Sensitivity of coupled cluster electronic properties on the reference determinant: Can Kohn–Sham orbitals be more beneficial than Hartree–Fock orbitals?

Zsolt Benedek\textsuperscript{1,2,3} | Paula Timár\textsuperscript{1,2} | Tibor Szilvási\textsuperscript{3} | Gergely Barcza\textsuperscript{1,3,4}

\textsuperscript{1}Wigner Research Centre for Physics, Budapest, Hungary
\textsuperscript{2}Department of Inorganic and Analytical Chemistry, Budapest University of Technology and Economics, Budapest, Hungary
\textsuperscript{3}Department of Chemical and Biological Engineering, The University of Alabama, Tuscaloosa, Alabama, USA
\textsuperscript{4}Department of Physics of Complex Systems, ELTE Eötvös Loránd University, Budapest, Hungary

Correspondence
Gergely Barcza, Wigner Research Centre for Physics, H-1525 Budapest, Hungary.
Email: barcza.gergely@wigner.hu

Funding Information
Nemzeti Kutatási Fejlesztési és Innovációs Hivatal, Grant/Award Number: 135496

Abstract
Coupled cluster calculations are traditionally performed over Hartree–Fock reference orbitals (HF-CC methodology). However, in the literature it has been repeatedly raised whether the use of a Kohn–Sham reference (KS-CC methodology) might result in improved performance relative to HF-CC. In the present study, we re-examine the relation of HF-CC and KS-CC methods by comparing the results of widely applied truncated CC calculations (CCSD, CCSD(T), CCSDT) to the limit of full configuration interaction (FCI), which serves as an undebatable reference point of accuracy. Based on a series of CC calculations on diatoms and transition metal complexes, we demonstrate that no systematic improvement of coupled cluster electronic energies, densities and chemical reaction energies is expected when changing from HF to a KS reference. Nevertheless, fortuitous error cancellations might occasionally result in illusory improvement compared to HF-CC. Altogether, the application of KS-CC is not advantageous over HF-CC, but it is also not unreasonable as the choice of reference has negligible influence on the results at sufficiently high CC levels. In addition, KS-CC can be a particularly useful alternative if difficulties are encountered in HF or HF-CC convergence. It is also notable that KS-CC results are found to be practically independent of the chosen density functional, which implies that almost any KS-CC method can be used in place of HF-CC.

Keywords
benchmark, computational chemistry, coupled cluster theory, Kohn–Sham referenced coupled cluster theory, transition metal complexes

1 | INTRODUCTION

Coupled cluster (CC) calculations\textsuperscript{1} are considered to be one of the most reliable computational methods in quantum chemistry and thus often used as benchmarks.\textsuperscript{2} Their application, however, is strictly limited to single-reference systems denoted $\Phi$, where the exponential CC ansatz, $\Psi_{\text{CC}} = e^T \Phi$, parametrizes all possible configurations to accurately describe the electronic structure, even if the excitation operator $T$ is truncated.

From the theoretical point of view, the all-electron CC solution is orbitally invariant in the limit of full configuration interaction (FCI) but the truncated CC is sensitive to the choice of reference orbitals. At first sight, it could be even presumed that distinguished orbital sets are exceptionally advantageous to obtain FCI solution at computationally accessible low level of truncation. These expectations are put into perspective by the Thouless theorem\textsuperscript{3} which proves that determinants built from distinct orbital sets are related by the single excitation operator of the exponential CC ansatz. The theorem also implies that truncated CC results involving monoexcitations can be practically insensitive to the choice of reference orbitals as demonstrated in References 4,5. (Roughly speaking, the single excitation operator $e^{T_1}$ performs “orbital relaxation” and sets an optimal reference of $e^{T_1} \Phi$ for

\textit{Received:} 1 March 2022  \textit{Revised:} 28 June 2022  \textit{Accepted:} 30 June 2022

\textit{DOI:} 10.1002/jcc.26996
double and higher excitation terms, which is rather similar for virtually any initial $\Phi$ determinants.) However, this insensitivity is not to be confused with strict invariance as the impact of orbital choice on the truncated CC calculations does depend on the investigated problem. In fact, it is essentially a size-extensive effect yielding truncated CC energies for the distinct references, which might spread in a chemically relevant range of kcal/mol even for diatom, as shown in Section 3.2.

The construction of the optimal orbital set for CC calculations, which provides the possibly most accurate solution at the given truncation scheme compared to the FCI limit, is still an open question. The CC ansatz was originally formulated for the Hartree–Fock (HF) orbital\(^6\) derived Slater determinant $\Phi_{HF}$, which is the predominant reference to date. Upon the increase of the multi-reference character of the system, that is, upon the decrease of the weight of the reference in the ansatz, the reliability of the commonly applied truncated CC methods—such as CCSD, CCSDT or the gold standard CCSD(T)\(^7\)—becomes increasingly questionable.

To tackle the latter issue, alternative CC ansatzes were proposed which aim to provide a more appropriate reference Slater determinant. In Brueckner theory, for example, the CC calculation is coupled to an orbital rotation which sets all $T_1$ amplitudes to zero.\(^8\) This methodology maximizes the orbital relaxation effects but its implementation to the CC framework is not straightforward. In addition, it also does not necessarily improve, and in some cases even deteriorates accuracy compared to HF-CC.\(^9\)

Later, motivated by the success of Kohn–Sham Density Functional Theory (KS-DFT) in the description of electron correlation, some researchers suggested the replacement of canonical HF orbitals to KS-DFT derived auxiliary Kohn–Sham one-electron orbitals\(^10\) (whose physical interpretation is considered somewhat vague\(^11\)) in the reference determinant. In the last two decades, as reviewed in the following, a considerable number of KS-CC studies appeared in the literature\(^12\)–\(^21\) which argued that the use of KS reference determinants might lead to improved results due to the decrease of the multi-reference character in some aspects.

To the best of our knowledge, KS-CC calculations were proposed for the first time by Harvey and Aschi in 2003.\(^12\) In their work on the interconversion of iron carbonyl complexes (Fe(CO)\(_4\) + CO $\rightarrow$ Fe(CO)\(_5\)) in various spin states, they initially utilized the usual HF-CCSD(T) methodology and calculated the $T_1$ diagnostic for the studied systems as a test of multi-reference character in some aspects.

To the best of our knowledge, KS-CC calculations were proposed for the first time by Harvey and Aschi in 2003.\(^12\) In their work on the interconversion of iron carbonyl complexes (Fe(CO)\(_4\) + CO $\rightarrow$ Fe(CO)\(_5\)) in various spin states, they initially utilized the usual HF-CCSD(T) methodology and calculated the $T_1$ diagnostic for the studied systems as a test of multi-reference character in some aspects.

$$T_1 = \frac{\|t_1\|}{n_{corr}},$$

(1)

where, $\|t_1\|$ is the Euclidean norm of the vector of $t_1$ amplitudes, that is, the coefficients of singly excited determinants in the CC ansatz, and $n_{corr}$ denotes the number of correlated electrons.\(^22\) Empirically, $T_1 < 0.02$ is recommended for reliable CCSD(T) results but 0.04–0.06 was obtained for iron carbonyls. To gain more reliable energies at the same computational cost, the authors also performed CCSD(T) calculations using BP86 Kohn–Sham orbitals,\(^23,24\) which resulted in improved, but borderline (0.02–0.025) $T_1$ diagnostics. In addition, considerable deviations of up to 4 kcal/mol were observed between the respective HF-CCSD(T) and KS-CCSD(T) net reaction energies and the authors suggested the KS reference results to be more accurate. Unfortunately, the latter hypothesis could not be tested against experimental energies which are only approximately determined.

Some subsequent studies put the emphasis on the equivalence of HF-CC and KS-CC approaches, as much lower deviations were observed for other molecular systems. In 2005, the relative stability of the low-lying electronic excited states of NiCH\(_2\)\(^+\) was studied by Bartlett et al.\(^25\) at HF-CCSD and KS-CCSD levels. In contrast to the case of iron carbonyls, the reference orbital related deviation was found to be only 0–1.5 kcal/mol in the investigated cases, which led the authors to the conclusion that the choice of the initial orbital set is non-critical. A few years later, the Harvey group applied KS-CCSD(T) methodology for Heme model iron complexes, in the case of which HF-CCSD(T) calculations turned out to be practically unfeasible due to bad convergence behavior\(^26\) whereas the use of B3LYP Kohn–Sham orbitals\(^26,27\) eliminated the convergence issues. Furthermore, in the few cases where both KS and HF referenced results are available, it was observed that KS shows decreased singles and doubles CC amplitudes; the electronic energy, however, only negligibly (<1 kcal/mol) altered compared to HF counterpart.

In 2019, as their most recent publication\(^20\) related to the topic, Harvey et al. extensively studied the effect of different settings in CCSD(T) calculations, including the choice of the reference orbital type. They selected a C-H activation reaction on a non-heme type iron complex, that is, $[\text{Fe(NH}_3\text{)}_2\text{O}]^{2+} + \text{CH}_4 \rightarrow [\text{Fe(NH}_3\text{)}_3\text{O-H-CH}_3]^2$ in various spin states, as model system and found significant, up to 4 kcal/mol deviations between HF-CCSD(T) and KS-CCSD(T) energetics. Interestingly, the choice of the KS functional (B3LYP, BP86, TPSS,\(^28\) and M06,\(^29,30\) were tested) appeared to be practically insignificant, as the different KS-CCSD(T) energies varied within a mere 1 kcal/mol. Again, the authors suggested the KS reference to be more probably reliable without external benchmarking.

In the meantime, relying on the above works, Shaik\(^35\) and Pierloot\(^16\) also published B3LYP referenced coupled cluster studies on various iron complexes. These papers already assumed that the KS-CCSD(T) approach is equally reliable to HF-CCSD(T), thus, the traditional HF referenced calculations were not even performed.

Moving beyond iron group complexes, Dixon et al.\(^17\) investigated the compounds of numerous $d$ and $f$ transition metals using both HF-CCSD(T) and KS-CCSD(T). At first, the group explored the energy profile of water addition to metal dioxides, that is, $\text{MO}_2 + \text{H}_2\text{O} \rightarrow \text{M(OH)}_2\text{O}$ and found that the change of reference orbitals (Hartree–Fock or PW91 Kohn–Sham\(^34\)) results in deviations of up to 5 kcal/mol in the energy level of intermediates and transition states. Since the $T_1$ diagnostic was considerably smaller with KS reference, the ambiguity between the methods was interpreted as the improvement of CCSD(T) energetics by the inclusion of KS orbitals. Unfortunately, experimental data cannot be used for the assessment of accuracy in this case, as they are only available for metal oxide surfaces, rather
than for separate dioxide molecules. In their subsequent works,18,19 novel computational protocols were developed for the prediction of the heat of formation of transition metal hydrides, oxides, sulfides and chlorides, which treats the electronic energy at CCSD(T) level of theory. Again, it was observed that the use of PW91 or B3LYP orbitals instead of the usual HF orbitals decreases the T1 diagnostic and also alters the energy (or the derived enthalpy) by up to several kcal/mol. In this case, although experimental dissociation energies and enthalpies are abundantly available in the literature, most of these come with error bars exceeding the orbital choice related deviation. The Dixon group still managed to find one example where the B3LYP-KS reference appears to be unequivocally superior to HF, that is, the enthalpy of the reaction of UCl6 + 3F2 → UF6 + 3Cl2, experimentally measured to be 278.0 ± 1.2 kcal/mol at room temperature, was predicted to be 285.9 kcal/mol and 277.8 kcal/mol by HF-CCSD(T) and KS-CCSD(T) based protocols, respectively.18 Even though KS-CCSD(T) clearly produces better numerical agreement, assessing the quality of the computational models for the single benchmark is still challenging. In fact, not only an incidental cancellation of numerical errors cannot be ruled out, but thermochemical correction factors of the temperature-dependent enthalpy might also be a source of error.

Altogether, in the above pioneering KS-CC papers, the suspected superior performance of the KS based approach is typically argued on the basis of an overlap criterion which poses the references with more favorable T1 diagnostics as more appropriate. This is however a questionable argument as the usefulness of T1 as a measure of the quality of CC results is disputed: the high degree of orbital relaxation indicated by large T1 amplitudes1,5 does not necessarily imply poor CC electronic energies,22,23 the accuracy of which is of utmost importance in high-level quantum chemistry. Note that observing the other T1 = 0 extremity of the diagnostic, as in case of Brueckner orbitals, also does not guarantee improved accuracy.9

Apart from the overlap criterion, the quality of reference orbitals can be evaluated from numerous other aspects. The Head-Gordon group, for example, compared the performance of HF-CC and KS-CC calculations based on the reproduction of experimental vibrational frequencies.13,21 On a database of diatomic molecules, KS-CCSD(T) was found to outperform HF-CCSD(T) in accuracy by nearly a factor of 5 regardless of the density functional, among which BLYP,24,26 B97M-rV,26 B97,35 oB97X-V,56 and oB97M-V37 were tested. This result was attributed to the smooth change of KS orbitals with nuclear displacement contrasted to the rapid change of HF orbitals.

Importantly, although HF-CCSD(T) is less efficient regarding the frequency criterion, it does not necessarily mean that HF based CC produces less accurate electronic energies at the optimal ground state nuclear configuration, which is of primary interest in typical coupled cluster studies. Namely, the comparison of two vastly different regions and minima on the potential energy surface (which is required for electronic energies and chemical reaction energies) is a significantly more complex task than the determination of the local structure of the surface (which is required for vibrational analysis).

Regarding the applications of the CC methodology in quantum chemistry, the most relevant criteria would be (i) the extent and (ii) the consistency of error in electronic energy, which directly determine the reliability of CC derived energetics of chemical reactions. However, in order to validate the relative performance of truncated HF-CC and KS-CC approaches with unquestionable mathematical rigor, high-level numerical reference quantities are needed. As the truncated CC methods approximate the FCI wave function, the upper limit of their accuracy is the FCI energy for the applied basis set which is invariant to the chosen reference orbitals, see Appendix A for details. Accordingly, the main goal of the present work is to compare the truncated KS-CC and HF-CC results to calculations of essentially FCI quality benchmark results, which—in contrast to ambiguous wave-function diagnostics or experimental data—allows an unequivocal evaluation of accuracy. As such an analysis is unfeasible for transition metal complexes using today’s classical computers, we selected diatomic systems of both small and large multi-reference character for this purpose. Nevertheless, we will demonstrate that it is possible to generalize the results to experimentally more relevant larger molecules.

The rest of the paper is structured as follows. In Section 2, the computational details are given and we present the selection of diatomic model systems in Section 3.1. We discuss the effect of orbital choice on the difference between FCI and truncated CC energies in Section 3.2. As an additional evaluation of the quality of the truncated CC results, we also investigate the orbital dependence of similarity between FCI and CC electron densities in Section 3.3. Furthermore, to compare the behavior of HF-CC and KS-CC from the practical points of view, we examine the reference dependence of the wave-function diagnostics in Section 3.4, the expected error of the routinely applied approximations in Section 3.5, and the propagation of molecular energy errors on concrete examples of chemical reactions in Section 3.6. The analysis of larger electronic system is found in Section 3.7. The conclusion, Section 4, is followed by the Appendix A on the invariance of FCI with respect to the reference orbital set.

2 | COMPUTATIONAL DETAILS

All quantum chemical calculations were carried out using the MRCC program (version 2020-02-22)38 which was linked to the Llibxc density functional library (version 4.3.4).39 Sample input files are found in Section 1.1 of the Supplementary Material (SM). Optimized geometry and multiplicity of the studied molecules was taken from previous theoretical works.20,40

We note here that the correct formulation of the coupled cluster equations with perturbative inclusion of (triple) excitations for general non-HF single determinant reference requires additional terms involving the off-diagonal elements of the Fock-matrix41 as implemented in MRCC.42

2.1 | Electronic energies

For all selected model systems, coupled cluster calculations were performed using STO-3G,43 6-31G,44 cc-pVDZ, and cc-pVTZ45 basis sets, at CCSD, CCSD, CCSD(T), CCSDT, CCSDTQ, and CCSDTQP level of theory. Calculations containing even higher order excitations would
require unreasonably large computational resources, even for diatomic systems; nevertheless, as pointed out in the next sections, CCSDTQ produces converged energies of practically FCI quality. The self-consistent field (SCF) calculations generating the reference orbitals were performed with restricted (RHF/RKS) formalism in the case of singlet molecules and with unrestricted (UHF/UKS) formalism in the case of higher multiplicities.

Since we aim to draw general conclusions on the effect of Kohn–Sham orbitals on CC accuracy, we selected various types of density functionals for KS-CC calculations. Our conception was to pick one functional from each rung of the Jacobs’s ladder of KS-DFT density functionals for KS-CC calculations. Our conception was to pick one functional from each rung of the Jacobs’s ladder of KS-DFT methods. Following this logic of hierarchy, we chose the Hartree, LDA, BP86, M06-L, and B3LYP formulae as representative examples for the 0th, 1st, 2nd, 3rd, and 4th rung, respectively. Double hybrid functionals were omitted from the present study as their SCF procedure, hence the resulting set of orbitals, does not differ from that of conventional hybrid functionals. However, we added a range-separated hybrid, HSE06, as the sixth functional, which although belongs to the 4th rung, shows distinct features from regular hybrids in many aspects. Note that CC calculations on top of range-separated hybrid KS orbitals require a slight modification in the MRCC source code, which is described in the SM (Section 1.2).

We note that the generation of Brueckner orbitals, which could provide additional information on the role of orbital relaxation effects, is unfortunately not implemented in the MRCC program. Nevertheless, Brueckner CC calculations are not required for the comparison of HF-CC and KS-CC energies, on which the present work is focusing. The investigation of the performance of Brueckner CC compared to FCI might be the topic of a future research.

The accuracy of the electronic energy can be easily evaluated by comparing it to the (near-)FCI reference value. Still, it must be kept in mind that the implemented CC methods are non-variational, which means that a more accurate energy does not necessarily indicate a more accurate electronic wave function. Thus, to better establish the role of reference molecular orbitals, we also analyzed their effect on the three-dimensional function of electron density.

2.2 | Electron density

The unrelaxed electron density\(^{51}\) of a \(\Psi_{\text{CC}}\) or \(\Psi_{\text{KS-CC}}\) many-body wave function, expanded in the basis of HF or KS molecular orbitals \(\{\phi_r\}\), reads

\[
\rho(r) = \sum_{ij} P^r_{ij} \phi_m(r) \phi_n(r),
\]

with \(P\) one-electron reduced density matrix (1RDM) where, \(i, j,\) and \(n\) denote the spatial orbital and the spin index, respectively.

MRCC provides both the 1RDM matrix and the molecular orbital information required for the computation of density according to Equation (2). In fact, 1RDM is obtained from the CCDENSITIES output in FCIDUMP format,\(^ {52}\) while orbital data are saved as a MOLDEN file.\(^ {53}\)

The density was mapped to a finite grid by Multiwfn,\(^ {54}\) whose relevant input, that is, the data of the corresponding natural orbitals,\(^ {51}\) was generated by our hand-coded routine (see Section 1.3 of SM for details). In order to ensure sufficient numerical accuracy, the density was calculated for a grid, which was set fine enough to keep the deviation of the integrated electronic charge from the theoretical value below 0.1%.

Note that the short- and long-range behavior of the electron density can be revealed by studying its higher moments.\(^ {55}\)

3 | RESULTS AND DISCUSSION

3.1 | Selection of model systems

As the first step of our investigation, we selected a set of diatomic molecules with versatile degrees of multi-reference character. We investigated the diatoms in the equilibrium geometry of their ground electronic state, corresponding to the scope of typical CC studies. (That is, we do not consider, e.g., the dissociation process of diatoms where symmetry breaking issues affect HF-CC and KS-CC differently.)

The selection of model systems was based on the W4-17 dataset where the atomization energy is computed for 200 molecules at various HF-CC levels up to sextuple excitations, using complete basis set (CBS) extrapolation.\(^ {40}\) In contrast to \(T_1\) or other wave-function diagnostics, the accuracy of total atomization energies (TAE) are investigated as an unambiguous indicator of multi-reference effects. In Reference 40, the commonly applied CCSD, CCSD(T) or CCSDT level TAE results were compared to the TAE for CCSDTQ6 of FCI quality. The reliability of truncated CC methods can be quantified by the proportion of correlation effects taken into account by the calculation. The percentage of correlation energy (\(\text{TAE}_{\text{corr}}\%\)) recovered by a particular truncated method \(X\) is given by

\[
\text{TAE}_{\text{corr}}(X) = \frac{\Delta\text{TAE}(\text{HF}-X)}{\Delta\text{TAE}(\text{HF-CCSDTQ6})}.
\]

Here, we introduced the relative TAE for method HF-CCX, which measures the TAE difference obtained for HF-CCX and for HF-SCF level, that is,

\[
\Delta\text{TAE}(\text{HF}-X) = \text{TAE}(\text{HF}-X) - \text{TAE}(\text{HF-SCF}).
\]

Equation (3) gives 100% for methods of FCI quality and 0% for plain Hartree–Fock SCF. The typically applied CC methods are expected to give a value between 0% and 100%, depending on the CC level and the degree of multi-reference. Nevertheless, as the implemented coupled cluster equations are not based on the variational principle, values over 100%, that is, the overestimation of correlation effects, are also not excluded. We note that the advantageous feature of
TAEcorr\% is that it quantifies the multi-reference character in a system size independent manner, allowing simple comparisons among different molecules.

As summarized in Table 1, nine diatomic systems were chosen for the study (BH, HF, CF, NO, C₂, OF, B₂, and BN). We note that four out of the nine studied molecules are open-shell (CF, NO, OF, and B₂) and these four molecules are versatile in the degree of multireference character. Thus, the study covers both closed-shell and open-shell cases. These molecules are all sufficiently small for near-FCI calculations with basis sets of reasonable size. According to the TAEcorr\% covered by lower CC levels, the multi-reference character gradually increases from BH to B₂, as ordered from left to right in Table 1. The order is primarily based on TAEcorr\% for CCSDT, but it is mostly consistent with the relation of CCSD and CCSD(T) data. The rationale for the reversed order of NO and OF regarding TAEcorr\%(CCSDT) is the significantly lower TAEcorr\%(CCSD) and TAEcorr\%(CCSD(T)) percentages of the latter.

### 3.2 Convergence of HF-CC and KS-CC absolute electronic energies to the FCI limit

We began our investigations by computing the electronic energies of the nine molecules in Table 1 at different HF-CC and KS-CC levels.

In theory, HF-CC and KS-CC converge to the same energy limit, that is, to the FCI limit for a given molecule and basis set, when gradually increasing the maximal number of excitations, in spite of the fact that considerably different energies might result from HF-CC and KS-CC at the commonly used truncated levels, such as the popular CCSD(T). A mathematical proof for the invariance of the FCI limit with respect to the choice of reference orbitals is provided in the Appendix A. Importantly, once this universal energy limit is determined, the evaluation of the effect of reference orbitals becomes straightforward as the reference enabling the fastest convergence to FCI, that is, the reference producing the closest energy to FCI at a given CC level, can be considered the most beneficial.

The main issue with the latter approach is that the routinely applied frozen core (FC) and density fitting (DF) approximations are need to be turned off, otherwise the FCI limit is not independent of the reference in practice, which makes any rigorous comparison between HF-CC and KS-CC absolute electronic energies meaningless. All-electron coupled cluster calculations without density fitting are, however, very expensive, even for diatomic systems. In the case of STO-3G, 6-31G, and cc-pVDZ bases, we managed to perform calculations up to CCSDTQP level, and the deviation between HF-CCSDTQP and KS-CCSDTQP electronic energies was found to be below 0.05 kcal/mol for all molecules. Taken this negligibly low value together with the expected negligible contribution of sextuple excitations to the electronic energy (according to the W4-17 dataset, <0.1 kcal/mol change in TAE was observed, even for pathologically multi-reference systems), it can be stated that CCSDTQP results essentially represent the FCI limit,

$$E(\text{FCI}) \approx E(\text{CCSDTQP}).$$

(5)

In the case of the largest studied basis set, cc-pVTZ, the levels of CCSDTQ and CCSDTQP turned out to be unreasonably expensive for the majority of the diatomic model systems. Thus, being inspired by Reference 40, we estimated the FCI/cc-pVTZ limit by computing the frozen-core CCSDTQ electronic energy with density fitting and correcting the error of FC and DF at CCSDT level of theory,

$$E(\text{FCI}) \approx E(\text{CCSDTQ}) + \Delta E,$$

(6)

with $\Delta E = E(\text{CCSDTQ}) - E(\text{FC, DF)-CCSDTQ})$. Naturally, Equation (6) provides slightly less reliable numbers than Equation (5); still, as the energy is already closely independent of the reference orbitals at CCSDTQ level (the maximal observed deviation between HF-CCSDTQ and KS-CCSDTQ is 0.2 kcal/mol), these are suitable for investigating the performance of the commonly applied low-level HF-CC and KS-CC calculations with the triple zeta basis.

The resulting CCS, CCSD, CCSDT, CCSDTQ, CCSDTQP, and CCSD(T) electronic energies are summarized in Tables S1–S4 of the Supplementary Material. Estimated FCI limits for cc-pVTZ (as given by Equation 6) are shown in Table S9.

As the size of the basis set was found to have little influence on the relation of HF-CC and KS-CC results, the present section—and also Sections 3.3–3.6 of the paper—only discusses the evaluation of cc-pVDZ calculations. The analogous discussion of STO-3G, 6-31G, and cc-pVTZ results is provided in the SM.

The effect of reference orbitals on the convergence of CC/cc-pVDZ energies to the FCI limit is illustrated in Figure 1. Each of the seven chosen reference orbital sets (generated using the habitual Hartree–Fock method and the six Kohn–Sham density functionals listed in the Computational Details) is assigned to a color, as depicted in Table 1. Diatomic model systems selected for the present study

| Molecule | BH | HF | CF | NO | OF | BN | C₂ | B₂ |
|----------|----|----|----|----|----|----|----|----|
| **TAEcorr\%(HF-CCSD)** | 97.8 | 95.0 | 89.5 | 89.8 | 89.9 | 88.2 | 83.3 | 84.5 | 75.7 |
| **TAEcorr\%(HF-CCSDT)** | 99.8 | 99.7 | 99.5 | 99.3 | 99.0 | 99.1 | 98.0 | 97.8 | 97.0 |
| **TAEcorr\%(HF-CCSD(T))** | 99.8 | 100.1 | 99.8 | 100.1 | 99.6 | 98.5 | 100.3 | 99.6 | 96.9 |

Note: As measures of multi-reference character, the correlation contribution to CCSD/CBS, CCSDT/CBS and CCSD(T)/CBS atomization energies are provided in terms of the correlation contribution to CCSDTQP6/CBS total atomization energy, according to the W4-17 database. Consult Section 3 of the Supplementary material for the details of the calculations.
by the legend on the very top. Going from the top black frame to the bottom one of Figure 1, we present the performance of the increasingly accurate and expensive non-perturbative coupled cluster methods of CCSD, CCSDT, CCSDTQ, and CCSDTQ. In each frame, the horizontal axis shows the diatomic model systems in the order of increasing multi-reference character (see also Table 1). On the top
vertical axes, the percentage of recovered correlation energy is depicted, which is defined—analagously to Equation (3), in a molecule size independent manner—as

\[ E_{\text{corr}}\% (X) = 100 \frac{\Delta E(X)}{\Delta E(\text{HF-CCSDTQP})}. \]  

(7)

using \( \Delta E(X) = E(X) - E(\text{HF-SCF}) \), where \( E \) stands for the electronic energy and \( X \) is the computational method to be characterized. Note that this expression gives 100\% for \( X = \text{HF-CCSDTQP} \), and should negligibly deviate from 100\% in the case of any KS-CCSDTQP energy, as any of these are practically equal to the FCI limit. For sake of brevity, we introduce notation \( E_{\text{corr}} = \Delta E(\text{HF-CCSDTQP}) \). It is notable that all electronic energies, including KS-CC results, were analyzed in terms of the dimensionless \( E_{\text{corr}}\% (X) \) using the HF referenced electronic energies. In this way, \( E(X) \) is the only quantity in the measure, which varies with the reference, accordingly it is ensured that a higher \( E_{\text{corr}}\% \) percentage indicates higher accuracy, that is, smaller deviation from the near-FCI energy.

The convergence to the FCI energy limit upon increasing the CC level can be observed on the change of the scale of the top vertical axes in Figure 1. At CCSD level (top frame of Figure 1) 89\%–99\% of the correlation energy is recovered, depending on the molecule and the reference orbitals. The inclusion of higher and higher excitations to the CC ansatz brings the depicted values closer and closer to the FCI limit of 100\%, that is, CCSDT: 98.7\%–100.0\%; CCSDTQ: 99.82\%–100.01\%. At CCSDTQP level, all values reach 100.0\% with a negligible variation of ±0.02\% in \( E_{\text{corr}}\% \), regardless of the multi-reference character of the molecule and the choice of reference orbitals.

The bottom vertical axes represent the alteration in the recovered correlation energy percentage upon changing the usual HF reference to KS orbitals, which we defined as

\[ \Delta E_{\text{corr}}\% (\text{KS-CC}) = \left| 100 - E_{\text{corr}}\% (\text{HF-CC}) \right| - \left| 100 - E_{\text{corr}}\% (\text{KS-CC}) \right|. \]  

(8)

A positive \( \Delta E_{\text{corr}}\% \) indicates the improvement effect of the given KS reference, that is, \( E_{\text{corr}}\% (\text{KS-CC}) \) approximate more accurately the FCI limit of 100\%, while a negative value means deterioration relative to the HF-CCX result. We note that the absolute values in Equation (8) are introduced due to the non-variational nature of the applied coupled cluster implementation, which might lead not only to the underestimation, but also to the overestimation of correlation energy (\( E_{\text{corr}}\% > 100\% \)) at truncated CC levels.

Benchmarks are summarized in Figure 1 which clearly shows that the change of the habitual HF reference (black bars) to KS (colored bars) does not yield any improvement relative to the FCI limit for the commonly used functionals (\( \Delta E_{\text{corr}}\% < 0 \)). In contrast to this, the most basic Hartree formula (pink bars) provides an opposite trend for some of the studied molecules and it might facilitate the convergence to FCI (\( \Delta E_{\text{corr}}\% > 0 \)). The interesting tendency might be interpreted as the consequence of the intimate relation of the Hartree and the Hartree–Fock functionals, which is also witnessed at low orders in the Møller-Plesset perturbation theory, which is closely connected to the CC expansion.51

In particular, at CCSD level (top black frame of Figure 1), we observe that the use of any of LDA, BP86, M06-L, B3LYP, and HSE06 references consistently decreases the accuracy of the electronic energy. The deviation from HF-CCSD in the unfavorable direction ranges from a minimal change of 0.0\%–0.1\% in terms of correlation energy (BH, HF) to up to 0.5\% in moderately multi-reference cases, and up to 1\% in the case of the four highly multi-reference systems.

Apart from the least correlated BH and HF molecules, the KS-CCSD energies—with the exception of those based on Hartree orbitals—are typically more similar to one another than to the Hartree-CCSD results, that is, the difference between the highest and lowest \( \Delta E_{\text{corr}}\% \) values is generally below 0.2\% and even the highest variations, observed at exceptionally high multi-reference character (BN, B$_2$), do not exceed 0.35\%. Even though this result is in complete agreement with previous findings,20 it is still a remarkable outcome considering the great variety of density functionals from local density approximation to hybrid formulae.

We also note that the order of functionals in KS-CC accuracy somewhat reflects the Jacob’s ladder of DFT, that is, LDA from the 1st rung (blue color) and HSE06 from the 4th rung (red color) appear to be the least and most favorable functionals, respectively. Nevertheless, even the most complex functional produces a less beneficial reference than the HF method. For example, using HSE06, a 2.3 kcal/mol discrepancy is found for BN, see Table S3 of the SM.

In the unique, deviate case of the Hartree functional (pink color), for most of the investigated systems, Hartree reference provides lower KS-CCSD energy than the HF-CCSD, that is, five molecules out of nine give positive \( \Delta E_{\text{corr}}\% \), among which the values of 1.9\% (C$_2$), 1.3\% (B$_2$), and 0.9\% (NO) are the most significant. On the other hand, the improvement is not systematic as Hartree-CCSD provided the least accurate result in the case of BF and three other negative \( \Delta E_{\text{corr}}\% \) values were also found.

Although the inclusion of triple excitations (CCSDT/cc-pVQZ frame of Figure 1) considerably decreases the variation of electronic energies as expected upon approaching the FCI limit, the above described trends are sustained. KS-CCSDT results, which remain practically independent of the functional choice, underperform HF-CCSDT by 0.1\%–0.3\% of the correlation energy, going from low to high multi-reference character. In particular, for BH and HF, \( \Delta E_{\text{corr}}\% \) decreases practically to zero, indicating that CCSDT is of near-FCI quality due to the low extent of correlation.

Taking another step further to CCSDTQ level, the first four molecules of BH, HF, CF, and BF are found to reach the convergence of electronic energy, regardless of the reference choice. The deviation of KS-CCSDTQ from HF-CCSDTQ is barely detectable (\( \Delta E_{\text{corr}}\% \leq 0.01\% \)). For NO and the four highest multi-reference cases, the previously noted trends apply, that is, HF-CCSDTQ provides the closest approximation of the FCI limit, slightly outperforming KS-CCSDTQ by 0.01\%–0.1\% of \( E_{\text{corr}} \), where the magnitude of deviation depends on
the degree of multi-reference character. The effect of orbital choice within the Kohn–Sham framework has negligible influence.

As the next step, we extended the established analyses for the most commonly applied CCSD(T) method. These results are presented in Figure 2. The recovered CCSD(T) correlation energy percentages are quite similar to those at CCSDT level. (We note that although CCSD(T) is designed as an approximation of CCSDT, the perturbative method quite often produces more accurate energetics in terms of $E_{\text{corr}}$% than CCSDT. This phenomenon has been reported for various molecular systems [see Reference 40], and can be attributed to a fortunate error cancellation in the (T) approximation.) $\Delta E_{\text{corr}}$% values, however, range down to –1%, which better recalls the CCSD results. The behavior of high multi-reference systems closely follows the patterns of Figure 1, that is, the change from HF to any KS reference does not result in any improvement of the electronic energy, and the extent of deviations between HF-CC and KS-CC has little dependence on the density functional. The case of low multi-reference character is slightly more complex, as certain systems show distinct CCSD(T) results, that is, positive $\Delta E_{\text{corr}}$% values appear with non-Hartree references, ranging up to +0.1% (for BH and BF). Still, the practical significance of this magnitude of improvement is debatable, as it seems to be rather incidental and it is peculiar to two of the least correlated systems, ranging up to $0.3\%$, $0.1\%$, $0.1\%$, and $0.02\%$ of $E_{\text{corr}}$ at CCSD, CCSD(T), CCSDT, and CCSDTQ level, respectively. The choice of the KS functional has minimal impact (with the exception of the hectically performing Hartree formula). It should not escape our attention that, given that these proportions also apply to larger molecules such as transition metal complexes, the reference dependent deviation might easily reach multiple kcal/mol units for practically relevant chemical reactions.

From the practical point of view, it is also important to consider the consistency of molecular energy errors. Namely, individual errors of similar size favorably cancel out during the calculation of energy differences (e.g., the energy change of a given chemical reaction), no matter whether this size is high or low. This leaves the possibility that a methodology producing higher molecular errors (in the present case, KS-CC) might be beneficial for theoretical chemical applications. To examine this question, we quantified the consistency of errors by the standard deviation of $E_{\text{corr}}(X)$ values among the nine studied molecules. These numbers are summarized in the bottom heatmap of Figure 3 for each CC level—reference orbital pair. Analogously to the above discussed average errors, all (non-Hartree) KS-CC standard deviations show a small but systematic increase relative to HF-CC ($0.3\%$, $0.1\%$, $0.1\%$, and $0.02\%$ of $E_{\text{corr}}$% at CCSD, CCSD(T), CCSDT, and CCSDTQ level, respectively). The clear analogy of the two heatmaps reveals that the depicted average and standard deviation values well correlate (see also Figure S4 of SM for a linear correlation plot). Consequently, KS-CC electronic energies are neither more accurate, nor more consistent in accuracy than HF-CC.

Finally, we also performed the same analysis for further basis sets, the corresponding $E_{\text{corr}}(X)$ data is found Section 4 of the SM. Numerical data obtained for STO-3G, 6-31G, cc-pVDZ, and cc-pVTZ basis sets are compiled in Tables S11–S14 of the SM, respectively. The figures
of the error percentages for STO-3G, 6-31G, and cc-pVTZ basis sets are given in Figures S1–S3 of the SM, respectively, which reveal essentially similar tendencies discussed for cc-pVQDZ set above.

3.3 Convergence of HF-CC and KS-CC electron densities to the FCI limit

It is notable that the applied CC implementation is not a variational approach hence lower CC energy does not necessarily indicate a more accurate description of the problem. Therefore, in the following, we attempt to assess the accuracy of the computations in an alternative way, that is, we benchmark the precision of the CC calculations in terms of the associated electron density instead of the numerically demanding direct comparison of the corresponding many-body wave functions. The application of this analysis is also motivated by getting physical insight on the effect of using various correlation-exchange functionals as we not only monitor densities obtained for the coupled cluster excitation levels but for Kohn–Sham theory as well.

In practice, densities to be compared are evaluated on the same finite grid. In general, the similarity of two vectors, such as electron densities to the FCI limit, is defined as

\[ d(X) = \frac{|\rho(X) - \rho(\text{HF-SCF})|}{|\rho(\text{HF-SCF})|}, \]

which measures the error in units of $D(\text{HF-SCF})$. By definition, $d(X)$ is exactly 100% and 0% for $X = \text{HF-SCF}$ and $X = \text{HF-CCSDTQP}$, respectively. For general $X$ cases, where the corresponding wave function recovers a significant portion of the correlation effects, $d(X)$ tends towards 0%. The obtained $d(X)$ values are presented in Figure 4 for various CC/cc-pVQDZ levels.

Comparing SCF level error measures, we find that the high-level Kohn–Sham theories provide typically the most accurate single-determinant approximation, that is, the corresponding errors are a small fraction of the HF-SCF error. Interestingly, molecules of intermediate multi-reference character are found to be represented the most faithfully by the high-level KS-SCF theories. Also in agreement with the expectations, the error of the solution Hartree-SCF, which neglects any exchange and correlation effects, is outstandingly large, that is, it is typically multiple of $D(\text{HF-SCF})$.

For increasing CC excitation levels, the tendencies of electron-density errors are generally in line with the conclusions of the correlation-energy analysis detailed in Section 3.2. (See Figure S7 of SM for a correlation plot between $E_{\text{cor}}(\%)(X)$ and $d(X)$ values.) Most importantly, the lowest deviations are typically found for CC results of Hartree and Hartree–Fock references, that is, the addition of computationally rather cheap CCSD excitations already compensate their limitations found at SCF level of theory reflecting the Thouless theorem. Similarly to the observations found for CC energies, the deviation of electron-density error is quite even for the various KS functionals compared to tendencies found for the aforementioned correlation-free references. In line with expectations corroborated by the findings for molecular energy, low-level CC theory also shows an overall tendency to provide more accurate electron density for single-reference molecules. Also remarkable that low-level CC theory applied on high-level functionals (B3LYP and HS06) results in little improvement compared to the outstandingly accurate SCF results, especially in the case of moderately correlated problems. In this sense, KS-CCSD can even be regarded as a counterproductive approach.

Furthermore, Figure 5 (which compares the average and the standard deviation of $d(X)$ values analogously to Figure 3) also
reveals that KS-CC results are less consistent in the error of density than those based on HF-CC. We observe near direct proportionality between averages and standard deviations (see Figure S8 of SM for a linear correlation plot), indicating that the less accurate methods are also less prone to favorable error cancellations in general.

For CCSDTQP level of theory, the error of the wave functions is already found to be practically zero confirming the convergence of the high-level CC calculations to the exact solution.

Details of the numerical data is found in Section 5 of the SM. In particular, d%\textit{X} data obtained for STO-3G, 6-31G, and cc-pVDZ basis sets are compiled in Tables S15–S17 of SM, respectively. The
corresponding figures of the error of the densities for STO-3G and 6-31G basis sets are given in Figures S5 and S6 of the SM, respectively, which essentially corroborate findings presented for cc-pVDZ.

3.4 The relation of $T_1$ and $\max(|t_2|)$ diagnostics to the accuracy of electronic energy

The $T_1$ diagnostic of Equation (1) and the $\max(|t_2|)$ value, that is, the largest absolute value among doubles amplitudes, are commonly applied in the literature to check the reliability of coupled cluster calculations. Thus, we were naturally interested in how the reference choice affects these wave-function diagnostics. Importantly, we do not intend to use these diagnostics to evaluate the quality of orbitals from another aspect besides energy and density—we would like to emphasize again that the $T_1$ based “overlap criterion” (which poses the Brueckner orbitals with $T_1 = 0$ as optimal reference) has limited relevance from the practical point of view. Instead, we investigate whether the accuracy trends described in the previous sections (especially the outstanding performance of HF-CC) are reflected in $T_1$ and $\max(|t_2|)$.

Figure 6A shows the $T_1$ diagnostic for the diatomic model systems at CCSD(T)/cc-pVDZ level of theory. The magnitude of singles CC amplitudes has remarkable reference dependence, that is, in most cases, $T_1$ value for HF based result are drastically larger relative to KS counterparts. We find that all types of non-Hartree density functionals typically decrease the $T_1$ diagnostic except for a minor number of increases with LDA and BP86 functionals in the case of the least multi-reference BH, HF, and BF systems. Furthermore, with the exception of BN and B2, the decreasing order of $T_1$ values from LDA
to HSE06 fairly corresponds to the Jacob’s ladder hierarchy. The most conspicuous example is probably OF, where the HF-based $T_1$ diagnostic of 0.04 is already reduced to 0.026 when using a simple LDA functional, and is decreased to as low as 0.007 with HSE06.

These findings are in line with the previous KS-CC studies reporting decreased $T_1$ values relative to HF-CC. Furthermore, we also observe an overall trend of increased $T_1$ values for the more correlated problems in the case of the common references. In contrast to this trend, in the case of CC solutions based on the Hartree reference, which completely neglects quantum exchange and correlation effects, the diagnostic produces exceedingly large value, that is, 0.11–0.15, regardless of the particular molecule.

The tendencies found for $T_1$ might be interpreted as the footprint of orbital relaxation effect induced by single excitations, that is, larger $T_1$ is detected for the less accurate single-determinant approximations of Hartree–Fock and—in particular—Hartree theories. This argument is corroborated by our systematic numerical investigation of energy and density, see Sections 3.2 and 3.3, which also shows that lower $T_1$ values do not indicate more accurate approximation of the molecule.

The reference dependence of the other widely used wave-function diagnostic, max($\langle t_2 \rangle$), is plotted in Figure 6B. In contrast to $T_1$ diagnostic, max($\langle t_2 \rangle$) is found to be practically independent of the reference, with a variation below 0.01 in most cases. Even the Hartree functional, which otherwise always shows distinctive behavior, fits the rest of the data points, though it slightly stands out at high multi-reference character. It is also notable that max($\langle t_2 \rangle$) systematically increases for problems of strongly correlated character, see the outstandingly large max($\langle t_2 \rangle$) values for BN, C$_2$, and B$_2$.

To conclude on the use of CC-amplitude based diagnostics, according to our numerical exploration, max($\langle t_2 \rangle$) is found to be of small magnitude, rather than $T_1$, could be applied to estimate the degree of multi-reference character, hence the reliability of lower-level truncated CC calculations. Nevertheless, we found no obvious relation between the reference dependent deviations in the diagnostic value and the relative accuracy of HF-CC and different KS-CC methods reflected by energy and electron density.

Further details can be found in Section 6 of the SM. Numerical data of the diagnostics obtained for STO-3G, 6-31G, cc-pVDZ, and cc-pVTZ basis sets are compiled in Tables S18–S21 of the SM, respectively. Diagnostics figures for STO-3G, 6-31G, and cc-pVTZ basis sets are given in Figures S9–S11 of the SM, respectively, which show similar tendencies discussed for cc-pVDZ basis set.

### 3.5 | Effects of frozen core and density fitting approximations

Until this point, we discussed coupled cluster calculations without any additional approximations, in order to focus purely on the effects of reference choice and exclude any other factors. In practice, however, frozen core (FC) approximation and density fitting (DF) are commonly applied to speed up calculations, which introduce additional sources of error in the electronic energy. In the case of FC, the core orbitals are kept doubly occupied in the CC calculations, thus the number of active orbitals can be significantly reduced for heavy elements by the negligible loss of correlation effects of the low-lying core electrons. DF is used to provide computationally cheap approximation for numerical integrals, correspondingly, we found that the error introduced by DF is detectable but rather marginal compared to the chemical accuracy scale.

To investigate the magnitude and the reference dependence of these typical error sources, we repeated the energy calculations described in Section 3.2 applying both FC and DF. The obtained CC electronic energies are summarized in Tables S5–S8 of the Supplementary Material.

As already discussed in Section 3.2, the use of FC and DF breaks down the theoretically established reference invariance of the FCI energy; thus, we firstly examined how the reference affects the approximated CCSDTQ/P/cc-pVDZ (practically FCI/cc-pVDZ) level energy. In Figure 7, we plotted the deviation of these energies from those computed at non-approximated CCSDTQ/P/cc-pVDZ level, expressing it in terms of the percentage of the exact correlation energy, $E_{\text{corr}}$.

Figure 7 clearly shows that the Hartree–Fock reference comes with the smallest error, which amounts approximately 1% of the correlation energy for all molecules. As for Kohn–Sham referenced methods, the error of approximations consistently follows the order

![Figure 7](image-url)
of M06-L < HSE06 < B3LYP < BP86 < LDA < Hartree. The first four functionals in the latter order, which produce errors between 1% and 2% of $E_{\text{corr}}$, show very similar behavior, in the sense that their approximated FCI limits are closer to each other than to that of HF-CC. In the case of LDA, the use of frozen core and density fitting causes 2% and 4% error. Most notably, due to the complete neglect of exchange processes of core electrons, 10%–30% portion of $E_{\text{corr}}$ is lost with Hartree references which also implies that the application of Hartree orbitals is severely limited for small electronic systems.

Of course, the magnitude of the error depends on the CC level using FC and DF approximations, but we find that the truncation of the CC ansatz leaves the results shown in Figure 7 practically unchanged, see SM for the error of approximations with CCSD, CCSD(T) and CCSDT. That is, the same amount of electronic energy (as given in Figure 7) is neglected by frozen core and density fitting at all CC levels, implying that the slower convergence of KS-CC methods to the FCI limit, as established in Section 3.2, also holds true for approximated CC. This tendency also reveals that the contribution of the cores is rather marginal in the overall correlation effects validating the applicability of the FC approximation.

Altogether, our results indicate that the KS-CC methods are not only unable to accelerate the convergence towards FCI, but they are also more sensitive to the common approximations. The consistent performance of the HF reference is rather natural as these orbitals are optimized for describing exchange effects, which can be also captured in the FC approximation at SCF level of theory.

Further numerical details can be found in Section 7 of the SM. In particular, data of the computed error percentages obtained for STO-3G, 6-31G, cc-pVDZ, and cc-pVTZ basis sets are compiled in Tables S22–S25 of the SM, respectively. The corresponding figures of the error percentages for STO-3G, 6-31G, and cc-pVTZ basis sets are given in Figures S12–S14 of SM, respectively, which imply similar features presented for cc-pVDZ basis set above.

### 3.6 Implications on the accuracy of reaction energies: Error propagation

As we have seen in the previous sections, the use of Hartree–Fock reference orbitals in truncated CC calculations is expected to lead to the closest approximation of the FCI limit of the molecular electronic energy $E$. In quantum chemical studies, however, we are typically interested in the accuracy of reaction energies ($\Delta E_r$), which are calculated as the difference of the algebraic sum of the individual $E$ energy of the product molecules and the reactants,

$$\Delta E_r = \sum_{\text{products}} E - \sum_{\text{reactants}} E.$$  \hspace{1cm} (12)

Consequently, the error of reaction energy, $\delta \Delta E_r$, can be expressed by the individual molecular errors obtained for method X, $\delta E = E(X) - E(\text{FCI})$, as

$$\delta \Delta E_r = \sum_{\text{products}} \delta E - \sum_{\text{reactants}} \delta E.$$ \hspace{1cm} (13)

Owing to the balancing nature of Equation (12) for $\Delta E_r$, the accuracy of the reaction energy is governed not only by the magnitude of the molecular errors involved in the reaction but also by their variance. Therefore, even though we have shown in the previous subsections that the truncated HF-CC model is more reliable than the KS-CC methods with standard functionals, it does not necessarily predict the most accurate reaction energy, that is, the smallest total error in Equation (13). In this section, by illustrative reactions constructed of the computed diatoms we investigate the accuracy of reaction energy to be expected.

A typical benchmark calculation in quantum chemistry is performed at CCSD(T) level, with FC and DF approximations turned on. The bar chart on top left of Figure 8 shows the size of molecular errors ($\delta E$) for this case, in kcal/mol unit. We note that we only obtained errors of positive sign, that is, all (FC,DF)-CCSD(T)/cc-pVDZ electronic energies are more positive than the corresponding FCI/cc-pVDZ limit. In line with results presented in Sections 3.2–3.5, KS-CC errors always exceed the corresponding HF-CC error, typically by 1–2 kcal/mol. When plotting the errors as a function of the reference orbital set (Figure 8, top right), it can also be observed that KS-CC methods increase not only the size, but also the variance of individual molecular errors: while $\delta E$s with HF-CC are located within an interval of merely 3 kcal/mol ($\delta E = 1–4$ kcal/mol), KS-CC with the commonly applied GGA, meta-GGA and hybrid functional types (BP86, M06-L, B3LYP, and HSE06) gives slightly larger values, distributed in a slightly wider range between 1 and 5 kcal/mol. In the case of LDA, even larger errors cover an even wider interval of 5 kcal/mol ($\delta E = 2–7$ kcal/mol). Finally, as the most extreme example, the use of Hartree orbitals as reference results in outstandingly high errors with outstanding variance ($\delta E = 15–42$ kcal/mol).

The impact of the distribution of errors on the accuracy of chemical reaction energies is visualized in the bottom graphs of Figure 8. Here, two hypothetical reactions among the diatomic species (OF + BN $\rightarrow$ NO + BF; BN + OF + BH $\rightarrow$ B$_2$ + HF + NO) are analyzed by plotting the molecular errors with the appropriate sign: as all $\delta E$ values are positive in Equation (13) in the case of (FC,DF)-CCSD(T)/cc-pVDZ calculations, the products and reactants contribute to the total error by positive and negative terms, respectively. As a result, the major part of molecular errors is canceled out and a relatively small final reaction energy error, shown in the lowest table rows of Figure 8, remains. Interestingly, although KS-CC clearly produces higher errors in both the initial and the product molecules in the case of both presented reactions, only OF + BN $\rightarrow$ NO + BF follows the expected pattern of (absolute) total errors (i.e., $|\delta \Delta E_+(\text{KS-CC})| > |\delta \Delta E_-(\text{HF-CC})|$ and the error of KS-CC decreases with the increasing complexity of the density functional.) For the other reaction, BN + OF + BH $\rightarrow$ B$_2$ + HF + NO, the Hartree reference remains the most inaccurate ($|\delta \Delta E_+| = 4.5$ kcal/mol) but unexpectedly small KS-CC total errors are obtained for other functionals: most KS-
CC reaction energy errors ($\delta \Delta E_r = 0.6\text{--}1.1 \text{ kcal/mol}$) half the HF-CC counterpart ($\delta \Delta E_r = 2.2 \text{ kcal/mol}$). As the KS-CC methodology is not more reliable than HF-CC in general (considering the molecular error contributions), it is clear that a case of a fortunate error cancellation was found.

It is important to note here that the probability of such error cancellations strongly depends on the variation of molecular errors: as KS-CC errors are spread in a wider range compared to HF-CC results, the chance for near-equal positive and negative KS-CC error contributions is expected to be relatively low. Nevertheless, a fortuitous cancellation of molecular errors occasionally results in the outperformance of KS-CC. These expectations are completely corroborated by our numerical exploration conducted on a set of around 100 chemical reactions allowed among the diatoms (see Tables S26--S28 of SM for details). In 53 cases out of 91 investigated cases, the HF-CCSD(T) calculation was found to provide better approximation for the FCI reaction energy than any KS-CCSD(T) method with conventional functionals. Statistics given in Table S27 is even more informative, that is, the mean absolute error of KS-CCSD(T) reaction energies (3.5--7.5 kcal/mol) is clearly above that of HF-CCSD(T) (2.9 kcal/mol). The same relation between HF-CC and KS-CC holds true for the standard deviation of errors, which further indicates that only occasional—and not systematic—improvements are expected upon changing to a KS reference.

It is notable that the insensitivity on the functional formula (beyond LDA) has great potential in practical coupled cluster
calculations: it implies that a KS-CC study does not require an extensive functional selection procedure, in other words, almost any arbitrary functional (including the cheapest GGA functionals) might be suitable for KS-CC investigations.

3.7 Can the results be generalized to larger systems? – Example of a practically relevant transition metal complex

Given our results obtained for diatomic model systems, it is naturally of interest whether similar trends are also valid for larger molecules including transition metal (TM) complexes. Thus, we searched for a practically relevant complex, which is however small enough to perform calculations of closely FCI quality. After deliberation and performing extensive tests on the limits of computational performance, we selected the helium model of Harvey. This tetragonal bipyramidal complex, structured as [FeHe5O]2+, is constructed in a way to resemble the spin density of C–H bond activating iron-oxo catalysts.

As the final part of our study, we decided on examining the effect of reference orbital choice on a simple chemical reaction involving [FeHe5O]2+ species. We selected the triplet-to-quintet spin transition of the iron complex for this purpose similar to Harvey et al. Considering basis sets of reasonable size, such as cc-pVDZ, all-electron CCSD(T) calculations, used as near-FCI reference for diatomic molecules, would be unfeasible even for such a small TM complex. Nevertheless, it is possible to calculate the electronic energy of triplet and quintet states at CCSD(T)/cc-pVDZ(Fe:cc-pVTZ) level with frozen core and density fitting approximations. The label of cc-pVDZ(Fe:cc-pVTZ) refers to a custom basis set consisting of cc-pVTZ functions on the central iron atom and cc-pVDZ functions on all other atoms. (We note that the use of triple zeta on Fe is inevitable to obtain realistic spin state energetics.)

Furthermore, corresponding to findings of Section 3.5, we attempted to approximate all-electron CCSD(T)/cc-pVDZ(Fe:cc-pVTZ) level energy as

$$E(\text{CCSD}(T)) \approx E(\text{FC, DF}) - E(\text{CCSD}(T)) \approx \Delta E_{\text{FC,DF}}.$$  (14)

that is, the error of FC and DF can be already estimated at low CC levels, such as

$$\Delta E_{\text{FC,DF}} = E(\text{CCSD}(T)) - E(\text{FC, DF}) - E(\text{CCSD}(T)).$$  (15)

Finding that the approximated energy of Equation (14) is nearly independent of the orbital choice for the investigated TM model, the numerical result can be considered of near-FCI quality.

More specifically, the difference between HF and B3LYP referenced triplet-quintet gaps obtained in this manner ($E_{\text{gap}}(\text{HF-FCI estimation}) = -5.5$ kcal/mol and $E_{\text{gap}}(\text{B3LYP-FCI estimation}) = -7.2$ kcal/mol, respectively) is 1.7 kcal/mol, which approaches the requirement of chemical accuracy and indicates close—albeit not full—convergence to the FCI limit. Note that the non-negligible discrepancy of the results for the high-level theory clearly indicates that the CC solution is relatively sensitive to the orbital choice in this few-atomic multi-reference problem. Still, the order of accuracy between HF and KS referenced truncated CC methods can be reasonably evaluated as described below.

In the following discussion, we treat the average of the latter estimated values (~6.4 kcal/mol, as depicted by the horizontal dashed red line in Figure 9) as the reference gap of near-FCI quality, keeping in mind its uncertainty of 1.7/2 ≈ 0.9 kcal/mol (corresponding to the dim red range in Figure 9). We note that only the popular B3LYP functional was tested here among the KS-CC methods due to the significant computational cost.

The top graph of Figure 9 visualizes the effect of orbital choice on the triplet-quintet gap at truncated levels of CCSD, CCSD(T), and CCSD(T) with FC and DF approximations. (Derivation of the data used for the preparation of the figure can be found in Tables S29–S32 of the SM in Section 9.) Though the FCI limit of the gap is not exactly known, the performance of (FC,DF)-B3LYP-CC is unequivocally not superior to (FC,DF)-HF-CC. Namely, (FC,DF)-HF-CC results converge to the FCI estimation (see black dots), while (FC,DF)-B3LYP-CC leans towards −10.1 kcal/mol (see orange dots), which is considerably far outside of uncertainty range of the FCI estimation. At the most commonly applied level of (FC,DF)-CCSD(T), the HF referenced gap of −5.6 kcal/mol is more consistent with the −6.4 ± 0.9 kcal/mol FCI estimation than the B3LYP referenced value of −8.6 kcal/mol.

![FIGURE 9](image-url)  
Analysis of the triplet and quintet states of the helium model of [FeHe5O]2+. Top: Energy gap computed at truncated CC models with FC and DF. All-electron FCI energy is estimated according to Equation (14). Bottom: Wave-function diagnostics using $T_1$ and $\text{max}(|\mathcal{J}_2|)$ measures.
The deterioration of the results by the B3LYP reference is attributed partly to the slower convergence of KS-CC electronic energies to the FCI limit and partly to the larger frozen-core error of (FC,DF)-B3LYP-CC (see Section 9 of the SM for a detailed analysis of molecular energies), which is in complete agreement with our observations on smaller diatomic systems discussed in Section 3.5.

Moreover, the reference dependence of wave-function diagnostics also follows the trends described in Section 3.4 as summarized in the bottom of Figure 9, that is, the $T_1$ diagnostic provides considerably lower values for B3LYP-CCSD(T) than for HF-CCSD(T) and max($l_2$) is essentially reference-free.

Altogether, these findings suggest that the conclusions drawn for the nine diatomic molecules are generally applicable for KS-CC. It can be stated with high certainty that this methodology does not describe the electronic structure of TM complexes or other molecular systems more reliably than HF-CC. Thus, the previously reported “improvement” of CC results upon changing to KS reference\textsuperscript{17–19} most likely results merely from the accidental cancellation of KS-CC errors and does not reflect the typical behavior of KS-CC.

4 | CONCLUSION

The present computational study investigated the effect of orbital choice on the accuracy of coupled cluster calculations compared to reference calculations of FCI quality for the first time to the best of our knowledge. We selected seven different sets of reference orbitals (HF and six different KS density functionals: LDA, BP86, M06-L, B3LYP, HSE06 and—for the sake of completeness—the Hartree formula) and four different basis sets (STO-3G, 6-31G, cc-pVQZ, and cc-pVTZ) to perform the latter calculations on nine small closed- and open-shell molecules with increasing multi-reference character (BH, HF, CF, BF, NO, OF, BN, C$_2$, and B$_2$). Additional analysis of high-level CC calculations on transition metal complexes were also performed to study larger and practically relevant molecules. According to our numerical exploration, the following points should be taken into consideration in future CC studies involving Kohn–Sham reference orbitals.

1. KS-CC calculations, regardless of the exchange-correlation density functional used for the generation of reference orbitals, do not produce more accurate molecular electronic energies and densities (relative to the FCI result) than the habitual HF-CC approach with equal CC level and basis set. HF-CC and KS-CC all-electron methods converge to the same FCI limit upon increasing the CC level, but the convergence behavior is different—according to our study, the convergence to FCI is somewhat slower with common KS references. What is more, the error of the routinely applied frozen core and density fitting approximations appears to be larger for KS-CC, which also confirms that one cannot obtain more reliable energetics by simply replacing the HF orbitals of the reference determinant to KS orbitals.

2. The choice of KS density functional has little impact on the result of the KS-CC calculation beyond the LDA approximation; that is, even the cheapest GGA functionals can be used for KS-CC studies with essentially equal reliability to complex hybrid functionals.

3. The $T_1$ diagnostic of KS-CC wave functions of typical functionals is considerably lower than that of the corresponding HF-CC wave functions in line with literature findings. Nevertheless, this seemingly favorable decrease in $T_1$ is neither corroborated by the improvement of correlation energy, nor by the accuracy of the corresponding density. On the other hand, the largest $t_2$ amplitudes seem as a reference-free indicator of multi-reference character.

4. In spite of the higher errors in individual molecular electronic energies, KS-CC might still outperform HF-CC in the accuracy of reaction energies corroborating previous observations as a result of cancellation of errors. If discrepancy is encountered between HF-CC and KS-CC reaction energies, both values should be handled cautiously, as no general strategy can be developed to determine the order of accuracy.

Although the application of KS-CC might seem pointless considering the above conclusions, we emphasize that this methodology is a very useful alternative of HF-CC in case convergence issues encountered in HF-SCF or HF-CC. At sufficiently high CC level, determined by the multi-reference character of the studied molecule, the difference between HF-CC and KS-CC results disappears in the calculations and both can be reliably applied.

ACKNOWLEDGMENTS

Discussions with Adám Gali, Mihály Kállay, and Péter R. Nagy as well as support from the National Research, Development and Innovation Office (NKFIH FK-20-135496) are greatly acknowledged. Paula Timár and Gergely Barcza thank the support from the Wigner Student Scholarship of the Wigner Research Centre for Physics and the Bolyai Research Scholarship of the Hungarian Academy of Sciences, respectively. We acknowledge KIFÜ for awarding us access to computational resources based in Hungary. Zsolt Benedek, Tibor Szilvási, and Gergely Barcza would like to thank the University of Alabama and the Office of Information Technology for providing high performance computing resources and support that have contributed to these research results. This work was made possible in part by a grant of high-performance computing resources and technical support from the Alabama Supercomputer Authority. Dedicated to the loving memory of professors Géza Tichy, János Pipek, and Gyula Radnay.

DATA AVAILABILITY STATEMENT

Most of the numerical data (energies and diagnostic values), which were analyzed in this study, are given in the supplementary material while the generated density data are available from the authors on reasonable request.

ORCID

Zsolt Benedek https://orcid.org/0000-0002-0174-9676
Tibor Szilvási https://orcid.org/0000-0002-4218-1570
Gergely Barcza https://orcid.org/0000-0001-5612-7447
In the following we give an illustrative mathematical derivation for the invariance of the FCI solution with respect to the choice of reference orbitals.

Let us consider two different sets of reference orbitals, for example, a set of Hartree–Fock and a set of Kohn–Sham molecular orbitals which will be referred to as $\phi^F$ and $\phi^K$ in the following, respectively. Considering that analogous formulae are derived for both orbital sets, for the sake of compactness, double notation is introduced in the following derivations, for example, $\phi^{(i)}$ denotes both $\phi^F$ and $\phi^K$ simultaneously.

The basis consisting of $m$ basis functions $\phi_j$ is used to expand the molecular orbitals as

$$\phi^{(i)} = \sum_{j=1}^{m} c^{(i)}_j \phi_j.$$  \hspace{1cm} (16)

Note that the HF and KS procedure creates $m$ molecular orbitals (MOs) from the $m$ atomic basis functions. However, constructing a...
Slater determinant $\Phi^{(i)}$, only $n$ MOs, corresponding to the number of electrons, are selected out of $m$ MOs. The $k$th Slater determinant of the wave function ansatz, containing the $k_1$th, $k_2$th, ..., $k_d$th orbitals reads as

$$
\Phi^{(i)} = \hat{A} \left( \prod_{i=1}^{n} \phi^{(i)}(l) \right),
$$

(17)

where, $\hat{A}$ is the antisymmetrizing operator and the parentheses indicate that each one-electron orbital contains the spatial coordinates of a single electron as variable. There are $\left( \begin{array}{c} m \\ n \end{array} \right)$ possible ways of orbital selection, that is, the order among orbitals does not make a difference because $\hat{A}$ permutes them anyway. Thus, FCI ansatz $\Psi^{(i)}$ is expanded in the basis of $\left( \begin{array}{c} m \\ n \end{array} \right)$ physically relevant Slater determinants,

$$
\Psi^{(i)} = \sum_{k=1}^{d} C_k^{(i)} \Phi^{(i)}.
$$

(18)

The Slater determinants in Equation (18) are expressed in terms of the basis functions using Equations (16) and (17) as

$$
\Psi^{(i)} = \sum_{k=1}^{d} C_k^{(i)} \hat{A} \left( \prod_{i=1}^{n} \prod_{j=1}^{m} \phi^{(i)}(l) \right).
$$

(19)

By expanding the products of Equation (19), the Slater determinants can be rewritten as a linear combination of $n$-factored products of basis functions, that is,

$$
\Psi^{(i)} = \sum_{k=1}^{d} C_k^{(i)} \sum_{S} W_{S}^{(i)} \prod_{l=1}^{n} \chi_{S}(i),
$$

(20)

where, $S$ denotes a selection scenario. By interchanging the order of summations, the coefficients of basis product $\prod_{l=1}^{n} \chi_{S}(i)$ corresponding to selection $S$ can be merged into a single variable, that is,

$$
W_{S}^{(i)} = \sum_{l=1}^{d} C_k^{(i)} w_{sk}^{(i)}.
$$

(21)

Finally a simple expression for the many-body wave function is written as

$$
\Psi^{(i)} = \sum_{S} W_{S}^{(i)} \prod_{l=1}^{n} \chi_{S}(i).
$$

(22)

In order to prove that the FCI wave functions $\Psi'$ and $\Psi''$ are identical (and hence so are their corresponding electronic energies), it must be shown that $W'$ and $W''$ are equal for all $S$ selections.

As a first step, we show that the number of the linearly independent $S$ scenarios is exactly $\left( \begin{array}{c} m \\ n \end{array} \right)$ due to the antisymmetrization. In particular, if any two (or more) $\chi$ functions within of selection $S$ are identical, the corresponding basis product drops out in the course of antisymmetrization, as the same expression is obtained with opposite sign upon variable exchange. Furthermore, $S$ selections that only differ in the order of $\chi$ functions necessarily have the same weight up to a sign as the antisymmetrization process interconverts them.

FCI defined in Equation (18) optimizes $d$ variational parameters explicitly in order to minimize energy. According to Equation (21), the reference-free expansion of FCI of Equation (22) implicitly varies parameters $W_{S}^{(i)}$ of dimension $\left( \begin{array}{c} m \\ n \end{array} \right)$ which implies the invariance of reference.

As for truncated CI and CC methods or frozen core approximation, the many-body solution is no longer independent of the reference molecular orbital set. The breaking of the invariance is owed to the fact that the dimension of $W_{S}^{(i)}$ equals $\left( \begin{array}{c} m \\ n \end{array} \right)$ independently of the applied computational method whereas the dimension of $C^{(i)}$ decreases for truncated approaches, that is, the summation in Equation (21) runs on a restricted configuration space of dimension $d < \left( \begin{array}{c} m \\ n \end{array} \right)$. This means that $W_{S}^{(i)}$ values become inevitably interdependent, which also prevents them from reaching their optimal values.