDT or are about neutral, though the (Q)/B correction always has smaller error than (Q)/A. The -3 iterative approximation fares much better, but its greater computational expense makes the similarly accurate Λ perturbative correction more appealing. Overall, the robust members of the CC hierarchy (exact CC models and Λ based corrections) show effectively systematic improvement with complexity, till we are well within the error bars of our benchmark numbers.

5.2.1 Hydrides

We next consider the individual hydrides by tabulating the magnitudes of errors in BDE predictions for a subset of CC methods in Fig 1.

This shows that CCSD(T) appears to encounter some difficulty with the low spin (quartet) FeH, CoH and NiH, but is otherwise able to predict BDEs with an accuracy roughly similar to the chemical accuracy limit of 4 kJ/mol (i.e 1 kcal/mol), despite the larger RMSE in Table 3. CCSD(2) is almost universally much better, but it still has unacceptably high error for CoH. Unfortunately, lack of complete CC data for CoH and TiH prevents us from explicitly determining what is the minimum level of CC theory that has errors below sub

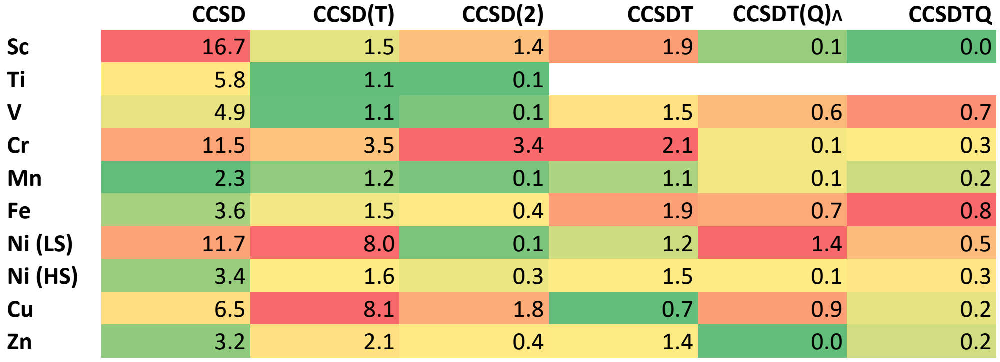


	CCSD	CCSD(T)	CCSD(2)	CCSDT	CCSDT(Q)Λ	CCSDTQ
Sc	5.8	1.2	1.7	0.7	0.2	0.1
Ti	3.7	4.9	4.6			
V	1.3	1.0	0.9	0.1	0.1	0.1
Cr	6.9	2.8	2.6	0.8	0.2	0.2
Mn	2.3	0.1	0.0	0.0	0.1	0.1
Fe (HS)	1.9	0.1	0.4	0.2	0.1	0.1
Fe (LS)	11.3	17.3	2.1	0.6	1.1	0.4
Co	27.2	22.7	12.8			
Ni	21.1	10.1	2.0	0.2	2.0	0.1
Cu	7.2	3.5	1.0	0.5	0.7	0.1
Zn	0.1	0.9	0.3	0.2	0.1	0.1

Figure 1: Magnitude of errors (in kJ/mol) of selected CC methods against the ASCI benchmark for all the hydrides. The highlighting is done on a per column basis (i.e. shows the relative level of difficulty encountered by a method for a given species, relative to all species.).

1 kJ/mol for these two species. CCSDT is adequate for the others, with CCSDTQ further lowering errors.

5.2.2 Fluorides



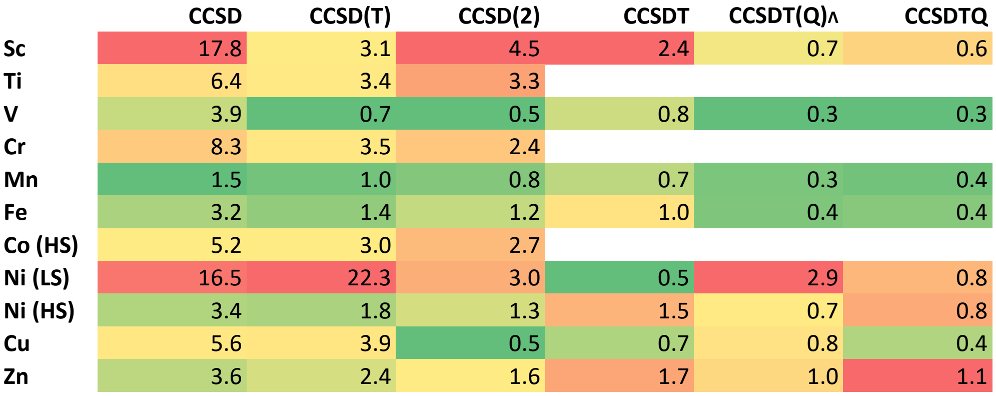
	CCSD	CCSD(T)	CCSD(2)	CCSDT	CCSDT(Q)Λ	CCSDTQ
Sc	16.7	1.5	1.4	1.9	0.1	0.0
Ti	5.8	1.1	0.1			
V	4.9	1.1	0.1	1.5	0.6	0.7
Cr	11.5	3.5	3.4	2.1	0.1	0.3
Mn	2.3	1.2	0.1	1.1	0.1	0.2
Fe	3.6	1.5	0.4	1.9	0.7	0.8
Ni (LS)	11.7	8.0	0.1	1.2	1.4	0.5
Ni (HS)	3.4	1.6	0.3	1.5	0.1	0.3
Cu	6.5	8.1	1.8	0.7	0.9	0.2
Zn	3.2	2.1	0.4	1.4	0.0	0.2

Figure 2: Magnitude of errors (in kJ/mol) of selected CC methods against the ASCI benchmark for all the fluorides. CoF is missing due to our inability to converge the benchmark to a 3σ estimated error of 1 kJ/mol.

Fig 2 and Table 3 show that fluorides behave very similarly to hydrides. Fig 2 shows that while CCSD(T) is challenged by low spin (doublet) NiF and CuF, it is generally very accurate for the rest of the species. CCSD(2) however remains the more robust performer overall. In contrast to the hydrides, CCSDT does not guarantee sub 1 kJ/mol error (possibly on account of metal electrons interacting with ligand lone pairs, which could not happen in hydrides) but

CCSDTQ appears to be always sufficient. While we were unable to converge the extrapolated ASCI+PT2 energy of CoF to the desired accuracy (3σ error below 1 kJ/mol), the variational ASCI solutions within the $M_s = 1$ subspace suggest that the high spin quintet state is the ground state within the def2-SVP basis, and not the experimentally observed triplet state. Furthermore, quintet CoF appears to be more challenging for CCSD(T) than any of the other species, and even CCSD(2) fails to significantly reduce error relative to CCSD(T). Prima-facie, triplet CoF is likely to be even more challenging (via analogy to chlorides, as discussed below).

5.2.3 Chlorides



	CCSD	CCSD(T)	CCSD(2)	CCSDT	CCSDT(Q)Λ	CCSDTQ
Sc	17.8	3.1	4.5	2.4	0.7	0.6
Ti	6.4	3.4	3.3			
V	3.9	0.7	0.5	0.8	0.3	0.3
Cr	8.3	3.5	2.4			
Mn	1.5	1.0	0.8	0.7	0.3	0.4
Fe	3.2	1.4	1.2	1.0	0.4	0.4
Co (HS)	5.2	3.0	2.7			
Ni (LS)	16.5	22.3	3.0	0.5	2.9	0.8
Ni (HS)	3.4	1.8	1.3	1.5	0.7	0.8
Cu	5.6	3.9	0.5	0.7	0.8	0.4
Zn	3.6	2.4	1.6	1.7	1.0	1.1

Figure 3: Magnitude of errors (in kJ/mol) of selected CC methods against the ASCI benchmark for all the chlorides. LS (triplet) CoCl is missing due to our inability to converge the benchmark to a 3σ estimated error of 1 kJ/mol.

The chlorides again exhibit behavior quite similar to hydrides and fluorides (as can be seen from Fig 3 and Table 3). CCSD(T) is satisfactory for most species, but gives quite disappointing performance for the low spin (doublet) NiCl states. Partially converged numbers (provided in the supporting information) suggest that LS (triplet) CoCl is very challenging for both CCSD(T) and CCSD(2), suggesting that it is quite a difficult molecule to model.

Table 4: Root Mean Squared Error (RMSE) and Mean Error (ME) in BDEs predicted by CC methods for divalent ligands, against the ASCI benchmark (in kJ/mol). The CC methods have been listed roughly in order of their complexity. VO, VS and NiS were not included as all data was not available.

Method	Oxide		Sulphide	
	RMSE	ME	RMSE	ME
CCSD	58.7	-55.4	44.7	-41.4
CCSD(T)	16.4	-12.3	9.7	-8.3
CCSD(T)$_\Lambda$	13.4	-9.3	14.3	-11.6
CCSD(2)	14.4	-11.7	14.4	-12.5
CCSDT-1a	24.5	17.9	9.5	2.9
CCSDT-1b	22.8	15.3	10.5	1.1
CCSDT-3	11.4	3.9	8.9	-2.3
CCSDT	6.2	-5.3	3.9	-3.4
CCSDT(Q)/A	25.0	-21.8	14.8	-12.7
CCSDT(Q)/B	6.7	-5.2	3.7	-0.2
CCSDT(Q)$_\Lambda$	2.5	0.1	0.9	0.2
CCSDTQ-1a	12.7	-11.5	8.3	-7.3
CCSDTQ-1b	3.8	-2.8	1.8	-0.9
CCSDTQ-3	3.9	-3.2	2.6	-2.1
CCSDTQ	1.2	0.3	0.9	0.3
CCSDTQ(P)/A	5.8	-5.2	4.7	-3.8
CCSDTQ(P)/B	7.1	-6.5	5.7	-4.7
CCSDTQ(P)$_\Lambda$	0.4	-0.2	0.5	-0.3

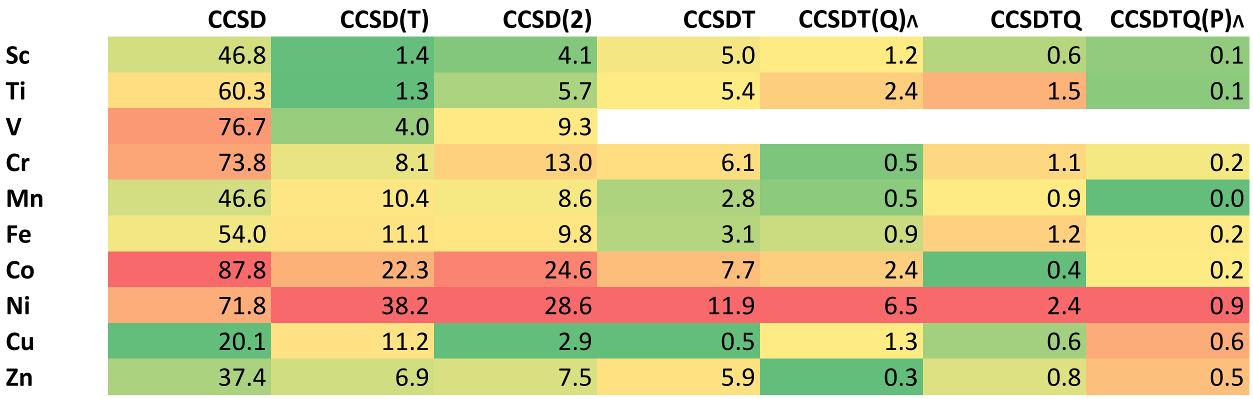
5.3 Divalent Ligands

The oxides and sulphides can be expected to be more challenging than hydrides and halides due to the potential of metal ligand double bonds, which would involve four electrons and thus could be difficult for even CCSDT to model. In fact, even partial triple bonds are potentially possible via donation of ligand lone pairs to the metal, possibly increasing the complexity of the interaction. The statistical values in Table 4 reveal that this is indeed the case. CCSD's performance is nothing short of disastrous, as it underestimates the bond energy by $\approx 40-55$ kJ/mol. Perturbative corrections help recover most of the missing correlation, but the BDEs are still systematically underestimated by $\approx$ 10 kJ/mol. Interestingly, CCSD(2) and CCSD(T)$_\Lambda$ do not represent significant improvements over CCSD(T) for oxides, and in fact they perform worse for sulphides (which was very rarely the case for the monovalent hydride and halide ligands). The CCSD(2) vs CCSD(T) performance difference for sulphides stems mostly from CCSD(2) underestimating BDEs relative to CCSD(T), which is perhaps an unintended byproduct of the renormalized one body terms present in the former to prevent divergent behavior.

The iterative triples approximations CCSDT-1a,1b and 3 yield quite disappointing performance as well, though they seem to have a bias towards overestimating bond energies versus underestimating them. Even the effect of full triples is insufficient, as CCSDT has an RMSE of 6 kJ/mol for oxides, and 4 kJ/mol for sulphides (most of which arose on account of systematically missing correlation energy coming from connected four electron interactions), which is quite sobering with regards to applying CC theory to arbitrary transition metal species involving metal-ligand multiple bonds. The (Q)/A perturbative correction significantly worsens performance, while the more robust (Q)/B correction merely manages to nearly break even relative to bare CCSDT. The much more expensive $(Q)_\Lambda$ correction however gives good performance, bringing RMSE quite a bit below the chemical accuracy limit. It is still not foolproof (as can be seen from the shabby performance for NiO shown in Fig 4) and the errors are too high to be of benchmark quality, but it is the most inexpensive method to give acceptable performance by itself.

Iterative quadruples corrections have a more mixed performance with CCSDTQ-1a being terrible, but CCSDTQ-1b and 3 improving upon CCSDT. Even CCSDTQ's full quadruples only results in RMSE ≈ 1 kJ/mol, indicating that the oxides are a challenging collection. CCSDTQ nonetheless appears to be an adequate level of theory for purposes other than DFT grade benchmark generation, since the metal-ligand interactions are at most a double bond with four electrons (if ligand lone pair donation is neglected), which the exact quadruples should be able to mostly handle. Pentuples corrections via $(P)_\Lambda$ improve results even further, though other perturbative or iterative approximations may not be worth the computational effort invested (as can be seen from the poor performance of (P)/A and (P)/B corrections).

5.3.1 Oxides



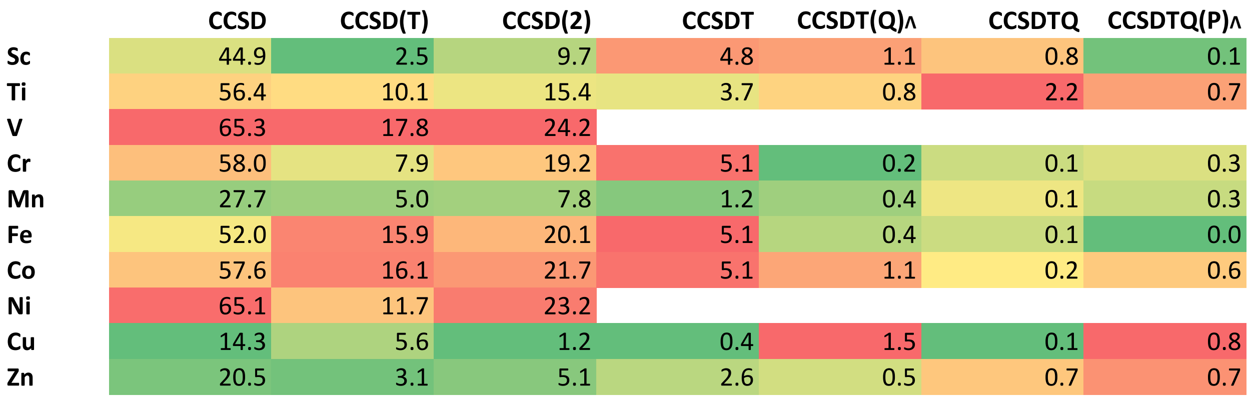
	CCSD	CCSD(T)	CCSD(2)	CCSDT	CCSDT(Q)Λ	CCSDTQ	CCSDTQ(P)Λ
Sc	46.8	1.4	4.1	5.0	1.2	0.6	0.1
Ti	60.3	1.3	5.7	5.4	2.4	1.5	0.1
V	76.7	4.0	9.3				
Cr	73.8	8.1	13.0	6.1	0.5	1.1	0.2
Mn	46.6	10.4	8.6	2.8	0.5	0.9	0.0
Fe	54.0	11.1	9.8	3.1	0.9	1.2	0.2
Co	87.8	22.3	24.6	7.7	2.4	0.4	0.2
Ni	71.8	38.2	28.6	11.9	6.5	2.4	0.9
Cu	20.1	11.2	2.9	0.5	1.3	0.6	0.6
Zn	37.4	6.9	7.5	5.9	0.3	0.8	0.5

Figure 4: Magnitude of errors (in kJ/mol) of selected CC methods against the ASCI benchmark for all the oxides.

Fig 4 shows individual errors predicted by relatively robust CC methods for all the oxides. CoO and NiO appear to be the most challenging species overall, although the rest can hardly be described as easy, especially at the CCSD(T) level. Even full CCSDTQ fails to bring error in NiO below 2 kJ/mol, and a $(\mathrm{P})_\Lambda$ correction proves necessary! This however is unlikely to be a consequence a partial triple bond resulting from oxygen lone pair donation to Ni, as Ni has a $3d^8 4s^2$ configuration and thus is unlikely to accept any more than two electrons from O. Interestingly enough, CCSDT does worse than CCSD(T) for ScO and TiO, indicating CCSD(T) low errors are a consequence of cancellation of errors between an overestimation of the perturbative triples correction and complete neglect of quadruples. This general behavior is already long known to be responsible for part of the extraordinary accuracy of CCSD(T) for main group chemistry.[64] Similar error cancellations might also be responsible for the relatively better performance of CCSD(T) vs CCSD(2) for a number of species.

5.3.2 Sulphides

Fig 5 shows that the sulphides act similar to oxides overall, even though the magnitude of the errors is smaller (possibly due to the weakness of the M—S bond relative to the M—O bond). CCSD(2) consistently underperforms relative to CCSD(T), likely due to cancellation



	CCSD	CCSD(T)	CCSD(2)	CCSDT	CCSDT(Q)Λ	CCSDTQ	CCSDTQ(P)Λ
Sc	44.9	2.5	9.7	4.8	1.1	0.8	0.1
Ti	56.4	10.1	15.4	3.7	0.8	2.2	0.7
V	65.3	17.8	24.2				
Cr	58.0	7.9	19.2	5.1	0.2	0.1	0.3
Mn	27.7	5.0	7.8	1.2	0.4	0.1	0.3
Fe	52.0	15.9	20.1	5.1	0.4	0.1	0.0
Co	57.6	16.1	21.7	5.1	1.1	0.2	0.6
Ni	65.1	11.7	23.2				
Cu	14.3	5.6	1.2	0.4	1.5	0.1	0.8
Zn	20.5	3.1	5.1	2.6	0.5	0.7	0.7

Figure 5: Magnitude of errors (in kJ/mol) of selected CC methods against the ASCI benchmark for all the sulphides.

of errors in the former. CCSDT itself predicts errors above chemical accuracy for a number of compounds, but appears fairly accurate for MnS and CuS. Full CCSDTQ is able to attain sub 1 kJ/mol error for all species with available data, other than TiS. Consistent sub kJ/mol accuracy can however only be obtained with the extremely expensive $(P)_\Lambda$ correction.

5.4 Polyvalent Ligands

The carbides and nitrides can be expected to have much more complex electronic structure relative to the preceding species, due to the possibility of metal-ligand triple bonds. Table 5 reveals that CCSD is catastrophically bad (RMSE 72-73 kJ/mol), due to systematic underestimation of bond correlation energy by 60-65 kJ/mol. CCSD(T) fares a lot better (RMSE of 11 kJ/mol for nitrides and 14 kJ/mol for carbides) but this is solely due to fortuitous error cancellation,[64] since full CCSDT winds up doing worse (RMSE 13 kJ/mol for nitrides and 21 kJ/mol for carbides, mostly on account of systematically missing 11-14 kJ/mol correlation energy on average)! It therefore appears that the perturbative (T) correction is a significant overestimate (possibly due to pseudo-divergent behavior of the perturbative contributions) that partially compensates for missing higher order terms, resulting in CCSD(T) not being as bad as one might expect. The poor performance of CCSDT however dooms the more robust $(T)_\Lambda$ correction, as well as all the iterative approximations to full triples.

Table 5: Root Mean Squared Error (RMSE) and Mean Error (ME) in BDEs predicted by CC methods for polyvalent ligands, against the ASCI benchmark (in kJ/mol). The CC methods have been listed roughly in order of their complexity. CrC, CoC, NiC, FeN and NiN were not included as all data was not available.

Method	Carbide		Nitride	
	RMSE	ME	RMSE	ME
CCSD	72.0	-60.3	72.8	-64.6
CCSD(T)	14.2	-11.9	10.8	-9.0
CCSD(T)$_\Lambda$	32.0	-21.2	17.3	-12.3
CCSD(2)	11.9	-10.4	9.9	-9.0
CCSDT-1a	11.2	-0.3	21.6	9.0
CCSDT-1b	33.2	-13.7	20.9	0.2
CCSDT-3	24.1	-13.1	12.8	-7.3
CCSDT	21.2	-13.9	13.4	-10.9
CCSDT(Q)/A	13.7	-0.8	13.1	-9.6
CCSDT(Q)/B	11.0	3.2	7.1	-1.0
CCSDT(Q)$_\Lambda$	1.7	-0.2	1.5	-1.2
CCSDTQ-1a	6.4	-3.6	7.6	-6.1
CCSDTQ-1b	2.6	-0.8	6.1	-4.2
CCSDTQ-3	5.4	-4.3	6.8	-5.6
CCSDTQ	5.7	-3.3	4.0	-3.0
CCSDTQ(P)/A	5.2	0.1	2.1	-1.2
CCSDTQ(P)/B	6.4	-0.2	2.2	-1.0
CCSDTQ(P)$_\Lambda$	1.6	-1.0	1.0	-0.7

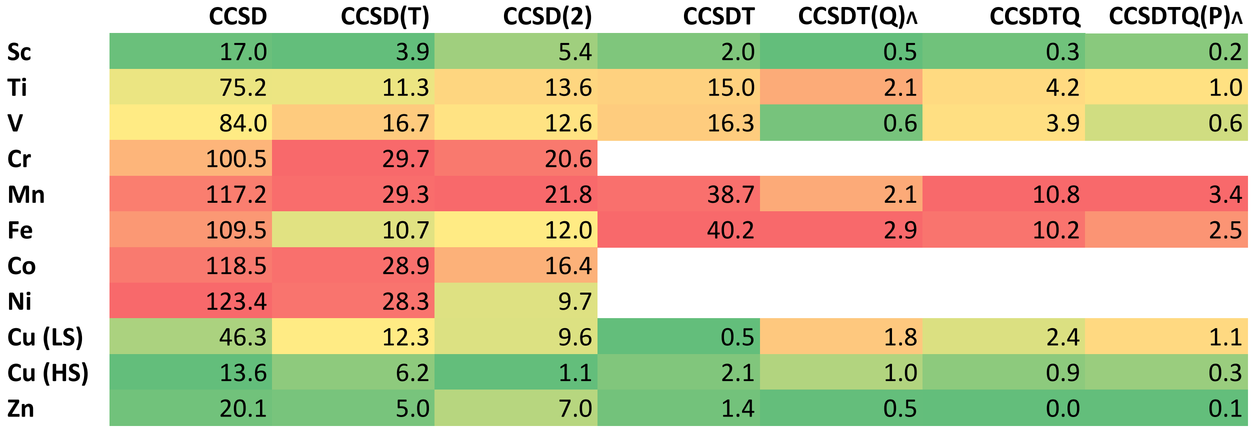
Very interestingly, CCSD(2) gives a much lower RMSE than CCSD(T)$_\Lambda$, even though the two methods had similar errors for all preceding cases. The higher accuracy of CCSD(2) is therefore likely a consequence of the perturbative quadruples terms contained in it. This, coupled with the less divergent behavior of CCSD(2), therefore suggests that it is getting the approximately correct answer for more rigorous reasons than CCSD(T), although it could also be a beneficiary of error cancellation. Overall, both CCSD(T) and CCSD(2) underestimate bond energies by roughly 10 kJ/mol due to missing correlation energy stemming from higher order connected terms.

Approximate quadruples corrections help reduce the error in CCSDT too, but this again likely originates from error cancellation, given that full CCSDTQ has higher RMSE than CCSDT(Q)$_\Lambda$ (and CCSDTQ-1b, for the case of carbides). The decent performance of the

(Q)/A and (Q)/B corrections is noteworthy as they seemed to worsen CCSDT performance for the more weakly correlated species considered earlier. The higher accuracy of CCSDT(Q)/B however owes a great deal to spurious overbinding relative to full CCSDTQ (as can be seen from mean errors in Table 5). Full CCSDTQ in fact has an RMSE of 4-6 kJ/mol (3 kJ/mol of which comes from systematic underbinding), which is far too large relative to the accepted chemical accuracy limit of 4 kJ/mol. Pentuples corrections prove necessary to lower the error further, with CCSDTQ$(P)_\Lambda$ having an RMSE of 1.6 kJ/mol for carbides and 1.1 kJ/mol for nitrides, most of it coming from systematically underestimating bond energy by 1 kJ/mol. Full CCSDTQP proved computationally intractable, and so it is not completely certain how much of CCSDTQ$(P)_\Lambda$'s great accuracy is a result of error cancellations. Interestingly, the (P)/A and (P)/B perturbation ansatze failed to improve upon CCSDTQ for the carbides, but were very accurate for the nitrides. This difference in behavior is quite difficult to understand, though it is possible that these corrections were acting in a more divergent manner for the more strongly correlated carbides (as evidenced by their low systematic error relative to the consistent underbinding behavior seen from nitrides). Overall however, it is quite clear that the metal-ligand triple bonds are extremely challenging for CC calculations.

5.4.1 Carbides

Fig 6 reveals that, while most of the carbides are quite challenging for CC methods, a few are relatively easily described. The performance of CCSD(T) for ScC, ZnC and high spin (quartet) CuC is not too poor, and the good performance of CCSDT for these species suggest that cancellation of errors had little role to play here. Low spin (doublet) CuC is also decently described at CCSDT and higher levels of theory. TiC and VC are considerably more challenging, but CCSDTQ can get errors down to chemical accuracy and the $(P)_\Lambda$ pentuples correction improves predictions further. MnC and FeC on the other hand are incredibly challenging, with CCSDTQ predicting errors around 10 kJ/mol, and even the

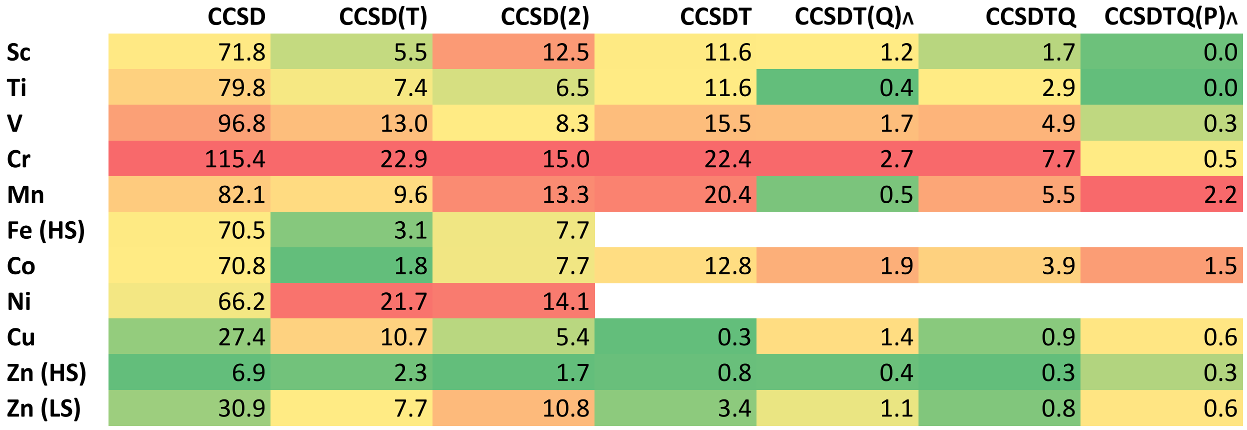


	CCSD	CCSD(T)	CCSD(2)	CCSDT	CCSDT(Q)Λ	CCSDTQ	CCSDTQ(P)Λ
Sc	17.0	3.9	5.4	2.0	0.5	0.3	0.2
Ti	75.2	11.3	13.6	15.0	2.1	4.2	1.0
V	84.0	16.7	12.6	16.3	0.6	3.9	0.6
Cr	100.5	29.7	20.6				
Mn	117.2	29.3	21.8	38.7	2.1	10.8	3.4
Fe	109.5	10.7	12.0	40.2	2.9	10.2	2.5
Co	118.5	28.9	16.4				
Ni	123.4	28.3	9.7				
Cu (LS)	46.3	12.3	9.6	0.5	1.8	2.4	1.1
Cu (HS)	13.6	6.2	1.1	2.1	1.0	0.9	0.3
Zn	20.1	5.0	7.0	1.4	0.5	0.0	0.1

Figure 6: Magnitude of errors (in kJ/mol) of selected CC methods against the ASCI benchmark for all the carbides.

$(P)_\Lambda$ correction results in 2-3 kJ/mol error.

5.4.2 Nitrides



	CCSD	CCSD(T)	CCSD(2)	CCSDT	CCSDT(Q)Λ	CCSDTQ	CCSDTQ(P)Λ
Sc	71.8	5.5	12.5	11.6	1.2	1.7	0.0
Ti	79.8	7.4	6.5	11.6	0.4	2.9	0.0
V	96.8	13.0	8.3	15.5	1.7	4.9	0.3
Cr	115.4	22.9	15.0	22.4	2.7	7.7	0.5
Mn	82.1	9.6	13.3	20.4	0.5	5.5	2.2
Fe (HS)	70.5	3.1	7.7				
Co	70.8	1.8	7.7	12.8	1.9	3.9	1.5
Ni	66.2	21.7	14.1				
Cu	27.4	10.7	5.4	0.3	1.4	0.9	0.6
Zn (HS)	6.9	2.3	1.7	0.8	0.4	0.3	0.3
Zn (LS)	30.9	7.7	10.8	3.4	1.1	0.8	0.6

Figure 7: Magnitude of errors (in kJ/mol) of selected CC methods against the ASCI benchmark for all the nitrides.

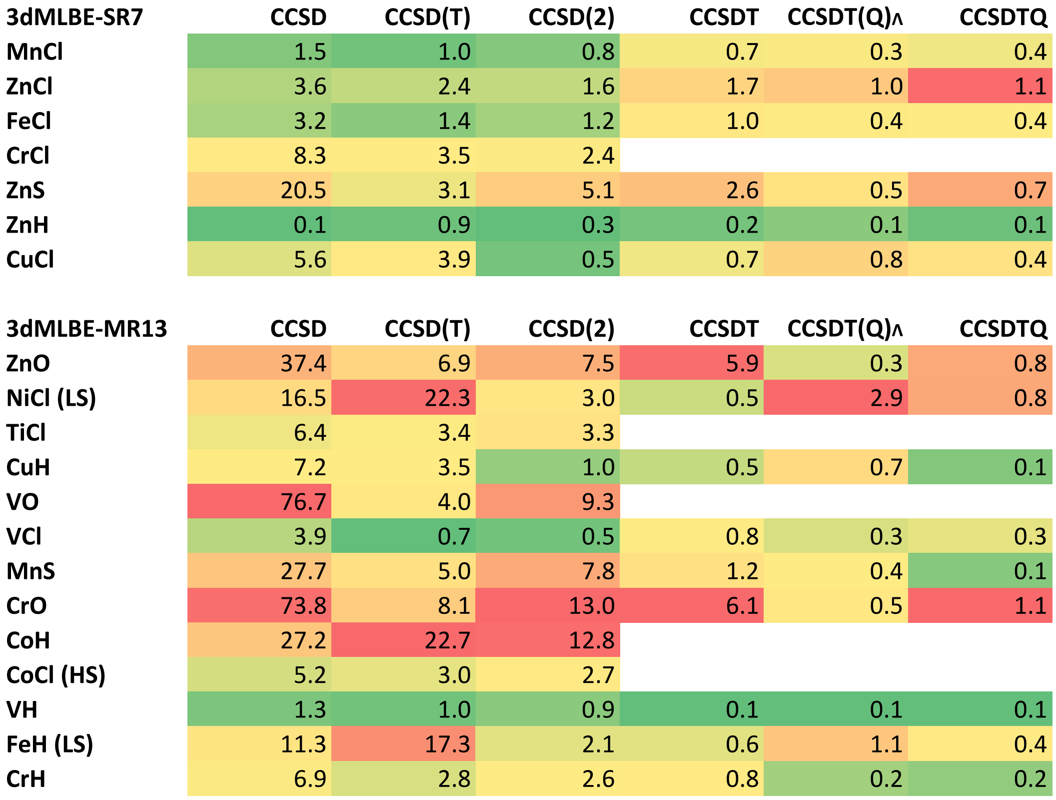
The MN compounds appear to have quite complex electronic structure overall, similar to the MC compounds. Only high spin (quartet) ZnN can be considered to be 'easy' from a CCSD(T) perspective, although the good performance of CCSDT for low spin (doublet) ZnN and CuN suggest that these compounds are not too challenging. At the other extreme, VN, CrN, MnN and CoN appear to be quite challenging as even CCSDTQ predicts around 4 kJ/mol error. The $(P)_\Lambda$ correction however gives good performance for VN and CrN. The less robust (P)/A and (P)/B corrections yield similar behavior as well, and this consistency between the three recipes suggest that the good behavior is not all due to error cancellation. MnN and CoN however appear to be challenging even with pentuples corrections, and therefore should probably be excluded from any DFT development datasets.

5.5 Behavior for the 3dMLBE20 dataset

The 3dMLBE20 dataset proposed by Xu et.al.[11] is a subset of twenty molecules out of our full dataset, which has been studied more extensively than others with wave function theory,[11,36,37] albeit mostly by comparison to experiment. The original study by Xu et.al.[11] claimed that CC theories were in general suboptimal for this subset, but this has been subsequently contested.[36,37] Ref 37 in fact presented benchmark BDEs using the HEAT protocol,[29] which uses CCSDTQ as the highest level of theory. We therefore decided to specifically examine how well various methods fare for this specific subset.

We begin by noting that the dataset is 45% chlorides, 30% hydrides, 15% oxides and 10% sulphides. The absence of any nitrides or carbides immediately gives CC theory a considerable advantage relative to our full dataset. However, a number of challenging cases like CoCl are present. Xu et.al. had attempted to classify the molecules into a 7 member single reference (SR) subset and a 13 member multireference subset, based on the $\hat{T}_1$ MR metric,[65] which is depicted in Fig 8, along with errors predicted by some representative CC methods. It is immediately apparent that the SR labeled species have BDE's that can be estimated by CCSD(T) to chemical accuracy.

The MR labeled cases give more interesting behavior, with CuH, VCl, VH and CrH showing essentially the same behavior as the supposedly SR species, and TiCl potentially exhibiting similar behavior (but not VO, due to the error increasing on moving to CCSD(2) from CCSD(T), suggesting cancellation of errors in the latter). CCSDT appears adequate for NiCl, MnS and quartet FeH, for which CCSD(T) predicts larger errors than the chemical accuracy limit. ZnO and CrO appear to require quadruples corrections to reach chemical



3dMLBE-SR7	CCSD	CCSD(T)	CCSD(2)	CCSDT	CCSDT(Q)Λ	CCSDTQ
MnCl	1.5	1.0	0.8	0.7	0.3	0.4
ZnCl	3.6	2.4	1.6	1.7	1.0	1.1
FeCl	3.2	1.4	1.2	1.0	0.4	0.4
CrCl	8.3	3.5	2.4			
ZnS	20.5	3.1	5.1	2.6	0.5	0.7
ZnH	0.1	0.9	0.3	0.2	0.1	0.1
CuCl	5.6	3.9	0.5	0.7	0.8	0.4

3dMLBE-MR13	CCSD	CCSD(T)	CCSD(2)	CCSDT	CCSDT(Q)Λ	CCSDTQ
ZnO	37.4	6.9	7.5	5.9	0.3	0.8
NiCl (LS)	16.5	22.3	3.0	0.5	2.9	0.8
TiCl	6.4	3.4	3.3			
CuH	7.2	3.5	1.0	0.5	0.7	0.1
VO	76.7	4.0	9.3			
VCl	3.9	0.7	0.5	0.8	0.3	0.3
MnS	27.7	5.0	7.8	1.2	0.4	0.1
CrO	73.8	8.1	13.0	6.1	0.5	1.1
CoH	27.2	22.7	12.8			
CoCl (HS)	5.2	3.0	2.7			
VH	1.3	1.0	0.9	0.1	0.1	0.1
FeH (LS)	11.3	17.3	2.1	0.6	1.1	0.4
CrH	6.9	2.8	2.6	0.8	0.2	0.2

Figure 8: The 3dMLBE20 dataset (with SR and MR classification), along with absolute errors (in kJ/mol) predicted by various CC approaches against our benchmark.

accuracy, while the incomplete information for CoH also pegs it as a difficult species. CoCl is a harder case, as we have energies for the high spin quintet state and were not able to converge the triplet ground state to sufficient accuracy. The HS CoCl state looks relatively easy, but the LS state appears to be much more difficult. We therefore think that CCSD(T) can get to chemical accuracy for 5 out of the 13 supposed MR species and all the SR labeled species. Three of the MR subset require full triples corrections, two others need some quadruples corrections and the remaining three appear to be challenging for CC theory, but insufficient data is available to determine what is the minimum level of CC theory needed to accurately find their BDE. However, the overall accuracy of CCSDTQ indicates that the values presented in Ref 37 are quite accurate, though around 1 kJ/mol error is possible for CrO, ZnCl, ZnO, NiCl and ZnS, and it is difficult to be fully certain about VO, CoH and CoCl. The uncertainties in typical experimental results however suggest that Ref 37 values should be well within those error bars.

5.6 Estimate of multireference character

Table 6: Effective number of unpaired electrons arising from broken electron pairs for species considered, as estimated from the ASCI variational wave function. Multiple spin states were considered for some species, resulting in values for high spin (HS) and low spin (LS) cases. IC stands for insufficiently converged (i.e. high error) cases that have not been reported.

	Sc	**Ti**	**V**	**Cr**	**Mn**	**Fe**	**Co**	**Ni**	**Cu**	**Zn**
H	0.67	1.16	0.35	0.39	0.27	0.29 (hextet) 1.12 (quartet)	1.06	0.63	0.40	0.31
F	0.71	0.40	0.41	0.44	0.36	0.38	IC	0.38 (quartet) 0.46 (doublet)	0.44	0.38
Cl	0.81	0.45	0.45	0.50	0.42	0.44	0.43 (quintet) IC (triplet)	0.43 (quartet) 0.62 (doublet)	0.47	0.43
O	0.66	0.73	0.80	0.96	0.90	0.89	1.08	0.95	0.60	0.89
S	0.81	0.88	0.98	1.27	0.82	1.00	0.97	0.88	0.57	0.69
C	0.78	1.01	1.17	1.57	1.79	1.51	1.52	1.46	0.54 (quartet) 0.89 (doublet)	0.90
N	1.04	0.91	1.03	1.40	1.44	1.02 (hextet) IC (quartet)	1.05	0.91	0.70	0.52 (quartet) 1.03 (doublet)

It is also interesting to consider the multireference character of the studied species. There are multiple measures for this in literature like the aforementioned T_1 metric,[65] but they are known to yield conflicting results.[11] We have therefore chosen to use the following simple approach to determine the multireference character of the studied systems, using an estimate for the number of unpaired electrons resulting from correlation driven uncoupling of electron pairs, that are strictly paired in a Slater determinant with restricted orbitals.

1. Construct alpha and beta density matrices $\mathbf{P}_{\boldsymbol{\alpha}}$ and $\mathbf{P}_{\boldsymbol{\beta}}$ from the ASCI variational wave function.
2. Diagonalize the spinless density matrix $\mathbf{P} = \mathbf{P}_{\boldsymbol{\alpha}} + \mathbf{P}_{\boldsymbol{\beta}}$ to obtain spinless natural orbitals.
3. Rotate $\mathbf{P}_{\boldsymbol{\alpha}}$ and $\mathbf{P}_{\boldsymbol{\beta}}$ to this natural orbital basis. The resulting diagonal elements $(\mathbf{P}_{\boldsymbol{\alpha}})_{ii}$ and $(\mathbf{P}_{\boldsymbol{\beta}})_{ii}$ would then be α and β natural orbital populations $n_{\boldsymbol{\alpha},i}$ and $n_{\boldsymbol{\beta},i}$.

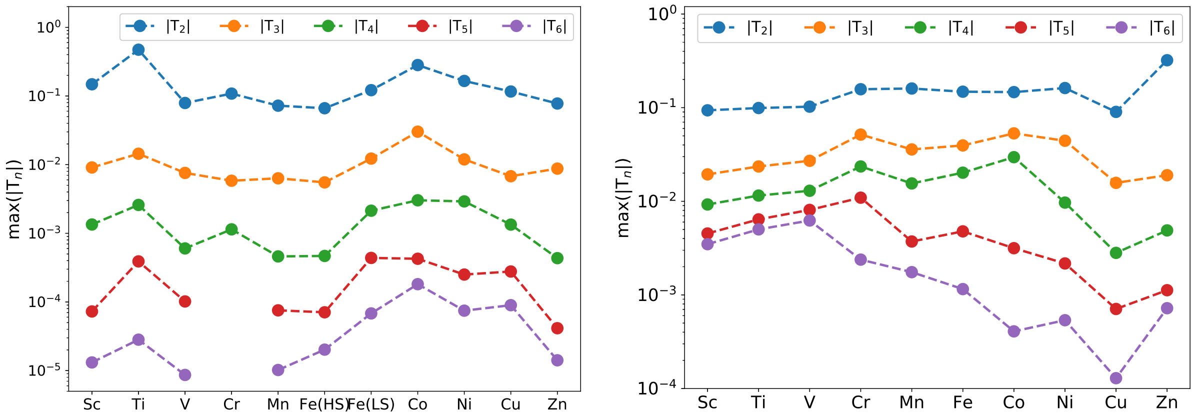
4. The estimated number of correlation driven unpaired electrons is then:
   $\sum_i \left(\min\left(n_{\boldsymbol{\alpha},i}, 1-n_{\boldsymbol{\alpha},i}\right)+\min\left(n_{\boldsymbol{\beta},i}, 1-n_{\boldsymbol{\beta},i}\right)\right)$

The resulting metric is consequently a generalization of earlier work for closed-shell species[66] and essentially involves summation of the difference in spin populations of the exact wave function and the ROHF type determinant in the natural orbital basis. This metric correctly predicts roughly 2 unpaired electrons for single bond dissociation (specifically, 2.04 for stretched ethane), roughly 4 for double bonds (specifically, 4.21 for stretched ethene) and 6 for triple bonds (specifically, 6.14 for stretched N_2). It also predicts approximately 1.94 unpaired electrons for singlet para-benzyne (which is functionally a biradical) and a value of 0.97 for O_3-a classic example of a main group species with strong correlation.[28] Values of this quantity for all the studied species is given in Table 6.

The values in Table 6 reveal that the species studied are not very multireference on the whole, typically having less than 1 correlation driven unpaired electrons. None appear to represent a situation where a full electron pair is broken because of correlation (i.e. a metric > 2). The most multireference species (> 1) are TiH, quartet FeH, CoH, CoO, CrS and FeS out of the mono and divalent ligand containing species (though CrO, NiO, VS and CoS are quite borderline). Many carbides and nitrides however have more than one correlation driven unpaired electron (with only a few like CuC being exceptions). Species with large values of this metric are amongst the most challenging species in the entire dataset (like CoH, NiO or MnC), although it would be unwise to solely draw conclusions about relative difficulty of species based solely on this factor, since tough species like LS NiCl can have smaller values like 0.62 (suggesting strong correlation sans significant multireference character).

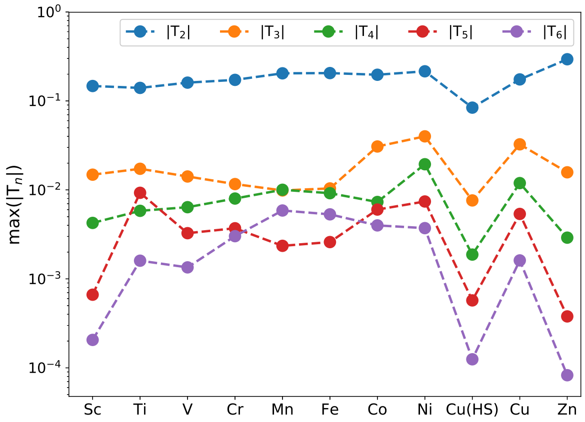
5.7 Cluster Decompositions

It is also instructive to consider what cluster amplitudes are present in the true many body wave function, in order to determine what is the best a given level of CC theory might perform. Information about this can be obtained by decomposing the ASCI variational



(a) MH

(b) MO



(c) MC

Figure 9: Largest T amplitudes for MX species, using ASCI wave functions.

wave function (which in principle contains nearly all the important degrees of freedom, and thereby nearly all strong correlation) into cluster amplitudes via the CLUSTERDEC program[67] (which has been described in Ref 68). The number of potential cluster amplitudes however is very large, and we therefore consider only the largest cluster amplitude of the entire decomposition, since strong correlation in such systems is likely to arise on account of interactions between very specific electrons, which would indicate outsized impact of specific cluster amplitudes. While most of the cluster amplitudes are likely a fair bit more approximate than the energies (since the PT2 and extrapolation induced gains in accuracy are not present), the largest ones are expected to be amongst the most converged due to the ASCI selection rule being effective at identifying important degrees of freedom.

Figs 9a, 9b and 9c present the largest cluster amplitudes generated by cluster decompositions of the ASCI variational wave functions (in the approximate natural orbital basis) for hydrides, oxides and carbides, as representative cases for the three ligand classes. There is no point in considering $\hat{T}_1$ since $e^{\hat{T}_1}$ serves to only perform orbital rotations.[69] Fig 9a shows that the magnitude of the largest cluster amplitude monotonically decays with increasing the order of excitation for the hydrides, and all excitation orders are well separated from each other. These species therefore could be considered to be well behaved, as there are no anamolous cases where $|T_4|$ jumps above $|T_2|$ in magnitude or the like. The largest $|T_5|$ and $|T_6|$ are below 10^{-3} suggesting that CCSDTQ should be adequate, which is indeed reflected by the energy errors in Fig 1. Similar behavior is seen for fluorides and chlorides. A monotonic convergence of exact CC methods with increasing number of allowed excitations is therefore expected (which is what we observe in energy errors till CCSDTQ).

The oxides (as shown in Fig 9b) exhibit somewhat similar behavior. However, several species have large T_5 and T_6 values, and the gap between the largest T_4, T_5 and T_6 values is often quite small, indicating that post CCSDTQ corrections might be needed in cases to have very high accuracy, as confirmed by Fig 4. Sulphides have very similar behavior as well. On the other hand, Fig 6 reveals significantly different behavior for the carbides. The gaps between the largest T amplitudes is often small, and the largest T_5 amplitude jumps above the largest T_4 for TiC, while the largest T_6 is larger than the largest T_5 for MnC and FeC. This suggests that such species would be quite challenging with CCSDTQ or even (P) corrections. Indeed, MnC and FeC have errors in excess of 2 kJ/mol even with CCSDTQ(P)$_\Lambda$. The nitrides also exhibit similar behavior, showing the challenges CC theory faces in describing interactions between metals and multivalent ligands.

We would however caution against overanalyzing the connection between the largest cluster amplitudes and the energy errors discussed earlier, as the latter involves collective interplay between individual amplitudes and Hamiltonian matrix elements, which can be glossed over by global metrics over all amplitudes. Indeed, while these plots can at times reveal periodic trends (relative 'easiness' of HS CuC indicated by low T magnitudes in Fig 9c are confirmed by energy errors in Fig 6, as an obvious case), they are often not very effective is distinguishing between species. For instance, NiO appears to have much smaller T_5 and T_6 values than ScO or TiO, and yet has a larger CCSDTQ error than the latter two. Use of different global metrics like the $||L||_2$ norm yielded a similar picture as well, indicating that further insight may not be readily obtained by other global metrics over amplitudes alone.

6 Conclusions

In summary, we have estimated the BDEs of 74 spin states of 69 3d transition metal binary compounds (out of the 70 possible species with the H, F, Cl, O, S, C and N ligands) with an expected 3σ error under 1 kJ/mol relative to the exact, Born-Oppenheimer non-relativistic results in the def2-SVP basis. This high level of accuracy was possible by using the ASCI approach to generate compact SCI wave functions, whose energies were augmented with a PT2 correction and were subsequently extrapolated to the full CASCI limit of zero PT2 energy. The resulting finite basis BDEs were used to assess the performance of standard CC models for transition metal systems.

We find substantial differences in behavior between three groups of ligands. The monovalent hydride and halide ligand containing species are on the whole relatively well described with basic CC theories like CCSD(T), barring some noticeable exceptions. These exceptions include all the Co containing compounds, as well as quartet FeH, NiH , CuF, doublet NiF and doublet NiCl. The remainder however are described to essentially chemical accuracy with CCSD(T). This is expected behavior, as the metal-ligand single bond here is a two electron interaction, which main group chemistry experience suggests should be quite well described by triples corrections on top of CCSD. It is nonetheless worth noting that CCSD(T) mostly tends to systematically underbind due to missing correlation. More robust corrections like the $(T)_\Lambda$ and (2) corrections do even better than (T), which suggests that developing local approximations to them in the manner of DLPNO-CCSD(T) would be a worthy endeavor towards more accurate transition metal quantum chemistry. CCSDT is nearly always sufficient in getting to around 1 kJ/mol accuracy (or better) for all of these molecules (including the challenging cases for CCSD(T) where data is available), and a robust quadruples correction (such as $(Q)_\Lambda$ or full Q) appears to nearly always reduce residual errors considerably.

The situation is somewhat different for the divalent O and S ligands. Four electron metal ligand interactions are possible for these species, which would indicate that even CCSDT could find them challenging. This is indeed borne out by the BDEs, with CCSDT having errors of around 6 kJ/mol for oxides and 4 kJ/mol for sulphides. Approximate triples corrections like CCSD(T) consequently perform worse, with errors in the order of 10-16 kJ/mol stemming from systematically missing correlation energy. The CCSD(2) and CCSD(T)_Λ methods either do not significantly improve upon CCSD(T) or sometimes perform worse, likely on the account of their more robust perturbative models damping the corrections too much relative to the more minimal (and more susceptible to divergence) CCSD(T). Perturbative $(Q)_\Lambda$ corrections on top of CCSDT lower error considerably, and RMSE around 1 kJ/mol is achieved when full CCSDTQ is used, suggesting that the challenge was indeed modeling four electron correlations. A pentuples $(P)_\Lambda$ correction is necessary to attain sub 1 kJ/mol accuracy for such species, highlighting the challenges involved with developing a DFT functional development quality benchmark purely from CC approaches.

The C and N ligands create even more complex electronic structure on account of their ability to form metal ligand triple bonds, whose six electron interactions are potentially beyond even CCSDTQ to model. CCSDT itself fails quite badly for such species, systematically underestimating BDEs by 10-18 kJ/mol. Somewhat miraculously, CCSD(T) has *lower* RMSE and ME than CCSDT, which is likely a consequence of an overestimated (T) perturbative correction compensating for the effect of missing quadruples. CCSD(T)_Λ on the other hand, winds up failing badly on account of trying to mimic CCSDT to a greater extent. CCSD(2) in contrast has somewhat lower error than CCSD(T) despite a less divergent triples correction, on account of the presence of some quadruples contributions. CCSD(2) therefore is likely getting a "better" answer than CCSDT for more physically correct reasons, even though nearly all its error stems from systematic underbinding. Full CCSDTQ itself has errors in the range of 3-6 kJ/mol, which are only further reduced by incorportation of a pentuples $(\mathrm{P})_\Lambda$ correction that is realistically unaffordable for most species. These species therefore represent extraordinarily challenging cases that can only be optimally addressed by explicitly MR methods, and are unlikely to be converged (in the complete basis set limit) to sufficient accuracy for inclusion in a functional development benchmark dataset.

In addition, we find some surprisingly poor performers amongst computationally intensive methods. The iterative approximations to higher order CC methods do not appear to offer any advantages over robust perturbative corrections like (2) or the $(n)_\Lambda$ ones, despite their greater asymptotic scaling. More surprisingly, the (Q)/A and (Q)/B perturbative corrections appear to degrade the performance of CCSDT for most cases, indicating that it is not really worth it to carry out these calculations in general. The corresponding pentuples (P)/A and (P)/B corrections appear to often be similarly problematic based on available data, suggesting that they too should probably be avoided. Overall, we recommend using only Λ based perturbative corrections beyond the triples level for transition metal systems, over iterative approximations or other perturbative models.

In conclusion, we have shown how different CC methods fare for different forms of ligand metal interaction over a dataset of 3d transition metal binary compounds. Looking to the future, this information could be used to generate a highly accurate composite scheme that combines CC and SCI data to obtain dissociation energies at the complete basis set limit. Such data (especially for monovalent ligands like H, F, Cl, Br, OH, NH_2, CH_3) etc would be quite useful for incorporating transition metal information to existing DFT functional development datasets. Work along these directions is currently in progress.

Supplementary Information

Geometries (zip), extrapolated ASCI energies with estimated errors and CC BDEs (xlsx).

Acknowledgment

This research was supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

A Justification for basis set choice

The def2-SVP basis was used for nearly all calculations in this work, despite its relatively small size. The computed BDEs for instance cannot be compared to experiment at all, and some spin-state orderings are incorrect because of basis set incompleteness errors. However, all the CC approaches being assessed are wave function theory techniques that attempt to approximate FCI solutions for *any* finite basis, and it is therefore fair to compare methods in this manner. Furthermore, precedent from main group thermochemistry indicates that high order CC contributions quickly reach the basis set limit around the double zeta level,[28,70] enabling protocols like HEAT[29] or W4[28] to only use double zeta basis sets at the CCSDTQ level (or beyond). We do however recognize that transition metal systems could potentially exhibit slower basis set convergence for high order CC contributions.

On a practical note, we observe that def2-SVP is the largest basis for which full CCSDTQ calculations could be performed with the resources available to us, which nonetheless took a while. The significant increase in number of virtual functions on moving to a triple-zeta basis would result in roughly a roughly 64 times longer calculation, which was unviable for the whole dataset. We have encountered little difficulty in getting ASCI numbers for many of the dataset species to chemical accuracy with the def2-TZVPP basis, and much higher accuracy gains appear readily achievable with modest increase of variational space to 10 or

20 million determinants. This is largely a consequence of the extra basis functions mostly adding dynamical correlation, whose effects could be easily estimated via the PT2 correction.

B Issues with ROCC frozen core

There are two possible choices of orbitals derived for ROHF for use in CC calculations. The ‘standard‘ approach employs the ROHF optimized orbitals straightaway, while the ‘semicanonical‘ orbitals are obtained by diagonalizing the alpha and beta Fock matrices generated from the ROHF density to have formally unrestricted orbitals. Furthermore, it is possible to neglect the core orbitals during the semi-canonicalization process. There are therefore often small differences in ROCC values between codes, as was noted in the original W4 paper, for instance.[28]

Semicanonical orbitals are nonetheless more widely used as they are the formally appropriate choice for perturbative corrections to CC. On the other hand, the Q-Chem ASCI code is set up to solely operate with standard orbitals, which dramatically simplifies the computation. So there is a slight difference in the CASCI orbital subspaces treated with ASCI+PT2 and the CC approaches using MRCC. The energetic impact of the difference can be well estimated by finding the difference between the CCSD BDEs from both sets of orbitals, as the resulting core effects are likely tiny. As anticipated, it is almost universally minuscule, with the typical values being $0.1 - 0.2$ kJ/mol, although there are a few cases where it goes as large as 1 kJ/mol. Attempting to ‘standardize’ MRCC energies with these numbers led to very similar errors, and thereby did not affect conclusions at all. This effect has consequently been neglected in the present study, though a full list of these differences can be found in the supporting information.

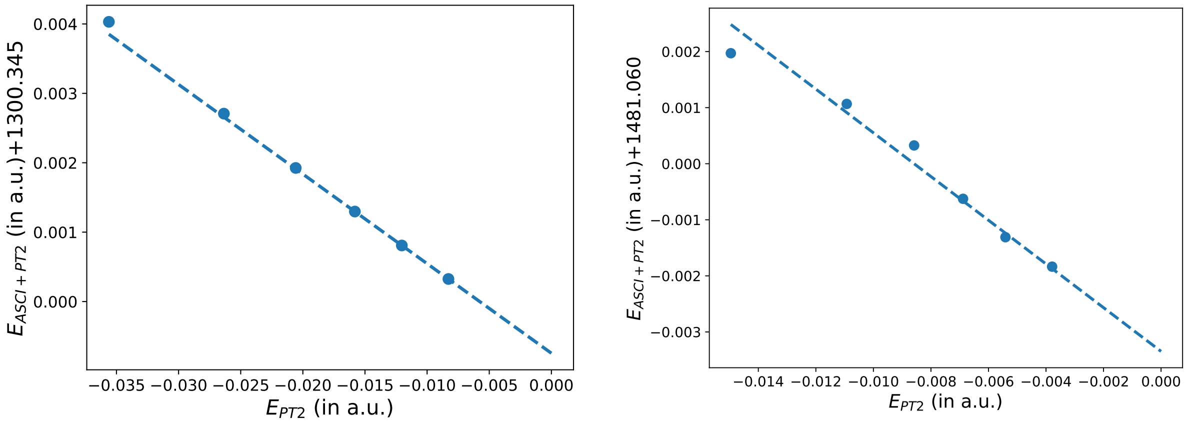
C Linear extrapolation and error estimation

The ASCI+PT2 energy ($E_{ASCI+PT2}$) was extrapolated against the PT2 correction (E_{PT2}) to the FCI limit of $E_{PT2} = 0$, in order to estimate the true FCI energy. An essentially linear behavior was seen in most cases (consistent with behavior seen for other SCI methods[49,50]), leading us to use a linear fit in order to minimize the possibility of overfitting. The standard uncertainty in the slope and intercept of linear fits are also quite well known (see Ref 71, for instance), providing a relatively reliable metric for the accuracy of our fit. Indeed, the standard deviation of the intercept for fitting data points $\{y_i\}$ against the variable $\{x_i\}$ is:

$$\sigma = \sqrt{(1-r^2)\left(\frac{\bar{x}^2}{\sigma_x^2}+1\right)\frac{\sigma_y^2}{n-2}} \tag{4}$$

where n is the number of data-points, $\bar{x}$ is the average of $\{x_i\}$, σ_x^2 and σ_y^2 are variances of $\{x_i\}$ and $\{y_i\}$ respectively, and r^2 is the square of the coefficient of correlation for the fit. Low intercept error σ can therefore be achieved by low $1-r^2$ (i.e. r^2 very close to 1) or low σ_y^2 (i.e. y almost flat and therefore has little variation). For our purposes, the latter condition indicates that $E_{ASCI+PT2}$ itself has nearly converged (has low variation).

The ligand atoms C,N,O,F,S and Cl had PT2 corrections below 10^{-8} a.u., and no extrapolation was therefore necessary. For all other systems, we performed linear regression with three points (typically corresponding to variational ASCI subspaces of 1 million, 2 million and 5 million determinants) to determine the intercept and the standard error σ associated with it. In some cases (most notably the V atom), the fit was not very linear at all, but $E_{ASCI+PT2}$ itself was very converged, leading to low σ_y^2 and therefore low error. For others, extrapolation recovered a substantial amount of energy (going upto 4 kJ/mol in cases), but very small $1-r^2$ led to low error predictions. An example for this has been depicted in Fig 10a. Certain species like CoF however had rather nonlinear behavior (as can be visually seen from Fig 10b) and insufficient convergence of $E_{ASCI+PT2}$, leading to large predicted error and subsequent exclusion from Table 2 and analysis.



(a) FeC

(b) quintet CoF

Figure 10: Examples of both excellent(left) and suboptimal (right) linear fits. The data has been fit only to the three points with smallest E_{PT2} by magnitude, but results from smaller variational subspaces have been added to show trends. The raw data has been supplied in Table 7.

Table 7: Raw data for fits in Fig 10 (in a.u.).

Variational Determinants	FeC		CoF (quintet)	
	E_{PT2}	$E_{ASCI+PT2}$	E_{PT2}	$E_{ASCI+PT2}$
100000	-0.035607	-1300.340970	-0.014955	-1481.058029
250000	-0.026355	-1300.342292	-0.010926	-1481.058933
500000	-0.020574	-1300.343075	-0.008585	-1481.059674
1000000	-0.015814	-1300.343703	-0.006883	-1481.060626
2000000	-0.012036	-1300.344192	-0.005410	-1481.061311
5000000	-0.008285	-1300.344675	-0.003787	-1481.061836
Fit		-1300.345744		-1481.063348

References

(1) Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133.

(2) Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Mol. Phys.* **2017**, *115*, 2315–2372.

(3) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Elec-*

tronic Structure Theory; Dover Publications, Inc.: Mineola, New York, 1996.

(4) Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291.

(5) Van Voorhis, T.; Head-Gordon, M. Benchmark variational coupled cluster doubles results. *J. Chem. Phys.* **2000**, *113*, 8873–8879.

(6) Hachmann, J.; Dorando, J. J.; Avilés, M.; Chan, G. K.-L. The radical character of the acenes: A density matrix renormalization group study. *J. Chem. Phys.* **2007**, *127*, 134309.

(7) Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184–32215.

(8) Hait, D.; Head-Gordon, M. How accurate is density functional theory at predicting dipole moments? An assessment using a new database of 200 benchmark values. *J. Chem. Theory Comput.* **2018**, *14*, 1969–1981.

(9) Hait, D.; Head-Gordon, M. How accurate are static polarizability predictions from density functional theory? An assessment over 132 species at equilibrium geometry. *Phys. Chem. Chem. Phys.* **2018**, *20*, 19800–19810.

(10) 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.

(11) Xu, X.; Zhang, W.; Tang, M.; Truhlar, D. G. Do Practical Standard Coupled Cluster Calculations Agree Better than Kohn–Sham Calculations with Currently Available Functionals When Compared to the Best Available Experimental Data for Dissociation

Energies of Bonds to 3 d Transition Metals? *J. Chem. Theory Comput.* **2015**, *11*, 2036–2052.

(12) Aoto, Y. A.; de Lima Batista, A. P.; Kohn, A.; de Oliveira-Filho, A. G. How to arrive at accurate benchmark values for transition metal compounds: Computation or experiment? *J. Chem. Theory Comput.* **2017**, *13*, 5291–5316.

(13) Shee, J.; Rudshteyn, B.; Arthur, E. J.; Zhang, S.; Reichman, D. R.; Friesner, R. A. On Achieving High Accuracy in Quantum Chemical Calculations of 3 d Transition Metal-Containing Systems: A Comparison of Auxiliary-Field Quantum Monte Carlo with Coupled Cluster, Density Functional Theory, and Experiment for Diatomic Molecules. *J. Chem. Theory Comput.* **2019**, *15*, 2346–2358.

(14) Vogiatzis, K. D.; Ma, D.; Olsen, J.; Gagliardi, L.; de Jong, W. A. Pushing configuration-interaction to the limit: Towards massively parallel MCSCF calculations. *J. Chem. Phys.* **2017**, *147*, 184111.

(15) Roos, B. O.; Taylor, P. R.; Si, P. E. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chem. Phys.* **1980**, *48*, 157–173.

(16) Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.* **1990**, *94*, 5483–5488.

(17) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* **1992**, *96*, 1218–1226.

(18) Olivares-Amaya, R.; Hu, W.; Nakatani, N.; Sharma, S.; Yang, J.; Chan, G. K.-L. The ab-initio density matrix renormalization group in practice. *J. Chem. Phys.* **2015**, *142*, 034102.

(19) Booth, G. H.; Thom, A. J.; Alavi, A. Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.* **2009**, *131*, 054106.

(20) Motta, M.; Zhang, S. Ab initio computations of molecular systems by the auxiliary-field quantum Monte Carlo method. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, e1364.

(21) Kim, J.; Baczewski, A. D.; Beaudet, T. D.; Benali, A.; Bennett, M. C.; Berrill, M. A.; Blunt, N. S.; Borda, E. J. L.; Casula, M.; Ceperley, D. M. et al. QMCPACK: an open source ab initio quantum Monte Carlo package for the electronic structure of atoms, molecules and solids. *J. Phys.: Condens. Matter* **2018**, *30*, 195901.

(22) Huron, B.; Malrieu, J.; Rancurel, P. Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions. *J. Chem. Phys.* **1973**, *58*, 5745–5759.

(23) Tubman, N. M.; Lee, J.; Takeshita, T. Y.; Head-Gordon, M.; Whaley, K. B. A deterministic alternative to the full configuration interaction quantum Monte Carlo method. *J. Chem. Phys.* **2016**, *145*, 044112.

(24) Holmes, A. A.; Tubman, N. M.; Umrigar, C. Heat-bath configuration interaction: An efficient selected configuration interaction algorithm inspired by heat-bath sampling. *J. Chem. Theory Comput.* **2016**, *12*, 3674–3680.

(25) Mahapatra, U. S.; Datta, B.; Mukherjee, D. A size-consistent state-specific multireference coupled cluster theory: Formal developments and molecular applications. *J. Chem. Phys.* **1999**, *110*, 6171–6188.

(26) Evangelista, F. A.; Simmonett, A. C.; Allen, W. D.; Schaefer III, H. F.; Gauss, J. Triple excitations in state-specific multireference coupled cluster theory: Application

of Mk-MRCCSDT and Mk-MRCCSDT-n methods to model systems. *J. Chem. Phys.* **2008**, *128*, 124104.

(27) Hanauer, M.; Köhn, A. Communication: Restoring full size extensivity in internally contracted multireference coupled cluster theory. *J. Chem. Phys.* **2012**, *137*, 131103.

(28) Karton, A.; Rabinovich, E.; Martin, J. M.; Ruscic, B. W4 theory for computational thermochemistry: In pursuit of confident sub-kJ/mol predictions. *J. Chem. Phys.* **2006**, *125*, 144108.

(29) Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vázquez, J.; Stanton, J. F. HEAT: High accuracy extrapolated ab initio thermochemistry. *J. Chem. Phys.* **2004**, *121*, 11599–11613.

(30) Tubman, N. M.; Freeman, C. D.; Levine, D. S.; Hait, D.; Head-Gordon, M.; Whaley, K. B. Modern Approaches to Exact Diagonalization and Selected Configuration Interaction with the Adaptive Sampling CI Method. *arXiv preprint arXiv:1807.00821* **2018**,

(31) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

(32) Jensen, K. P.; Roos, B. O.; Ryde, U. Performance of density functionals for first row transition metal systems. *J. Chem. Phys.* **2007**, *126*, 014103.

(33) Moltved, K. A.; Kepp, K. P. Chemical bond energies of 3d transition metals studied by density functional theory. *J. Chem. Theory Comput.* **2018**, *14*, 3479–3492.

(34) Harrison, J. F. Electronic structure of diatomic molecules composed of a first-row transition metal and main-group element (H- F). *Chem. Rev.* **2000**, *100*, 679–716.

(35) Kulik, H. J.; Marzari, N. Systematic study of first-row transition-metal diatomic molecules: A self-consistent DFT+ U approach. *J. Chem. Phys.* **2010**, *133*, 114103.

(36) Fang, Z.; Vasiliu, M.; Peterson, K. A.; Dixon, D. A. Prediction of bond dissociation energies/heats of formation for diatomic transition metal compounds: CCSD (T) works. *J. Chem. Theory Comput.* **2017**, *13*, 1057–1066.

(37) Cheng, L.; Gauss, J.; Ruscic, B.; Armentrout, P. B.; Stanton, J. F. Bond dissociation energies for diatomic molecules containing 3d transition metals: benchmark scalar-relativistic coupled-cluster calculations for 20 molecules. *J. Chem. Theory Comput.* **2017**, *13*, 1044–1056.

(38) Scemama, A.; Garniron, Y.; Caffarel, M.; Loos, P.-F. Deterministic construction of nodal surfaces within quantum Monte Carlo: the case of FeS. *J. Chem. Theory Comput.* **2018**, *14*, 1395–1402.

(39) Bender, C. F.; Davidson, E. R. Studies in configuration interaction: The first-row diatomic hydrides. *Phys. Rev.* **1969**, *183*, 23.

(40) Smith, J. E.; Mussard, B.; Holmes, A. A.; Sharma, S. Cheap and near exact CASSCF with large active spaces. *J. Chem. Theory Comput.* **2017**, *13*, 5468–5478.

(41) Levine, D. S.; Tubman, N. M.; Hait, D.; Head-Gordon, M.; Whaley, K. B. CASSCF with Extremely Large Active Spaces using the Adaptive Sampling Configuration Interaction Method. *In preparation* **2018**,

(42) Evangelisti, S.; Daudey, J.-P.; Malrieu, J.-P. Convergence of an improved CIPSI algorithm. *Chem. Phys.* **1983**, *75*, 91–102.

(43) Tubman, N. M.; Mejuto-Zaera, C.; Epstein, J. M.; Hait, D.; Levine, D. S.; Huggins, W.; Jiang, Z.; McClean, J. R.; Babbush, R.; Head-Gordon, M. et al. Postponing the orthog-

onality catastrophe: efficient state preparation for electronic structure simulations on quantum devices. *arXiv preprint arXiv:1809.05523* **2018**,

(44) Levine, D. S.; Tubman, N. M.; Hait, D.; Head-Gordon, M.; Whaley, K. B. Nuclear gradients with the Adaptive Sampling Configuration Interaction Self Consistent Field Method and appplication to bioinorganic chemistry. *In preparation* **2018**,

(45) Schriber, J.; Evangelista, F. Communication: An adaptive configuration interaction approach for strongly correlated electrons with tunable accuracy. *J. Chem. Phys.* **2016**, *144*, 161106.

(46) Epstein, P. S. The stark effect from the point of view of Schroedinger's quantum theory. *Phys. Rev.* **1926**, *28*, 695.

(47) Nesbet, R. Configuration interaction in orbital theories. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1955**, *230*, 312–321.

(48) Tubman, N. M.; Levine, D. S.; Hait, D.; Head-Gordon, M.; Whaley, K. B. An efficient deterministic perturbation theory for selected configuration interaction methods. *arXiv preprint arXiv:1808.02049* **2018**,

(49) Holmes, A. A.; Umrigar, C.; Sharma, S. Excited states using semistochastic heat-bath configuration interaction. *J. Chem. Phys.* **2017**, *147*, 164111.

(50) Loos, P.-F.; Scemama, A.; Blondel, A.; Garniron, Y.; Caffarel, M.; Jacquemin, D. A mountaineering strategy to excited states: Highly-Accurate reference energies and benchmarks. *J. Chem. Theory Comput.* **2018**,

(51) Noga, J.; Bartlett, R. J. The full CCSDT model for molecular electronic structure. *J. Chem. Phys.* **1987**, *86*, 7041–7050.

(52) Kucharski, S. A.; Bartlett, R. J. Recursive intermediate factorization and complete

computational linearization of the coupled-cluster single, double, triple, and quadruple excitation equations. *Theor. Chim. Acta* **1991**, *80*, 387–405.

(53) Kállay, M.; Gauss, J. Approximate treatment of higher excitations in coupled-cluster theory. II. Extension to general single-determinant reference functions and improved approaches for the canonical Hartree–Fock case. *J. Chem. Phys.* **2008**, *129*, 144101.

(54) Urban, M.; Noga, J.; Cole, S. J.; Bartlett, R. J. Towards a full CCSDT model for electron correlation. *J. Chem. Phys.* **1985**, *83*, 4041–4046.

(55) Kucharski, S. A.; Bartlett, R. J. Noniterative energy corrections through fifth-order to the coupled cluster singles and doubles method. *J. Chem. Phys.* **1998**, *108*, 5243–5254.

(56) Kállay, M.; Gauss, J. Approximate treatment of higher excitations in coupled-cluster theory. *J. Chem. Phys.* **2005**, *123*, 214105.

(57) Gwaltney, S. R.; Head-Gordon, M. A second-order perturbative correction to the coupled-cluster singles and doubles method: CCSD (2). *J. Chem. Phys.* **2001**, *115*, 2014–2021.

(58) Gwaltney, S. R.; Sherrill, C. D.; Head-Gordon, M.; Krylov, A. I. Second-order perturbation corrections to singles and doubles coupled-cluster methods: General theory and application to the valence optimized doubles model. *J. Chem. Phys.* **2000**, *113*, 3548–3560.

(59) Lee, Y. S.; Kucharski, S. A.; Bartlett, R. J. A coupled cluster approach with triple excitations. *J. Chem. Phys.* **1984**, *81*, 5906–5912.

(60) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X. et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184–215.

(61) Kállay, M.; Rolik, Z.; Csontos, J.; Ladjánszki, I.; Szegedy, L.; Ladóczki, B.; Samu, G.; Petrov, K.; Farkas, M.; Nagy, P. et al. MRCC, a quantum chemical program suite. *URL: http://www. mrcc. hu, accessed August 26th* **2016**,

(62) Rolik, Z.; Szegedy, L.; Ladjánszki, I.; Ladóczki, B.; Kállay, M. An efficient linear-scaling CCSD (T) method based on local natural orbitals. *J. Chem. Phys.* **2013**, *139*, 094105.

(63) Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. Natural triple excitations in local coupled cluster calculations with pair natural orbitals. *J. Chem. Phys.* **2013**, *139*, 134101.

(64) Ruden, T. A.; Helgaker, T.; Jørgensen, P.; Olsen, J. Coupled-cluster connected-quadruples corrections to atomization energies. *Chem. Phys. Lett.* **2003**, *371*, 62–67.

(65) Lee, T. J.; Rice, J. E.; Scuseria, G. E.; Schaefer, H. F. Theoretical investigations of molecules composed only of fluorine, oxygen and nitrogen: determination of the equilibrium structures of FOOF,(NO) 2 and FNNF and the transition state structure for FNNF cis-trans isomerization. *Theor. Chim. Acta* **1989**, *75*, 81–98.

(66) Head-Gordon, M. Characterizing unpaired electrons from the one-particle density matrix. *Chem. Phys. Lett.* **2003**, *372*, 508–511.

(67) Lehtola, S. ClusterDec. 2017; `https://github.com/susilehtola/clusterdec`.

(68) Lehtola, S.; Tubman, N. M.; Whaley, K. B.; Head-Gordon, M. Cluster decomposition of full configuration interaction wave functions: A tool for chemical interpretation of systems with strong correlation. *J. Chem. Phys.* **2017**, *147*, 154105.

(69) Thouless, D. Stability conditions and nuclear rotations in the Hartree-Fock theory. *Nuc. Phys.* **1960**, *21*, 225–232.

(70) Karton, A.; Taylor, P. R.; Martin, J. M. Basis set convergence of post-CCSD contributions to molecular atomization energies. *J. Chem. Phys.* **2007**, *127*, 064104.

(71) Bingham, N.; Fry, J. M. Linear Regression. In *Regression*; Springer, 2010.

Graphical TOC Entry