Calculation of Metallocene Ionization Potentials via Auxiliary Field Quantum Monte Carlo: Towards Benchmark Quantum Chemistry for Transition Metals

Benjamin Rudsteyn,[†] John L. Weber,[†] Dilek Coskun,[†] Pierre A. Devlaminck,[†]
Shiwei Zhang,^{*,‡,¶} David R. Reichman,^{*,†} James Shee,^{*,§} and Richard A.
Friesner^{*,†}

[†]*Department of Chemistry, Columbia University, 3000 Broadway, New York, NY, 10027*

[‡]*Center for Computational Quantum Physics, Flatiron Institute, 162 5th Avenue, New
York, NY 10010*

[¶]*Department of Physics, College of William and Mary, Williamsburg, VA 23187*

[§]*Department of Chemistry, University of California, Berkeley, CA 94720*

E-mail: szhang@flatironinstitute.org; drr2103@columbia.edu; jshee@berkeley.edu;

raf8@columbia.edu

Abstract

The accurate *ab initio* prediction of ionization energies is essential to understanding the electrochemistry of transition metal complexes in both materials science and biological applications. However, such predictions have been complicated by the scarcity of gas-phase experimental data, the relatively large size of the relevant molecules, and the presence of strong electron correlation effects. In this work, we apply all-electron phase-less auxiliary-field quantum Monte Carlo (ph-AFQMC) utilizing multi-determinant trial wavefunctions to six metallocene complexes to compare the computed adiabatic and vertical ionization energies to experimental results. We find mean averaged errors (MAE) of 1.69 ± 1.02 kcal/mol for the adiabatic energies and 2.85 ± 1.13 kcal/mol for the vertical energies. AFQMC significantly outperforms density functional theory (DFT), which has MAE’s of 1.89 to 11.59 and 3.26 to 15.74 kcal/mol, as well as a localized coupled cluster approach (DLPNO-CCSD(T)), which has MAEs of 4.96 and 6.08 kcal/mol, respectively. We also test the reliability of DLPNO-CCSD(T) and DFT on acetylacetonate (acac) complexes for adiabatic energies measured in the same manner experimentally, and find much higher MAE’s, ranging from 4.24 kcal/mol to 16.11 kcal/mol (with a different ordering) for DFT and 6.97 kcal/mol for DLPNO-CCSD(T), indicating that none of these approaches can be considered benchmark methods for standard organometallic and coordination complexes. We thus demonstrate that ph-AFQMC can handle realistically-sized, redox-relevant transition metal complexes with the advantage of systematically improvable results. By utilizing experimental solvation energies, we show that accurate reduction potentials in solution can be obtained.

Introduction

Quantum chemical methodology has made tremendous progress in both accuracy and computational efficiency during the past three decades.¹ The early 1990’s saw revolutionary improvements in density functional theory (DFT) via gradient corrected and then hybrid formalisms, yielding remarkable reductions in the mean unsigned errors in predicted bond

energies of organic molecules as, for example, assessed using Pople’s G3 database, from 85.27 kcal/mol (LDA) to 4.27 kcal/mol (B3LYP).² In parallel, wavefunction based *ab initio* techniques, in particular the CCSD(T) variant of coupled cluster theory, enabled the attainment of chemical accuracy (~ 1 kcal/mol MAE) for these same data sets, albeit at a much higher computational cost.³

Since these initial breakthroughs, reductions driven by Moore’s law of the cost/performance of computing, coupled with continued progress on theoretical models, algorithms, and software implementations, have greatly expanded the domain of applicability of both the DFT and wavefunction based approaches. Thousands of new DFT functionals have been created and tested, a number of which have demonstrated significant robustness in addressing many of the outlier cases which had plagued PBE,⁴ B3LYP,^{5–7} and related models.^{8–12}

It is now possible to routinely apply DFT calculations to systems containing hundreds to thousands of atoms, including transition metal containing species, and quite often obtain chemically accurate and useful results. The development of localized coupled cluster formulations by a number of research groups (e.g. those of Werner^{13,14} and Neese^{15–20}) has made it possible to routinely perform CCSD(T) computations for systems containing tens to hundreds of atoms; in many cases, the localization approximations have been shown to have a minimal effect on the accuracy that can be achieved. Furthermore, via the use of mixed quantum mechanics/molecular mechanics algorithms, DFT based approaches can be applied to very large and complex systems such as enzymes,²¹ and corrected (if necessary) by CCSD(T) cluster calculations on the reactive core of the system.²²

The accuracy of both DFT and CCSD(T) for transition metals has been a much more controversial topic than for their performance for typical organic systems. CCSD(T) yields highly precise results for transition metal atoms (for example for ionization energies),²³ but the agreement of results with experiment for diatomic molecule bond energies has been the subject of considerable debate.^{24–30} DFT results vary widely depending upon the functional that is used and the specific systems being treated. In many cases, the results for metal

complexes are surprisingly accurate, and at the very least enable considerable insight to be obtained into reaction mechanisms. However, no one has yet rigorously demonstrated, using large and diverse data sets, that any DFT functional achieves reliable performance for transition metal containing systems even at the level of “near-chemical” accuracy (3-4 kcal/mol errors). The problem is in part due to the paucity of high quality gas phase experimental data for transition metal containing systems,^{24,31–40} in contrast to organic molecules where hundreds to thousands of such data points are available for a variety of important thermochemical properties. Additionally, calculations involving solvent and other complicating factors (which typically necessitate the use of heavily parametrized models^{41,42}), or reference reactions, which can be used to take advantage of error cancellation,^{43,44} make it extremely difficult to render an accurate assessment of the performance of DFT based on a small data set of condensed phase experiments. Furthermore, the experiments can be difficult to interpret, an issue compounded by the fact that many transition metal species have a number of close-lying low-energy spin states. For example, assigning ground and vertical state multiplicity in photoelectron spectroscopy can be complicated.⁴⁵

In a series of recent publications, we have made substantial progress in addressing many of the above problems related to transition metal quantum chemistry via the use of auxiliary field quantum Monte Carlo (AFQMC) calculations. The AFQMC methodology, developed originally in the physics community,^{46,47} has a number of potential advantages as compared to traditional wavefunction based *ab initio* methods, including a more favorable formal scaling with system size N , of N^3 (with planewaves⁴⁶) or N^4 (with Gaussian-type orbitals⁴⁸) [vs. N^7 for CCSD(T)⁴⁹], a non-perturbative and multi-reference nature, and the ability to utilize a multiconfigurational SCF trial wavefunction straightforwardly. The ability to use a sophisticated multi-determinant trial is crucial for the treatment of transition-metal-containing systems. Early AFQMC algorithms suffered from a very large prefactor, restricting applications to relatively small systems. Recent technical advances, including vastly improved efficiency for multideterminantal trial wavefunctions,^{23,50} utilization of correlated sampling

(CS) in the Monte Carlo protocol to directly compute energy differences,⁵¹ and implementation on GPU hardware,^{23,52} has made it feasible to treat significantly larger systems.⁵³⁻⁵⁵

These advances have allowed systematic studies of three classes of small transition metal containing species (atoms, diatomic molecules, and tetrahedral, square planar, or octahedral complexes containing simple small molecule ligands) with highly encouraging results. Atomic ionization potentials,²³ diatomic bond energies,²⁹ and complex ligand dissociation energies,⁵⁴ have all been computed with a MAE of less than 1.5 kcal/mol across relatively large experimental gas phase data sets. These results substantially outperformed both localized coupled cluster methods and the best DFT functionals. Furthermore, in all cases, the maximum outlier error was less than 3.5 kcal/mol, in contrast to alternative methods where errors in the 5-10 kcal/mol range were routinely observed.⁵⁴ Improved agreement with precise, state-of-the-art experimental data,⁵⁶ some of which was measured after the calculations were carried out, further validated the robustness of the AFQMC approach. Given the uncertainties of the experiments (and the inherently greater difficulty in making reliable measurements on transition metal containing species), we believe that, to date, AFQMC results are as close to benchmark quality for transition metal chemistry as one is likely to be able to achieve for the near future.

While the data sets enumerated above contained many very challenging electronic structure problems for which the accuracy of coupled cluster methods is expected to be lower than for organic molecules,^{28,30,57} one could argue that the molecular structures that were studied are not representative of those considered relevant by inorganic chemists to biology, catalysis, and materials science. Firstly, the diatomic systems are small and coordinatively unsaturated; one could argue that the failures of DFT for some of these cases are not relevant to practical applications. Secondly, the vast majority of cases involve low oxidation states of the metal which are rarely if ever seen in chemically relevant molecular species. The question then remains: can AFQMC deliver benchmark quality results for more prototypical larger and more complex systems with typical (higher) metal oxidation states?

In the present paper, we study the adiabatic and vertical ionization energies of a series of six first row transition metal metallocenes, in which the metal (V through Ni) is in the II oxidation state, using AFQMC, DFT, and DLPNO-CCSD(T) methodologies. These systems are small enough to enable a large number of computational experiments to be carried out in order to explore which, if any, AFQMC protocols are necessary and sufficient to yield good agreement with experiment. They are also representative molecules for evaluating the expected performance for typical inorganic chemistry applications. Indeed, ferrocene oxidation is often used as a reference reaction in electrochemical measurement of redox potentials.⁵⁸ Finally, adiabatic gas phase experimental data, measured with electron transfer equilibrium (ETE) by Richardson and co-workers⁵⁹⁻⁶¹ with relatively low experimental uncertainties, as well as vertical gas phase experimental data, as measured with photoelectron spectroscopy by Green and co-workers,^{45,62} exist for this series, which can form the basis to evaluate the accuracy of the various quantum chemical approaches for prototypical organometallic species.

A number of significant conclusions emerge from the metallocene calculations presented here. First, we see again that the localized coupled cluster approach that we have employed (DLPNO-CCSD(T)) performs poorly, with a number of large outliers and an overall MAE that is inferior to a wide range of DFT functionals (with a large gap between this approach and the top DFT performers). Second, and again in agreement with our previous series of papers, the B3LYP functional performs well for several of the molecules in our test set, but displays serious errors for others. The same is true of a related functional, B3LYP*, which incorporates a smaller percentage of Hartree-Fock exchange, which ostensibly improves performance for transition metals.⁶³ This data contradicts the claims that have been made over the years that B3LYP* can reliably provide near chemical accuracy (typically 3-4 kcal/mol)⁶⁴ for realistic transition metal containing systems.⁶⁵ Thirdly, the PBE0,¹⁰ ω B97M-V,⁶⁶ M06,⁶⁷ and ω B97X-V⁶⁸ functionals display a level of accuracy which is quite impressive for the metallocene data set, with MAEs of 1.89, 2.07, 2.09, and 2.35 kcal/mol, respectively. However,

there are still a number of outliers that remain outside of the bounds of near chemical accuracy. Additionally, these functionals do not do as well for vertical energies with MAE’s of 5.60, 5.66, 4.57, and 5.52 kcal/mol respectively, showing the difficulty that these functionals would have with non-relaxed structures.

Motivated by these results, we also tested these DFT functionals and DLPNO-CCSD(T) on a second set of gas phase ionization energies measured by the same group,⁶⁹ for triacetylacetonate (acac) systems (V through Co). Here, we find two molecules, namely $[\text{Cr}(\text{acac})_3]^{1-}$ and $[\text{Mn}(\text{acac})_3]^{1-}$, for which very large outliers are obtained, confirming the initial picture that one can often obtain quite accurate results, but that major failures occur regularly as well. In other words, these functionals cannot be considered to be “blackbox” benchmark approaches for transition metal thermochemistry. We do not believe that these errors are due to issues with the experiments, as will be discussed in more detail below.

Finally, we are able to develop a robust AFQMC protocol which provides accurate results for the metallocene series, considering the level of the experimental uncertainties. This is achieved by overcoming significant challenges arising from the greater size and complexity of the metallocene series (as compared to molecules in earlier publications).^{23,29,54,55,70} Prior work⁵⁴ has demonstrated that it is essential to employ an appropriate multiconfigurational trial wavefunction; calculations based on a single determinant (mainly UHF) trial sometimes yield energy differences comparable to DFT for these systems. Deployment of an appropriate multiconfigurational trial function in the application of the AFQMC approach to transition metal containing systems remains essential if one is aiming at robust, benchmark quality results.

The second key component of our protocol is a novel synthesis of the CS simulation algorithm with the more standard approach in which two independent MC simulations are employed. In previous papers, we have referred to this second approach as one dependent upon “population control” (PC), and we continue to refer to such simulations as PC calculations to maintain continuity with prior work. When MC sampling is carried out independently

for two different states, lengthy imaginary time simulations are required, and PC along with the phaseless constraint are needed to control the phase problem which would otherwise make it difficult to converge the Monte Carlo integration. A CS simulation typically can be carried out for a much shorter duration, during which the need for population control of AFQMC walkers is minimized and the accrual of phase constraint error is mitigated. For this approach to be justified, it is necessary that the imaginary time runs produce a “quasi-plateau” in the targeted energy difference for relatively short imaginary times which better approximates the unbiased result. In the cases we have investigated, we empirically note such a stable regime in imaginary time. How robust this type of behavior is when considering a wider class of systems remains an open question worthy of more systematic future investigation.

We have shown in previous work that the use of CS is highly effective in obtaining accurate energy differences between different electronic surfaces for a number of very challenging cases, and can in fact produce more accurate results than pure PC in cases we have investigated.^{29,54} However, CS calculations are most effective if the geometry of the two surfaces is held constant. Previously, we computed the relevant quantities for adiabatic ionization potentials (energy differences between the potential minima on each surface) by using DFT to evaluate the reorganization energy in going from the vertical excitation geometry on the final surface to the energy minimum on that surface.⁵⁴ This approximation, however, produces very poor results for one case (manganocene), as we will discuss further below, as there is a large change in equilibrium geometry between the II and III states. Evaluating the reorganization energy with AFQMC using PC for the two geometries remedies this problem, and yields highly satisfactory agreement with experiment for all cases. This combination of the CS and PC versions of AFQMC calculations (CS to measure energy differences for vertical transitions, and PC to measure differences between two geometries on the same surface) appears to be a very promising approach that can be used in an automated fashion going forward. Of course, considerable additional comparison with experiment will be required in order to rigorously

assess errors across a wide range of relevant transition metal containing systems.

This paper is organized as follows. In Section II, we discuss the experimental data for the metallocene and acac series that we will be focusing on in our computational work. In Section III, we briefly review the AFQMC methodology and the previous results obtained using it, and describe the DLPNO-CCSD(T) and DFT methods employed. Section IV presents results of AFQMC, various DFT functionals, and DLPNO-CCSD(T) calculations for the metallocene series, as well as DFT and DLPNO-CCSD(T) results for the acac series, which will be studied by AFQMC in future work. We also show that using experimentally derived solvation free energies in concert with accurate gas-phase predictions leads to accurate solution-phase reduction potentials. In Section V, we consider the implications of our results for the utility of DFT, DLPNO-CCSD(T), and AFQMC in addressing transition metal chemistry. Finally, in Section VI we conclude with a summary of our results and outline future directions.

II. Experimental Data Sets

The gas-phase ionization energy experiments that we investigate below center on first row transition metals in the +2 or +3 oxidation state, with either anionic cyclopentadienyl (Cp) ligands (metallocenes) or acetylacetonate ligands (acac series), depicted schematically in Figs. 1a and 1b, respectively.

The metallocenes we investigate are vanadocene, chromacene, manganocene, ferrocene, cobaltocene, and nickelocene. The ionization processes we study are for the II oxidation state (charge = 0) to the III oxidation state (charge = +1). All of the metallocene molecules in these two oxidation states are low spin complexes, except for Mn(II), which is a sextet, and Mn(III), which is a quintet for the vertical experiments (as discussed below). The acac species in the II oxidation state (charge = -1) are ionized to form the III oxidation state (charge = 0). All of the acac molecules in these two oxidation states are high spin complexes,

except for Co(III) which is a singlet. The ground state multiplicities of all species are given in Table 1.^{41,62}

Table 1: The ground state multiplicities of the metallocene or acac complexes in III or II oxidation states.

Oxidation State	III	II	III	II
Ligand	Cp	Cp	acac	acac
Metal				
V	3	4	3	4
Cr	4	3	4	5
Mn	3	6	5	6
Fe	2	1	6	5
Co	1	2	1	4
Ni	2	3	-	-

The adiabatic experimental values come from the electron transfer equilibrium (ETE) measurements of Richardson and co-workers.^{59–61} These experiments utilize Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) to determine adiabatic ionization energetics for organic and inorganic species near room temperature. From the measured equilibrium constants for the electron-transfer reactions, the free energies of reaction are determined, and from this information the free energies of ionization can be inferred. The energies in solution are obtained using the solution phase potentials and the 4.44 V absolute potential of the standard hydrogen electrode (SHE) in water at 298.15 K, derived using various thermodynamic quantities, such as the solvation energy of a proton.⁷¹ Because these experiments were performed at 350 K, we use this temperature value in our calculated free energy corrections.

Figure 1: The structure of the metal complexes studied.

The vertical experimental values come from the photoelectron spectroscopy (PES) measurements of Green et al.,^{45,62} who used UV and X-ray photons to ionize samples via the photoelectric effect and measure the kinetic energy (KE) of the ejected photon with a frequency ν . The binding energy (BE) is then determined by the equation $BE = h\nu - KE$, where h is Planck’s constant.

III. Computational Details

The geometries, certain reorganization energies (*vide infra*), and ideal gas free energy corrections⁷² were obtained with the B3LYP functional⁵⁻⁷ utilizing the cc-pVTZ-DKH⁷³⁻⁷⁶ basis set and DKH2 relativistic corrections.⁷⁷ These calculations were performed using the ORCA program package.⁷⁸ In the cc-pVTZ-DKH basis set, metallocenes typically have around 508 basis functions and 95 electrons. The acac complexes are roughly twice as large from an electronic structure perspective with 992 basis functions and 183 electrons. Information regarding integration grids and other theoretical details can be found in the SI.

Unrestricted DLPNO-CCSD(T_0) (which we refer to as DLPNO-CCSD(T); T_0 refers to the semi-canonical approximation)¹⁸ calculations were performed with quasi-restricted orbitals (QROs) generated from unrestricted B3LYP reference orbitals and “NormalPNO” (moderate energy cutoff criteria for correlation between localized orbitals) localization parameters using ORCA. These calculations utilized cc-pV x Z-DKH and auxiliary cc-pV x Z\C

basis sets, where x is the cardinal number of the basis set (i.e. $x=3,4$ for TZ,QZ), as built into ORCA,⁷⁸ as discussed in the SI. We extrapolate to the complete basis set (CBS) limit for absolute energies using exponential and $\frac{1}{x^3}$ dependence for the HF and correlation energies, respectively,⁷⁹ The DLPNO procedure automatically freezes the core orbitals.

Electron repulsion integrals and the Hamiltonian for ph-AFQMC were obtained with PySCF.⁸⁰ The exact-two-component (x2c) relativistic Hamiltonian⁸¹ was used in place of DKH2, as the latter is not implemented in PySCF.

The ph-AFQMC propagation utilized an imaginary time step of $0.005 E_h^{-1}$ (these units are also referred to Ha^{-1} or β in the literature), which in our experience is sufficiently small such that errors from the Trotter decomposition are negligible given our target statistical accuracy.⁵¹ We utilize single precision (sp) rather than double precision (dp) floating point arithmetic, as discussed in the SI. Walkers were either initialized with a RHF/ROHF determinant or according to a distribution of CASSCF determinants weighted by their respective CI coefficients. Initialization with restricted orbitals ensures spin-purity, even in the case of a UHF trial wavefunction.⁸² Additional details are given in the SI.

For the main ph-AFQMC computations using the cc-pVTZ-DKH basis set, we utilized CASSCF trial wavefunctions. The default active space for generating the trial wavefunction was automatically determined via the atomic valence active space (AVAS) procedure, where only those orbitals that overlap significantly ($\sim 10\%$) with the 3d and 4d orbitals (as defined from the Atomic Natural Orbital (ANO-RCC) basis set)⁸³ of the metal ion were included.⁸⁴ This active space thus targets the static correlation of the metal rather than the ligands. These active spaces were typically around 14 electrons in 15 orbitals, typically with 98% of the CI weight retained (the minimum was 96%), resulting in about 300 determinants. RCAS/AFQMC calculations were determined to be converged with respect to the active space size by testing active spaces of increasing size until the resulting calculations were equivalent within statistical accuracy. If the natural orbital occupation numbers (NOONs) resulting from this approach were not physical (e.g. the fractional change in occupation in

the occupied orbitals is not reflected in the virtual orbitals), alternate active spaces were selected using other approaches, such as using the frontier orbitals without modification, or using the MINAO basis set with 3d and/or 2p_z orbitals, as described in the SI.

The CBS limit for the ph-AFQMC calculations was estimated using an approach similar to that described in previous work.⁵⁴ Briefly, we extrapolate the ionization energy computed with ph-AFQMC PC with an inexpensive trial wavefunction such as UHF or a CASSCF wavefunction with a small active space with the cc-pVxZ-dkh basis sets (x=3,4 for T,Q). The UHF ionization energy is extrapolated using an exponential form. The contribution to the ionization energy from the correlation energy computed by AFQMC is extrapolated using a $1/x^3$ functional form. This method is equivalent to fitting the procedure discussed above for the CBS extrapolation of DLPNO-CCSD(T), though there we extrapolate the *absolute* rather than the *relative* energy. This "low-level" result is used, in turn, to extrapolate the "high-level" ionization energy computed with a large CAS trial in the $x=3$ basis. A scaling factor, ρ the ratio of the correlation energies between low and high levels of AFQMC is used to translate the basis dependence of the least sophisticated trial to a result that approximates one with a better trial function,

$$\rho = \frac{\Delta E^{corr}(\text{TZ, high-level})}{\Delta E^{corr}(\text{TZ, low-level})}. \quad (1)$$

Whether or not a CAS trial is used for the extrapolation is determined by identifying which cases appear to exhibit significant multireference character, as can be flagged by deviations of the CASSCF NOONs and $\langle S^2 \rangle_{B3LYP}$ from ideal values. When needed, we run separate ph-AFQMC calculations on the optimized structures of both states of interest. Our independent sampling (as opposed to CS) runs apply population control (PC), in which walkers with large weights are duplicated while those with small weights are destroyed with appropriate probability via a "comb" algorithm.⁴⁸

Our correlated sampling approach⁵¹ (CS) enables fast convergence of vertical energy dif-

ferences between similar states of a system, e.g. reduced/oxidized states, by employing a shared set of auxiliary fields, effectively leading to a cancellation of statistical error and, in many cases, also fast (quasi-)equilibration. The lack of PC in this scheme leads to exponentially increasing statistical noise as a function of projection time, due to the increasing fluctuations in weights of the walkers. CS thus require more walkers, but faster equilibration allows for a much shorter propagation time. Empirically, we find that $15 E_h^{-1}$ allows full equilibration of the accuracy differences while providing excellent statistical accuracy.

CS has been shown to improve the accuracy of the calculated energy differences in certain situations, due to correlated and faster convergence of the energy differences, which avoids the full onset of the phaseless constraint error.^{23,29,54} For MnCp₂, we also checked that the use of an alternate, more rigorous approach to CS produces the same ionization energy as the original algorithm (details provided in the SI, and in a future publication). Results across three different types of data sets^{29,51,54} suggest that the CS methodology achieves a greater reliability than PC approaches in comparing different electronic surfaces to within near chemical accuracy.

The ionization energy (IE), is computed by ph-AFQMC with CS by

$$IE = G(III) - G(II) - \lambda, \tag{2}$$

where G is the free energy obtained by correcting the electronic energy with terms that account for temperature effects. The reorganization energy, λ , is defined as the difference in energy between the III product in its optimal geometry and in the II reactant geometry. λ is computed via either B3LYP/cc-pVTZ-DKH or the PC ph-AFQMC with either a UHF or CAS trial wavefunction, as described in the Results section. The calculation of IEs is illustrated in Fig. 2.

Figure 2: Schematic of the ionization and reorganization energy calculations performed in this work. CS and PC indicate the energy was measured by the CS and PC approaches, respectively. The v labels for vibrational energy levels are arbitrary.

For our CS calculations, we typically ran 30 repeats, with 6624 walkers ($\simeq 200,000$ walkers in total) and 276 GPUs (46 nodes) each until $15 E_h^{-1}$. For our PC calculations, we typically ran for $2000 E_h^{-1}$ (or shorter for QZ calculations) using 3312 walkers and 552 GPUs (92 nodes). To give a sense of the required computational cost, a CS ph-AFQMC calculation for $[Mn^{III}Cp_2]^0$ requires about 1,231 node hours on Summit (< 1 hour walltime), using a truncated (99.5% of the weight) CASSCF trial wave function containing 282 determinants. These settings typically allowed us to obtain statistical error bars in the energy difference below 2 kcal/mol.

We compare ph-AFQMC with the local GGA BP86,^{85,86} hybrids B3LYP,⁵ B3LYP*,⁶³ PBE0,¹⁰ and B97,^{8,9} the meta-GGA hybrids M06⁶⁷ and TPSSh,^{87,88} the semilocal meta-GGA (non-hybrid) B97M-V,⁸⁹ the range-separated hybrid meta-GGA ω B97M-V,⁶⁶ the range-separated hybrid ω B97X-V functionals,⁶⁸ and the double-hybrid DSD-PBEB86.^{90,91} These calculations were done in ORCA. We perform single-point energy calculations on the B3LYP geometries using the large “Grid7” option in ORCA (see SI).

IV. Results

Vertical Ionization Energies

The vertical IE results using wavefunction methods are given in Table 2. The AFQMC PC/UHF result is a significant improvement on average as compared to DLPNO-CCSD(T), though with large outlier cases such as CrCp₂ and NiCp₂. It is clear that DLPNO-CCSD(T) only performs well for FeCp₂ and CoCp₂. Both AFQMC PC and AFQMC CS with CAS trials are significant improvements over AFQMC PC/UHF. The CS results have a lower maximum error. The PC results show a greater dependence on the quality of the trial wavefunction.

The CASSCF NOONs of both ⁵Mn(III) and Ni(III) show a small but notable fractional occupation of the lowest unoccupied natural orbital (LUNO) of at least +0.165, which can be traced to a loss of occupation of at least -0.16 in one of the occupied orbitals. In both cases, these deviations from ideal NOONs are accompanied by a deviation of $\langle S^2 \rangle_{B3LYP}$ from the expected value. The Mn and Ni cases deviate from the ideal values by 5.54% and 6.38%, respectively. Non-integer CASSCF NOONs and spin-symmetry breaking in unrestricted DFT calculations have been put forth as complementary diagnostics of static correlation in transition metal compounds, as they reflect a wavefunction that is a superposition of more than one spin state made possible by a near-degeneracy of energy levels.⁹² Indeed, the calculated M-Cp lengths (Table S29) and experimental gas-phase homolytic M-Cp dissociation energies (Table S30)^{60,61} together imply that bonding is weakest in the Mn and Ni complexes across the 3d series. In general, as the strength of a bond weakens, the energy splitting between spin states narrows, setting the stage for static correlation. While most simply illustrated when diatomics such as H₂ or N₂ are stretched, an analogous situation has previously been reported for a weakly bound tetraammine Mn cation in the gas phase.⁹² While the degree of multireference character implied by these methods can be sensitive to the active space and DFT functional employed, respectively, we propose (and certainly find in this dataset) that this procedure has utility for pinpointing particularly difficult cases, independent from any

comparison to experiment.

Due to the presence of particularly extensive correlation in the Mn and Ni cases, we apply an improved extrapolation, using AFQMC PC/CAS rather than with AFQMC PC/UHF. We used AVAS and the MINAO basis set to generate active spaces of 8-10 electrons in 8-10 orbitals for $\text{Mn}^{II}\text{Cp}_2$, $^5\text{Mn}^{III}\text{Cp}_2$, and all three NiCp_2 species in both TZ and QZ basis sets for use in AFQMC. Using these trial wavefunctions instead of the UHF trial wavefunctions to extrapolate the vertical AFQMC CS/CAS value resulted in a value of 153.14 ± 1.72 kcal/mol for the Mn case, which agrees reasonably well with the experimental value of 150.1 ± 2.31 kcal/mol, considering the uncertainties in both values. The value with a UHF trial extrapolation is 157.08 ± 1.72 kcal/mol, which is clearly in worse agreement with experiment and outside of the joint error bars of theory and experiment. The Ni case was also significantly improved. Finally, as a control, we ran calculations using the CASSCF-trial CBS extrapolation for CoCp_2 , which did not require it according to our criteria, and obtained very similar results to those found with the use of the UHF trial (see SI).

The accuracy of these vertical excitation results implies that, coupled with accurate calculation of the reorganization energy, we should find accurate adiabatic results. Indeed, the photoelectron spectroscopy results may be more difficult to interpret than the adiabatic experiments, due to a variety of factors. For example, we do not attempt to compute vibronic contributions or to include temperature effects. Additionally, the vertical experiments exhibit an increased uncertainty due to the difficulty in interpreting the spectra in terms of line width. Moreover, the “excited” vertically ionized state, in a distorted, nonequilibrium geometry, is naturally harder for electronic structure methods to compute as compared to the equilibrium geometry of the ground state.

The DFT results for vertical ionization energies, using the various functionals enumerated in section III, are given in Fig. 3 and Table 3.

The hybrid functionals B3LYP B3LYP*, M06, B97, PBE0, TPSSh, $\omega\text{B97X-V}$ and $\omega\text{B97M-V}$, as well as the meta-GGA B97M-V all have MAE’s between 4.5 and 5.7 kcal/mol indicating

Table 2: Vertical ionization energies as a function of metallocene type and wavefunction methodology. The mean absolute errors (MAE), maximum errors (MaxE), root-mean-square deviations (RMSD), and the mean signed errors (MSE) are included. All units are in kcal/mol.

	PES Expt	DLPNO- CCSD(T)	AFQMC PC/UHF	AFQMC PC/CAS	AFQMC CS/CAS
$V(Cp)_2$	156.3 ± 2.31	162.93	158.95 ± 1.59	161.4 ± 2.16	155.34 ± 1.51
$Cr(Cp)_2$	131.4 ± 2.31	122.70	123.49 ± 1.31	129.29 ± 1.86	127.27 ± 1.85
$Mn(Cp)_2$	159.3 ± 2.31	168.50	159.29 ± 0.72	158.51 ± 1.00	156.56 ± 0.78
$Fe(Cp)_2$	158.7 ± 2.31	154.12	161.09 ± 1.92	161.1 ± 2.3	155.38 ± 1.88
$Co(Cp)_2$	128 ± 2.31	128.60	126.23 ± 1.87	127.5 ± 1.89	129.22 ± 1.44
$Ni(Cp)_2$	150.1 ± 2.31	156.89	157.89 ± 1.33	153.19 ± 0.97	153.14 ± 1.72
MAE		6.08	3.75 ± 1.13	2.33 ± 1.13	2.57 ± 1.13
$ MaxE $		9.20	7.91 ± 1.13	5.1 ± 1.13	4.13 ± 1.13
RMSD		6.73	4.82	2.79	2.81
MSE		1.66	0.53 ± 1.13	1.2 ± 1.13	-1.15 ± 1.13

similar performance, given the uncertainty of the experiments. The double hybrid functional DSD-PBEP86 also does not perform very well, especially for $CoCp_2$, which may be due to the difficulty that MP2-based methods have for organometallic complexes, as discussed in Ref. 92. The lower RMSD observed for the BP86 functional must be considered fortuitous in view of the very large errors obtained for the adiabatic calculations in Table 6 below. Among different functionals, the maximum error often occurs at different metallocenes, indicating a lack of predictable reliability.

Figure 3: Performance of ph-AFQMC (CAS trial, CS), DFT functionals (QZ basis), DLPNO-CCSD(T), and experiment for prediction of experimental vertical gas-phase ionization energies for metallocenes. The range is limited to deviations of -15 to 15 kcal/mol. The DSD-PBEP86 result for $\text{Fe}(\text{Cp})_2$ and $\text{Ni}(\text{Cp})_2$ are out of range with deviations of about -30 and 17 kcal/mol respectively. The gray band indicates the uncertainty of the experiments i.e. 2.31 kcal/mol.

Table 3: DFT (QZ basis) vertical ionization energies, as a function of metallocene and methodology. The mean absolute errors (MAE), maximum errors (MaxE), root-mean-square deviations (RMSD), and the mean signed errors (MSE) are included. All units are in kcal/mol.

	Expt.	BP86	B3LYP	B3LYP*	PBE0	TPSSh	M06	B97	B97M-V	ω B97X-V	ω B97M-V	DSD-PBEP86
$V(Cp)_2$	156.3 ± 2.31	157.11	152.01	151.88	155.18	151.77	153.57	151.55	153.03	154.25	155.69	153.41
$Cr(Cp)_2$	131.4 ± 2.31	132.26	124.60	125.84	123.06	125.77	121.16	123.79	124.00	126.15	125.69	124.04
$Mn(Cp)_2$	159.3 ± 2.31	148.26	152.81	150.65	157.07	149.44	157.80	153.00	152.18	161.81	162.51	166.63
$Fe(Cp)_2$	158.7 ± 2.31	159.50	149.05	150.79	149.04	150.61	153.55	149.47	150.20	151.55	151.08	131.33
$Co(Cp)_2$	128 ± 2.31	125.36	128.14	126.09	132.58	125.42	130.79	127.66	125.31	133.45	133.70	95.67
$Ni(Cp)_2$	150.1 ± 2.31	146.67	152.20	149.39	157.77	148.90	155.09	152.19	150.23	160.83	161.20	167.28
MAE		3.26	4.91	4.86	5.60	5.31	4.57	5.05	4.85	5.52	5.66	15.74
$ MaxE $		11.04	9.65	8.65	9.66	9.86	10.24	9.23	8.50	10.73	11.10	32.33
RMSD		4.88	5.84	5.66	6.44	6.10	5.38	5.91	5.71	6.25	6.55	19.17
MSE		-2.44	-4.17	-4.86	-1.52	-5.31	-1.97	-4.36	-4.81	0.71	1.01	-7.57

Reorganization Energies

Table 4 gives the reorganization energies along the III potential energy surface (except for MnCp_2 where we use the low spin surface) for the various metallocenes using B3LYP, AFQMC with PC and a UHF trial (AFQMC PC/UHF), and AFQMC with PC and a CAS trial (AFQMC PC/CAS). We see that B3LYP reorganization energies are outside AFQMC error bars in all cases. The reorganization energies are fairly similar for VCp_2 , CrCp_2 , and FeCp_2 . The AFQMC results disagree with B3LYP for CoCp_2 and the AFQMC/CAS result shows significant differences the other two for NiCp_2 .

The discrepancy between all methods is very large for MnCp_2 . The reorganization energy for the Mn system is expected to be large given the large geometry change in going from II to III.⁶¹ In particular, the B3LYP M-Cp ring centroid distance decreased from 2.08 Å to 1.78 Å. To explore this further, we rigidly changed the Mn^{III} -Cp centroid distance and calculated the energy along this coordinate (optimizing other degrees of freedom) using B3LYP and AFQMC/UHF. Fig 4 shows that the equilibrium position from both methods are similar, but at higher distances, where the Mn^{II} geometry would be found, the PES curves differ significantly, with the AFQMC/UHF curve well above the B3LYP curve, indicating its reorganization energy will be higher, as we observe. We use different AFQMC reorganization energies to calculate adiabatic ionization energies and compare the results to experiment in the next section.

Table 4: Reorganization energies calculated in the TZ basis along the III potential energy surface as a function of metallocene and methodology. All units are in kcal/mol.

	B3LYP	AFQMC PC/UHF	AFQMC PC/CAS
$V(\text{Cp})_2$	0.37	1.73 ± 0.58	1.76 ± 0.52
$\text{Cr}(\text{Cp})_2$	3.04	0.14 ± 0.5	-0.19 ± 0.52
$\text{Mn}(\text{Cp})_2$	28.46	52.38 ± 0.7	41.5 ± 0.68
$\text{Fe}(\text{Cp})_2$	3.17	0.82 ± 0.71	0.87 ± 0.69
$\text{Co}(\text{Cp})_2$	6.57	12.45 ± 0.79	10.75 ± 0.66
$\text{Ni}(\text{Cp})_2$	6.30	7.61 ± 0.64	11.04 ± 0.57

Figure 4: Rigid scans of the Cp-Cp distance in $[Mn^{III}Cp_2]^{1+}$ where red = B3LYP and black = ph-AFQMC/UHF. The x-axis is the deviation from the B3LYP optimal Cp to Cp ring distance of 3.55 Å. The dashed vertical lines indicate what the reorganization energy would be at the optimal ring distance of $[Mn^{II}Cp_2]^0$.

Adiabatic Ionization Energies

The adiabatic ionization energies using different AFQMC methods for the reorganization energy for the CS results for are compared to equivalent results from DLPNO-CCSD(T) as well as experiment in Table 5. With AFQMC CS, the adiabatic ionization energy is computed with the two-step procedure illustrated in Fig. 2, while with DLPNO-CCSD(T) and AFQMC PC, it is computed directly as a two-point energy difference (purple line). It is immediately apparent that DLPNO-CCSD(T) as well as AFQMC using PC with either a UHF trial or a CAS trial does not perform well, with similar MAE's, absolute maximum errors, and other statistical measures. Much of the poor performance is for $MnCp_2$. The poor performance of AFQMC using PC is likely due to an inferior trial wavefunction, which

is perhaps not converged with respect to active space, particularly for the III oxidation state, and potentially poor error cancellation. The results using AFQMC with CS, CAS trial, and B3LYP reorganization energies, called AFQMC CS (1), performs similarly poorly due to this outlier. However, using AFQMC-generated reorganization energies results in much better agreement with experiment. We interpret this success as follows: The vertical excitation from the II to the III state has the largest change in electronic structure and hence is most demanding; CS succeeds because the energy difference can be converged before the full bias due to the phaseless constraint appears, and exploits cancellation of error in the Monte Carlo sampling on the II and III surfaces. For the reorganization energy, DFT methods appear to have difficulty obtaining accurate results for these transition metal containing systems at geometries that are substantially distorted from the minimum; AFQMC/PC provides accurate results for such distortions, at least for the present systems.

Table 5: Wavefunction adiabatic ionization energies at the complete basis set limit as a function of metallocene and methodology. The mean absolute errors (MAE), maximum errors (MaxE), root-mean-square deviations (RMSD), and the mean signed errors (MSE) are included. All units are in kcal/mol. (1) = B3LYP reorganization energy and (2) = AFQMC PC/CAS reorganization energy. All units are in kcal/mol.

	ETE Expt	DLPNO- CCSD(T)	AFQMC PC UHF Trial	AFQMC PC CAS Trial	AFQMC CS (1) CAS Trial	AFQMC CS (2) CAS Trial
$V(Cp)_2$	154.5 ± 1.5	161.29	157.64 ± 1.55	160.25 ± 2.15	155.44 ± 1.61	153.7 ± 1.52
$Cr(Cp)_2$	127.5 ± 1.5	123.81	127.04 ± 1.43	133.54 ± 1.96	127.63 ± 1.81	131.18 ± 2.00
$Mn(Cp)_2$	142.5 ± 1.5	153.56	156.29 ± 1.54	151.72 ± 2.00	158.56 ± 1.97	146.12 ± 2.37
$Fe(Cp)_2$	153.1 ± 1.5	156.19	158.41 ± 1.84	158.37 ± 2.32	150.34 ± 1.75	152.78 ± 2.02
$Co(Cp)_2$	123.5 ± 1.5	122.43	119.08 ± 1.6	122.2 ± 1.85	128.12 ± 1.52	124.37 ± 1.83
$Ni(Cp)_2$	143.8 ± 1.5	147.83	145.14 ± 1.42	143.88 ± 1.64	148.64 ± 1.7	144.64 ± 2.15
MAE		4.96	4.75 ± 0.89	4.61 ± 1.02	4.89 ± 0.93	1.69 ± 1.02
$ MaxE $		11.06	13.79 ± 2.15	9.22 ± 2.5	16.06 ± 2.47	3.68 ± 2.5
RMSD		5.90	6.45	5.54	7.20	2.19
MSE		-3.37	3.12 ± 0.89	4.18 ± 1.02	3.97 ± 0.93	1.31 ± 1.02

DFT results for the metallocenes using a variety of functionals are given in Table 6 using the QZ basis set, as is recommended in Ref. 93. It is clear that the local GGA BP86 functional performs very poorly, particularly for $MnCp_2$, as opposed to its performance for the vertical energies, while B3LYP performs much better. B3LYP*, which is often claimed to perform

better than B3LYP due to a slightly smaller admixture of HF exchange (15% vs. 20%),⁶³ in fact performs more poorly in this case, particularly for MnCp_2 . The alternative hybrid functionals PBE0 (25% HF exchange) is actually the most accurate, and B97 fares slightly worse than does B3LYP. The meta-GGA hybrid M06 (27% HF exchange) also performs very well, as opposed to the similar TPSSh (10% HF exchange), which fails catastrophically for MnCp_2 . It does not appear that M06 is trained on metallocene reference data,⁶⁷ though it was tested on metallocene M-Cp dissociation energies (not ionization energies).^{94,95} Even if it had been, the metallocene experiments would not be a large enough set to evaluate whether parameterization on metallocenes is responsible for M06’s success. The semilocal meta-GGA functional B97M-V performs similarly to TPSSh. The two functionals, ω B97M-V (15% short range HF exchange) and ω B97X-V (16.7% short range HF exchange), are an improvement over B97 and B97M-V with the latter performing slightly better than the former. Finally the double-hybrid DSD-PBEP86 (32 % HF exchange), performs poorly as well, particularly for CoCp_2 , reaching a maximum error of almost 30 kcal/mol. Additionally, spin contamination was particularly severe with this functional as shown in the SI, which might have been expected given the relatively high level of HF exchange in this functional.

There is not enough data to declare a “best” functional or to note a trend of MAE vs. % HF exchange, especially considering this is not the only difference between functionals except in the case of the comparison between B3LYP and B3LYP*. Nonetheless, it is clear that hybrid functionals are needed to perform well on metallocenes. The performance for adiabatic IEs in general is better than for vertical energies. All the results using a selection of functionals and ph-AFQMC data are summarized in Fig. 5.

If we consider the adiabatic metallocene results for the two DFT functionals which perform best on large scale quantum chemistry benchmark sets excluding transition metals, namely, the range-separated hybrids ω B97X-V and ω B97M-V,⁹³ we can make two interesting observations. Firstly, the MAEs for both functionals, (2.35 and 2.07 kcal/mol, respectively, vs. 1.69 ± 1.02 kcal/mol for AFQMC), are quite respectable. Secondly, both functionals

Figure 5: Performance of ph-AFQMC, DFT functionals (QZ basis), and DLPNO-CCSD(T) for prediction of experimental adiabatic gas-phase ionization energies for metallocenes. The range is limited to deviations of -15 to 15 kcal/mol. The BP86, TPSSh, and B97M-V values for $\text{Mn}(\text{Cp})_2$ are out of the range of the plot with deviations of about -25, -20, and -15 kcal/mol, respectively. The DSD-PBEP86 results for $\text{Fe}(\text{Cp})_2$ and $\text{Co}(\text{Cp})_2$ involve a deviation of about -20 and -40 kcal/mol, respectively, so they are also beyond the range of the plot. The gray band indicates the uncertainty of the experiments i.e. 1.5 kcal/mol.

have a number of significant errors, greater than 4 kcal/mol, but these errors are not so large as to definitively rule out a major contribution from experimental noise. If performance at this level were to be observed for a wide range of transition metal containing systems, these functionals would in fact be in the range of performance reported for “hard” thermochemistry cases in Ref. 93 on standard DFT benchmarking sets. Resolving this ambiguity requires locating other transition metal cases where rigorous comparisons with experiment and/or benchmark quantum chemical results are possible. For this reason, we decided to carry out preliminary DFT and DLPNO-CCSD(T) calculations for the acac series of organometallic complexes described above. These systems have nearly double the number of electrons as the metallocenes, which presented difficulties with regard to obtaining results using our current AFQMC code. We have recently made major improvements to the code’s performance and scaling with system size, and so it is likely we will be able to report converged results for the acac series in the near future. Even without AFQMC data, the DFT results can provide a

preliminary assessment as to how robust the best performing DFT functionals in Table 6 are when extended to a new series of complexes. Results for our entire suite of DFT functionals, along with those for DLPNO-CCSD(T), are presented in Table 7 and Fig. 6 below. In what follows, we focus attention on adiabatic ionization potentials, as we have not evaluated vertical ionization energies for the acac series. We use the same QZ basis as was employed in studying the metallocenes. Results utilizing diffuse functions in the basis set, which one might argue could be significant because the II state in the acac series is an anion, yield very similar results, as is discussed in the SI.

A number of striking features of the data are immediately apparent. Firstly, none of the DFT methods performs nearly as well as the best performers do for the metallocene series; the MAE and RMSD are in all cases far from near-chemical accuracy. This performance may in part be due to the net negative charge residing on the molecule, which can delocalize onto the metal in the III state, resulting in overbinding (as has been observed in organic systems). It should also be noted that the two acac cases with anomalously large errors (Cr and Mn) are the only ones which involve ionization from singly occupied e_g orbitals. Previous work has shown that the errors in removing or adding electrons to DFT orbitals depends significantly upon whether the orbital is singly or doubly occupied, and what type of orbital is involved.^{41,42,96}

The double hybrid functional DSD-PBEP86 yields the best results, at 4.24 kcal/mol MAE, but it must be noted that this functional is by far the worst performing for the metallocenes, with a MAE of 11.59 kcal/mol. Across both data sets, all of the DFT functionals exhibit multiple failures with quite large errors. Secondly, the DLPNO-CCSD(T) results are no better than the DFT results for, e.g., the range-separated hybrids. Thirdly, some of the DFT error appears to be systematic in character, with similar trends being manifested for many of the DFT functionals. A particularly extreme example can be found for the $[\text{Cr}(\text{acac})_3]^{1-}$ species, for which most of the DFT functionals yield an adiabatic ionization potential that is ~ 15 kcal/mol smaller than experiment. Assuming that the acac experi-

ments have error bars that are similar to those of the metallocene experiments (the latter having been validated by the close agreement of the experimental data with the AFQMC results), we can conclude from the above results that both DFT and DLPNO-CCSD(T) cannot reliably produce benchmark level thermochemical data for transition metal containing systems. It would be surprising if the acac experiments, which were carried out by the same group, using the same apparatus and protocols, as for the metallocene experiments, were qualitatively less accurate than the latter, but some caution is warranted in the absence of confirming quantum chemical calculations.

Table 6: DFT (QZ basis) adiabatic ionization energies, including thermal corrections at 350 K, as a function of metallocene and methodology. The mean absolute errors (MAE), maximum errors (MaxE), root-mean-square deviations (RMSD), and the mean signed errors (MSE) are included. All units are in kcal/mol.

	Expt.	BP86	B3LYP	B3LYP*	PBE0	TPSSh	M06	B97	B97M-V	ω B97X-V	ω B97M-V	DSD-PBEP86
$V(Cp)_2$	154.5 ± 1.5	156.59	151.17	150.94	153.83	150.47	152.10	150.46	151.80	156.52	158.65	152.02
$Cr(Cp)_2$	127.5 ± 1.5	135.13	126.09	127.61	126.00	128.93	124.60	125.88	127.11	129.64	128.95	127.82
$Mn(Cp)_2$	142.5 ± 1.5	116.30	138.10	131.34	140.98	121.03	144.87	136.47	127.14	138.15	142.63	135.03
$Fe(Cp)_2$	153.1 ± 1.5	159.40	147.39	149.42	149.85	151.55	155.14	148.48	151.39	152.94	152.00	135.04
$Co(Cp)_2$	123.5 ± 1.5	117.11	122.22	119.74	122.55	115.17	120.96	120.78	114.60	122.45	123.58	91.05
$Ni(Cp)_2$	143.8 ± 1.5	137.79	145.20	142.05	147.27	138.04	144.10	144.04	139.12	148.18	149.31	152.52
MAE		9.10	2.92	4.00	1.89	7.09	2.09	3.21	5.62	2.35	2.07	11.59
$ MaxE $		26.20	5.71	11.16	3.47	21.47	2.90	6.03	15.36	4.38	5.51	32.45
RMSD		12.01	3.38	5.29	2.18	9.87	2.26	3.74	7.61	2.83	2.91	15.90
MSE		-3.76	-2.46	-3.97	-0.74	-6.62	-0.52	-3.13	-5.62	0.50	1.70	-8.57

Table 7: DFT (QZ basis) and DLPNO-CCSD(T) extrapolated with TZ/QZ basis sets, including thermal corrections at 350 K, adiabatic ionization energies as a function of acac and methodology. The mean absolute errors (MAE), maximum errors (MaxE), root-mean-square deviations (RMSD), and the mean signed errors (MSE) are included. All units are in kcal/mol.

	Expt.	BP86	B3LYP	B3LYP*	PBE0	TPSSh	M06	B97	B97M-V	ω B97X-V	ω B97M-V	DSD- PBEP86	DLPNO- CCSD(T)
$[V(acac)_3]^{1-}$	24.9 ± 2.4	26.79	22.00	21.96	21.91	20.77	26.21	22.35	24.07	18.36	20.25	22.34	22.50
$[Cr(acac)_3]^{1-}$	20 ± 3	2.82	5.97	3.71	5.61	-1.74	10.92	7.24	-3.23	4.37	5.92	10.01	7.79
$[Mn(acac)_3]^{1-}$	59 ± 5	35.04	45.20	41.18	46.98	36.31	49.80	46.62	38.57	46.77	49.32	57.56	57.13
$[Fe(acac)_3]^{1-}$	43 ± 2.4	39.49	39.22	38.10	37.21	36.06	32.80	38.46	32.93	40.63	40.79	43.40	34.05
$[Co(acac)_3]^{1-}$	47 ± 4	18.19	39.94	32.60	44.45	26.96	41.90	40.95	20.99	41.33	41.44	53.84	56.45
MAE		15.07	8.32	11.27	7.55	15.11	6.98	7.66	16.11	8.49	7.24	4.24	6.97
$ MaxE $		28.81	14.03	17.82	14.39	22.69	10.20	12.76	26.01	15.63	14.08	9.99	12.21
RMSD		18.52	9.59	12.83	8.95	17.06	7.73	8.72	18.63	9.74	8.36	5.57	8.10
MSE		-14.31	-8.32	-11.27	-7.55	-15.11	-6.45	-7.66	-16.11	-8.49	-7.24	-1.35	-3.20

Figure 6: Performance of DFT functionals (QZ basis) and DLPNO-CCSD(T) for prediction of experimental adiabatic gas-phase ionization energies for acac complexes. The range is limited to deviations of -15 to 15 kcal/mol. Numerous outliers (-30 to -15 kcal/mol) for $[Cr(acac)_2]^{1-}$, $[Mn(acac)_2]^{1-}$, and $[Co(acac)_2]^{1-}$ are not visible in the plot. The gray band indicates what is often given as the chemical accuracy for transition metals: 3 kcal/mol.

Reduction Potentials in Solution

Richardson and co-workers^{60,61} derive THF/acetonitrile differential solvation energies from their gas-phase and solution phase measurements. Therefore, we use their solvation energies to see if our results yield accurate reduction potentials. We leave the investigation of the proper simulation of solvation energies for a future publication.

The gas phase ionization energy is given by $\Delta G_{(g)}$, which is computed with ph-AFQMC and includes ideal gas free energy corrections. The differential solvation energy is given by $\Delta G_{solv,II} - \Delta G_{solv,III}$, for which we use the experimental value given in Ref. 8. Details on how the potential are derived using a thermodynamic cycle is given in the SI. Table 8 gives the resulting potentials using our best ph-AFQMC method as well as the best performing (in terms of MAE) DFT methods for metallocenes, namely PBE0. While the uncertainty of the experimentally-derived differential solvation energy clouds the interpretation of the results, ph-AFQMC clearly seems to perform well. The results show that accurate potentials can be

obtained by properly describing both the gas-phase and solvation parts. Alternatively, good results can otherwise be obtained using empirical corrections^{42,97–99} or error cancellation schemes.⁴⁴ Thus ph-AFQMC can be used as a microscopic approach for the computation of the gas-phase part and help isolate errors due to the solvation model.

Table 8: Computed reduction potentials (V) using the experimentally derived differential solvation energies ($E_{solv}(\text{II}) - E_{solv}(\text{III})$) (kcal/mol)^{59–61} as a function of metallocene and methodology for our best ph-AFQMC method as well as the best performing functional from above for metallocenes PBE0. The mean absolute errors (MAE), maximum errors (MaxE), root-mean-square deviations (RMSD), and the mean signed errors (MSE) are included.

	Expt. Differential Solvation Energy	Expt. Potential	AFQMC/CAS CS	PBE0
Units	kcal/mol	V	CAS Trial V	V
$V(\text{Cp})_2$	60 ± 4	-0.31	-0.38 ± 0.19	-0.37
$\text{Cr}(\text{Cp})_2$	36 ± 4	-0.43	-0.31 ± 0.19	-0.54
$\text{Mn}(\text{Cp})_2$	38 ± 4	0.11	0.25 ± 0.2	0.03
$\text{Fe}(\text{Cp})_2$	35 ± 4	0.65	0.67 ± 0.19	0.54
$\text{Co}(\text{Cp})_2$	38 ± 4	-0.60	-0.69 ± 0.19	-0.77
$\text{Ni}(\text{Cp})_2$	38 ± 4	0.25	0.18 ± 0.2	0.30
MAE			0.08 ± 0.19	0.10
$ \text{MaxE} $			0.14 ± 0.26	0.17
RMSD			0.09	0.11
MSE			0.01 ± 0.19	-0.08

V. Implications of the Results for Transition Metal Quantum Chemistry

We restrict our focus for the purpose of discussion to the adiabatic ionization energies, as there are not enough theory/experiment comparisons in our previous work on AFQMC for non-geometry-relaxed structures to warrant a more in-depth analysis. With regard to the performance of various quantum chemical methods, the results discussed above are consistent with those we have obtained for two earlier data sets comprised of transition metal

diatomics,²⁹ and simple organometallic complexes.⁵⁴ First, DLPNO-CCSD(T) calculations display MAEs that are generally in the vicinity of 3-5 kcal/mol (going up to 7 kcal/mol for the acac series), with the largest outliers in each data set typically in the range of 9-12 kcal/mol (reaching as high as 17 kcal/mol for one diatomic case). While these results are surprising in view of the chemically accurate results obtained by CCSD(T) approaches for organic systems, the data at this point seems unambiguous: while many individual cases calculated with DLPNO-CCSD(T) appear to be accurate, and average absolute errors across larger data sets can be in respectable agreement on aggregate with experiments, DLPNO-CCSD(T) calculations do not robustly yield even near-chemical accuracy, and errors sufficiently large to preclude drawing useful conclusions concerning reaction mechanisms are routinely observed.

It is possible that CCSD(T) methods which do not invoke localization approximations and use more accurate approaches to perturbative triples^{15,20} would perform better; however, full, nontruncated CCSD(T) scales as N^7 with a relatively large prefactor, and hence would be difficult to apply to the larger systems such as the acac series. It is possible that multireference coupled cluster approaches could achieve greater accuracy; however, when one implementation of MR-CCSD(T) was applied to a large data set of metal containing diatomic molecules, little improvement in the quality of results was seen as compared to single reference CCSD(T) when careful comparisons of theory and experiment were made,^{28,29} and the computational cost of MR-CCSD(T) calculations is large. The excellent agreement between our AFQMC results and experiment has enabled clean benchmarking sets to be extracted from an often confusing array of experimental and theoretical papers; these can be used to test alternative single and multireference CCSD(T)-based methodology going forward. These benchmarks should also be helpful in assessing other advanced wavefunction approaches, such as multiconfigurational pair-density functional theory (MC-PDFT),¹⁰⁰ other types of QMC such as diffusion Monte Carlo (DMC),¹⁰¹ nonorthogonal configuration interaction with second order perturbation theory (NOCI+PT2),¹⁰² and so on. When doing so, a balanced assessment of both accuracy and computational efficiency will be necessary.

The DFT results that we have obtained are also consistent with our prior work. Semi-local GGA functionals appear to be incapable of obtaining reliable results for organometallic complexes of the type we have studied to date, although performance may be better for other metal-containing systems.¹⁰³ The performance of hybrid functionals of various types (including range-separated hybrids) is highly variable, with many cases yielding results that agree well with experiment, while others appear as significant outliers with errors in the 5-10 kcal/mol range. Some outliers are specific to the functional in question, but others present problems across the entire range of alternatives that we have examined, for example the Cr and Mn complexes with acac ligands. Finally, the one double hybrid functional that we tested, DSD-PBEP86, while marginally better for the acac ligands, overall displays very large average and maximum errors, in line with the poor performance in previous studies.

We conclude from these results that DFT methods for transition metals are very promising, but need to be optimized using a much larger database of benchmark experimental and “beyond CCSD(T)” theoretical results for relevant transition-metal containing systems. Similar large scale optimization has succeeded in reducing the number of outliers present in modern DFT functionals, such as the latest range-separated hybrids, to a very substantial degree as compared to earlier generations of functionals such as B3LYP and PBE0,⁹³ although it should be noted that outliers have not been entirely eliminated. We expect that similar progress can be made for transition metal containing systems. Newer approaches, such as the use of machine learning methods to create better functional forms for the DFT functional,^{104–109} may also prove to be useful in the optimization process.

Going forward, we see the role of AFQMC for transition metal quantum chemistry as:

- (1) Generating benchmark data sets for assessment of various coupled cluster and other wavefunction approaches, and for optimization of a next generation of DFT functionals.
- (2) Obtaining results for unique, challenging systems of importance in biological and materials science, for example the Mn water splitting cluster in Photosystem II,^{110,111} or the CuO planes in high T_c superconductors.^{112,113}

Both of these applications will require validation of the accuracy of AFQMC methods for increasingly larger and more complex systems (e.g. those containing multiple metal centers), as well as improvements in computational efficiency to enable larger molecules, and larger data sets, to be effectively addressed. The present paper, while a step in this direction, has also been focused on pointing out the need for a true benchmark approach, via its assessment of the existing coupled cluster and DFT alternatives.

VI. Conclusions

We have developed an AFQMC protocol which yields results for the ionization of a series of metallocenes (normal oxidation state, coordinatively saturated organometallic complexes) that are essentially within experimental error bars for both vertical and adiabatic ionization energies. The protocol has the following key ingredients: (1) A multideterminantal trial wavefunction, based on CASSCF calculations (2) The use of a CS algorithm to compute energy differences between electronic surfaces, and a PC algorithm to calculate energy differences on the same surface, namely the reorganization energy. (3) Methods for extrapolation to the CBS limit which are upgraded to a higher quality trial function as indicated by CASSCF NOONs and spin-symmetry breaking at the B3LYP level.

We are optimistic that this protocol will enable accurate AFQMC results to be obtained in the future for a wide range of challenging and important transition metal containing species, relevant to biological and materials science problems. Tests on a range of difficult systems, such as the acac series, will be necessary to confirm this hypothesis. It is important to note that failure to implement any one of these three components can degrade the AFQMC results in some cases, to a level comparable to the best DFT calculations, thus obviating the purpose of deploying a method that is much more computationally expensive.

Our investigation of DLPNO-CCSD(T) and DFT approaches confirms our previous conclusions, that even the best of these methods, while yielding good agreement with experiment

a considerable fraction of the time, are subject to frequent large uncontrolled errors for transition metal containing systems. We postulate that training DFT functionals on a greatly expanded transition metal-focused training set of benchmark data can improve this situation significantly.

Acknowledgement

B.R. acknowledges funding from the National Institute of General Medical Sciences of the National Institutes of Health under award number F32GM136105. D.R.R. acknowledges funding from NSF CHE-1954791. S.Z. acknowledges funding from DOE DE-SC0001303. An award of computer time was provided by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program. This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562. In particular, we used San Diego Computing Center’s Comet resources under grant number TG-CHE190007 and allocation ID COL151. The Flatiron Institute is a division of the Simons Foundation. We would like to thank Joonho Lee, the QMC group at the Flatiron Institute, Prof. Martin Head-Gordon, Diptarka Hait, and Elvira Sayfutyarova for helpful discussions.

Supporting Information Available

Tabulated DLPNO-CCSD(T), ph-AFQMC vertical and adiabatic energies in the TZ basis set, scaling factors for the CBS extrapolations, free energy corrections, active space information, CASSCF energies, NOON’s, $\langle S^2 \rangle$ values, metal spin density values, details on the convergence of ph-AFQMC with respect to active space, details on the DFT integration grids, stability calculations, how CBS extrapolations are done for DLPNO-CCSD(T), a workflow

in terms of how other calculations complement the ph-AFQMC calculations, alternate experimental values, additional methodology details for ph-AFQMC, details on how potentials are calculated in solution, details on the calculation of statistical measures to compare to experiment, explanation of CS algorithms, ionization energies of acac complexes including diffuse functions on the O atoms, B3LYP M-Cp distances in the metallocenes, literature experimental homolytic bond dissociation energies for the metallocenes, and additional references (PDF) as well as B3LYP-optimized coordinates for transition metal complexes (TXT).

References

- (1) Friesner, R. A. Ab Initio Quantum Chemistry: Methodology and Applications. *Proc. Nat. Acad. Sci.* **2005**, *102*, 6648–6653.
- (2) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. Assessment of Gaussian-3 and Density Functional Theories for a Larger Experimental Test Set. *J. Chem. Phys.* **2000**, *112*, 7374–7383.
- (3) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Baboul, A. G.; Pople, J. A. Gaussian-3 Theory Using Coupled Cluster Energies. *Chem. Phys. Lett.* **1999**, *314*, 101–107.
- (4) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.
- (5) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (6) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211.

- (7) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula Into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785.
- (8) Becke, A. D. Density-Functional Thermochemistry. V. Systematic Optimization of Exchange-Correlation Functionals. *J. Chem. Phys.* **1997**, *107*, 8554–8560.
- (9) Schmider, H. L.; Becke, A. D. Optimized Density Functionals From the Extended G2 Test Set. *J. Chem. Phys.* **1998**, *108*, 9624–9631.
- (10) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods Without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
- (11) Ernzerhof, M.; Perdew, J. P. Generalized Gradient Approximation to the Angle-and System-Averaged Exchange Hole. *J. Chem. Phys.* **1998**, *109*, 3313–3320.
- (12) Hamprecht, F. A.; Cohen, A. J.; Tozer, D. J.; Handy, N. C. Development and Assessment of New Exchange-Correlation Functionals. *J. Chem. Phys.* **1998**, *109*, 6264–6271.
- (13) Ma, Q.; Schwilk, M.; Koppl, C.; Werner, H.-J. Scalable Electron Correlation Methods. 4. Parallel Explicitly Correlated Local Coupled Cluster With Pair Natural Orbitals (PNO-LCCSD-F12). *J. Chem. Comput. Chem.* **2017**, *13*, 4871–4896.
- (14) Ma, Q.; Werner, H.-J. Accurate Intermolecular Interaction Energies Using Explicitly Correlated Local Coupled Cluster Methods [PNO-LCCSD (T)-F12]. *J. Chem. Theory Comput.* **2019**, *15*, 1044–1052.
- (15) Liakos, D. G.; Sparta, M.; Kesharwani, M. K.; Martin, J. M. L.; Neese, F. Exploring the Accuracy Limits of Local Pair Natural Orbital Coupled-Cluster Theory. *J. Chem. Theory Comput.* **2015**, *11*, 1525–1539.
- (16) 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.

- (17) Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. Sparse Maps – A Systematic Infrastructure for Reduced-Scaling Electronic Structure Methods. II. Linear Scaling Domain Based Pair Natural Orbital Coupled Cluster Theory. *J. Chem. Phys.* **2016**, *144*, 024109.
- (18) Guo, Y.; Riplinger, C.; Becker, U.; Liakos, D. G.; Minenkov, Y.; Cavallo, L.; Neese, F. Communication: An Improved Linear Scaling Perturbative Triples Correction for the Domain Based Local Pair-Natural Orbital Based Singles and Doubles Coupled Cluster Method [DLPNO-CCSD(T)]. *J. Chem. Phys.* **2018**, *148*, 011101.
- (19) Saitow, M.; Becker, U.; Riplinger, C.; Valeev, E. F.; Neese, F. A New Near-Linear Scaling, Efficient and Accurate, Open-Shell Domain-Based Local Pair Natural Orbital Coupled Cluster Singles and Doubles Theory. *J. Chem. Phys.* **2017**, *146*, 164105.
- (20) Flöser, B. M.; Guo, Y.; Riplinger, C.; Tuzek, F.; Neese, F. Detailed Pair Natural Orbital-based Coupled Cluster Studies of Spin Crossover Energetics. *J. Chem. Theor. Comput.* **2020**,
- (21) Lin, H.; Truhlar, D. G. QM/MM: What Have We Learned, Where Are We, and Where Do We Go From Here? *Theor. Chem. Acc.* **2007**, *117*, 185.
- (22) Bistoni, G.; Polyak, I.; Sparta, M.; Thiel, W.; Neese, F. Toward Accurate QM/MM Reaction Barriers With Large QM Regions Using Domain Based Pair Natural Orbital Coupled Cluster Theory. *J. Chem. Theor. Comput.* **2018**, *14*, 3524–3531.
- (23) Shee, J.; Arthur, E. J.; Zhang, S.; Reichman, D. R.; Friesner, R. A. Phaseless Auxiliary-Field Quantum Monte Carlo on Graphical Processing Units. *J. Chem. Theory Comput.* **2018**, *14*, 4109–4121.
- (24) Jiang, W.; DeYonker, N. J.; Determan, J. J.; Wilson, A. K. Toward Accurate Theoretical Thermochemistry of First Row Transition Metal Complexes. *J. Phys. Chem. A* **2011**, *116*, 870–885.

- (25) 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 3d Transition Metals? *J. Chem. Theory Comput.* **2015**, *11*, 2036–2052.
- (26) 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.
- (27) 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.
- (28) Aoto, Y. A.; de Lima Batista, A. P.; Köhn, A.; de Oliveira-Filho, A. G. S. How to Arrive at Accurate Benchmark Values for Transition Metal Compounds: Computation or Experiment? *J. Chem. Theory Comput.* **2017**, *13*, 5291–5316.
- (29) Shee, J.; Rudsteyn, B.; Arthur, E. J.; Zhang, S.; Reichman, D. R.; Friesner, R. A. On Achieving High Accuracy in Quantum Chemical Calculations of 3d 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.
- (30) Williams, K. T. et al. Direct Comparison of Many-Body Methods for Realistic Electronic Hamiltonians. *Phys. Rev. X* **2020**, *10*, 011041.
- (31) Bross, D. H.; Hill, J. G.; Werner, H.-J.; Peterson, K. A. Explicitly Correlated Composite Thermochemistry of Transition Metal Species. *J. Chem. Phys.* **2013**, *139*, 094302.

- (32) Manivasagam, S.; Laury, M. L.; Wilson, A. K. Pseudopotential-Based Correlation Consistent Composite Approach (Rp-ccCA) for First-and Second-Row Transition Metal Thermochemistry. *J. Phys. Chem. A* **2015**, *119*, 6867–6874.
- (33) Jiang, W.; Laury, M. L.; Powell, M.; Wilson, A. K. Comparative Study of Single and Double Hybrid Density Functionals for the Prediction of 3d Transition Metal Thermochemistry. *J. Comput. Theory Comput.* **2012**, *8*, 4102–4111.
- (34) Zhang, W.; Truhlar, D. G.; Tang, M. Tests of Exchange-Correlation Functional Approximations Against Reliable Experimental Data for Average Bond Energies of 3d Transition Metal Compounds. *J. Chem. Theory Comput.* **2013**, *9*, 3965–3977.
- (35) Moltved, K. A.; Kepp, K. P. Chemical Bond Energies of 3d Transition Metals Studied by Density Functional Theory. *J. Chem. Theory Comput.* **2018**,
- (36) Carlson, R. K.; Li Manni, G.; Sonnenberger, A. L.; Truhlar, D. G.; Gagliardi, L. Multiconfiguration Pair-Density Functional Theory: Barrier Heights and Main Group and Transition Metal Energetics. *J. Chem. Theor. Comp.* **2014**, *11*, 82–90.
- (37) Bao, J. L.; Odoh, S. O.; Gagliardi, L.; Truhlar, D. G. Predicting Bond Dissociation Energies of Transition-Metal Compounds by Multiconfiguration Pair-Density Functional Theory and Second-Order Perturbation Theory Based on Correlated Participating Orbitals and Separated Pairs. *J. Chem. Theory Comput.* **2017**, *13*, 616–626.
- (38) Bao, J. L.; Zhang, X.; Xu, X.; Truhlar, D. G. Predicting Bond Dissociation Energy and Bond Length for Bimetallic Diatomic Molecules: A Challenge for Electronic Structure Theory. *Phys. Chem. Chem. Phys.* **2017**, *19*, 5839–5854.
- (39) Sharkas, K.; Gagliardi, L.; Truhlar, D. G. Multiconfiguration Pair-Density Functional Theory and Complete Active Space Second Order Perturbation Theory. Bond Dissociation Energies of FeC, NiC, FeS, NiS, FeSe, and NiSe. *J. Phys. Chem. A* **2017**, *121*, 9392–9400.

- (40) Tran, L. N.; Iskakov, S.; Zgid, D. Spin-Unrestricted Self-Energy Embedding Theory. *J. Phys. Chem. Lett.* **2018**,
- (41) Hughes, T. F.; Friesner, R. A. Development of Accurate DFT Methods for Computing Redox Potentials of Transition Metal Complexes: Results for Model Complexes and Application to Cytochrome P450. *J. Chem. Theor. Comput.* **2012**, *8*, 442–459.
- (42) Coskun, D.; Jerome, S. V.; Friesner, R. A. Evaluation of the Performance of the B3LYP, PBE0, and M06 DFT Functionals, and DBLOC-Corrected Versions, in the Calculation of Redox Potentials and Spin Splittings for Transition Metal Containing Systems. *J. Chem. Theory Comput.* **2016**, *12*, 1121–1128.
- (43) Roy, L. E.; Jakubikova, E.; Guthrie, M. G.; Batista, E. R. Calculation of One-Electron Redox Potentials Revisited. Is It Possible to Calculate Accurate Potentials With Density Functional Methods? *J. Phys. Chem. A* **2009**, *113*, 6745–6750.
- (44) Konezny, S. J.; Doherty, M. D.; Luca, O. R.; Crabtree, R. H.; Soloveichik, G. L.; Batista, V. S. Reduction of Systematic Uncertainty in DFT Redox Potentials of Transition-Metal Complexes. *J. Phys. Chem. C* **2012**, *116*, 6349–6356.
- (45) Green, J. C. *Bonding Problems*; Springer, 1981; pp 37–112.
- (46) Zhang, S.; Krakauer, H. Quantum Monte Carlo Method Using Phase-Free Random Walks With Slater Determinants. *Phys. Rev. Lett.* **2003**, *90*, 136401.
- (47) Al-Saidi, W. A.; Zhang, S.; Krakauer, H. Auxiliary-Field Quantum Monte Carlo Calculations of Molecular Systems With a Gaussian Basis. *J. Chem. Phys.* **2006**, *124*, 224101.
- (48) Motta, M.; Zhang, S. Ab initio Computations of Molecular Systems by the Auxiliary-Field Quantum Monte Carlo Method. *WIREs Comput. Mol. Sci.* **2018**, *8*, e1364.

- (49) Bartlett, R. J.; Musiał, M. Coupled-Cluster Theory in Quantum Chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291.
- (50) Shi, H.; Zhang, S. Some Recent Developments in Auxiliary-Field Quantum Monte Carlo for Real Materials. *J. Chem. Phys.* **2021**, *154*, 024107.
- (51) Shee, J.; Zhang, S.; Reichman, D. R.; Friesner, R. A. Chemical Transformations Approaching Chemical Accuracy via Correlated Sampling in Auxiliary-Field Quantum Monte Carlo. *J. Chem. Theory Comput.* **2017**, *13*, 2667–2680.
- (52) Malone, F. D.; Zhang, S.; Morales, M. A. Accelerating Auxiliary-Field Quantum Monte Carlo Simulations of Solids with Graphical Processing Units. *J. Chem. Theor. Comput.* **2020**, *16*, 4286–4297.
- (53) Lee, J.; Malone, F. D.; Morales, M. A. Utilizing Essential Symmetry Breaking in Auxiliary-Field Quantum Monte Carlo: Application to the Spin Gaps of the C₃₆ Fullerene and an Iron Porphyrin Model Complex. *J. Chem. Theory Comput.* **2020**, *16*, 3019–3027.
- (54) Rudsteyn, B.; Coskun, D.; Weber, J. L.; Arthur, E. J.; Zhang, S.; Reichman, D. R.; Friesner, R. A.; Shee, J. Predicting Ligand-Dissociation Energies of 3d Coordination Complexes with Auxiliary-Field Quantum Monte Carlo. *J. Chem. Theor. Comput.* **2020**, *16*, 3041–3054.
- (55) Weber, J. L.; Churchill, E. M.; Jockusch, S.; Arthur, E. J.; Pun, A. B.; Zhang, S.; Friesner, R. A.; Campos, L. M.; Reichman, D. R.; Shee, J. In Silico Prediction of Anihilators for Triplet–Triplet Annihilation Upconversion via Auxiliary-Field Quantum Monte Carlo. *Chem. Sci.* **2020**, *In Press*.
- (56) Morse, M. D. Predissociation Measurements of Bond Dissociation Energies. *Acc. Chem. Res.* **2019**, *52*, 119–126.

- (57) Hait, D.; Tubman, N. M.; Levine, D. S.; Whaley, K. B.; Head-Gordon, M. 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. *J. Chem Theory Comput.* **2019**, *15*, 5370–5385.
- (58) Pavlishchuk, V. V.; Addison, A. W. Conversion Constants for Redox Potentials Measured Versus Different Reference Electrodes in Acetonitrile Solutions at 25 °C. *Inorganica Chimica Acta* **2000**, *298*, 97–102.
- (59) Richardson, D. E. *Organometallic Ion Chemistry*; Springer, 1996; pp 259–282.
- (60) Ryan, M. F.; Eyler, J. R.; Richardson, D. E. Adiabatic Ionization Energies, Bond Disruption Enthalpies, and Solvation Free Energies for Gas-Phase Metallocenes and Metallocenium Ions. *J. Am. Chem. Soc.* **1992**, *114*, 8611–8619.
- (61) Ryan, M. F.; Richardson, D. E.; Lichtenberger, D. L.; Gruhn, N. E. Gas-Phase Ionization Energetics, Electron-Transfer Kinetics, and Ion Solvation Thermochemistry of Decamethylmetallocenes, Chromocene, and Cobaltocene. *Organometallics* **1994**, *13*, 1190–1199.
- (62) Cauletti, C.; Green, J. C.; Kelly, M. R.; Powell, P.; van Tilborg, J.; Robbins, J.; Smart, J. Photoelectron Spectra of Metallocenes. *J. Elec. Spec. Rel. Phen.* **1980**, *19*, 327–353.
- (63) Salomon, O.; Reiher, M.; Hess, B. A. Assertion and Validation of the Performance of the B3LYP* Functional for the First Transition Metal Row and the G2 Test Set. *J. Chem. Phys.* **2002**, *117*, 4729–4737.
- (64) DeYonker, N. J.; Peterson, K. A.; Steyl, G.; Wilson, A. K.; Cundari, T. R. Quantitative Computational Thermochemistry of Transition Metal Species. *J. Phys. Chem. A* **2007**, *111*, 11269–11277.

- (65) Siegbahn, P. E. M. The Performance of Hybrid DFT for Mechanisms Involving Transition Metal Complexes in Enzymes. *J. Bio. Inorg. Chem.* **2006**, *11*, 695–701.
- (66) Mardirossian, N.; Head-Gordon, M. ω B97M-V: A Combinatorially Optimized, Range-Separated Hybrid, Meta-GGA Density Functional With VV10 Nonlocal Correlation. *J. Chem. Phys.* **2016**, *144*, 214110.
- (67) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
- (68) Mardirossian, N.; Head-Gordon, M. ω B97X-V: A 10-Parameter, Range-Separated Hybrid, Generalized Gradient Approximation Density Functional With Nonlocal Correlation, Designed by a Survival-of-the-Fittest Strategy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904–9924.
- (69) Sharpe, P.; Richardson, D. E. Metal-Ligand Bond Energies and Solvation Energies for Gas-Phase Transition-Metal Tris (Acetylacetonate) Complexes and Their Negative Ions. *J. Am. Chem. Soc.* **1991**, *113*, 8339–8346.
- (70) Shee, J.; Arthur, E. J.; Zhang, S.; Reichman, D. R.; Friesner, R. A. Singlet–Triplet Energy Gaps of Organic Biradicals and Polyacenes with Auxiliary-Field Quantum Monte Carlo. *J. Chem. Theory Comput.* **2019**,
- (71) Trasatti, S., et al. The Absolute Electrode Potential: An Explanatory Note (Recommendations 1986). *Pure Appl. Chem* **1986**, *58*, 955–966.
- (72) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*; John Wiley & Sons, 2013.

- (73) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron Through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*.
- (74) Woon, D. E.; Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. III. The Atoms Aluminum Through Argon. *J. Chem. Phys.* **1993**, *98*.
- (75) de Jong, W. A.; Harrison, R. J.; Dixon, D. A. Parallel Douglas-Kroll Energy and Gradients in NWChem: Estimating Scalar Relativistic Effects Using Douglas-Kroll Contracted Basis Sets. *J. Chem. Phys.* **2001**, *114*.
- (76) Balabanov, N. B.; Peterson, K. A. Systematically Convergent Basis Sets for Transition Metals. I. All-Electron Correlation Consistent Basis Sets for the 3d Elements Sc–Zn. *J. Chem. Phys.* **2005**, *123*, 064107.
- (77) Pantazis, D. A.; Chen, X.-Y.; Landis, C. R.; Neese, F. All-Electron Scalar Relativistic Basis Sets for Third-Row Transition Metal Atoms. *J. Chem. Theory Comput.* **2008**, *4*, 908–919.
- (78) Neese, F. The ORCA Program System. *WIREs Comput. Mol. Sci.* **2012**, *2*, 73–78.
- (79) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-Set Convergence of Correlated Calculations on Water. *J. Chem. Phys.* **1997**, *106*, 9639–9646.
- (80) Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; Wouters, S.; Chan, G. K. PySCF: the Python–Based Simulations of Chemistry Framework. *WIREs Comput. Mol. Sci.* **8**, e1340.
- (81) Liu, W.; Peng, D. Exact Two-Component Hamiltonians Revisited. *J. Chem. Phys.* **2009**, *131*, 031104.
- (82) Purwanto, W.; Al-Saidi, W.; Krakauer, H.; Zhang, S. Eliminating Spin Contamination

- in Auxiliary-Field Quantum Monte Carlo: Realistic Potential Energy Curve of F_2 . *J. Chem. Phys.* **2008**, *128*, 114309.
- (83) Roos, B. O.; Lindh, R.; Malmqvist, P.-Å.; Veryazov, V.; Widmark, P.-O. New Relativistic ANO Basis Sets for Transition Metal Atoms. *J. Phys. Chem. A* **2005**, *109*, 6575–6579.
- (84) Sayfutyarova, E. R.; Sun, Q.; Chan, G. K.-L.; Knizia, G. Automated Construction of Molecular Active Spaces From Atomic Valence Orbitals. *J. Chem. Theory Comput.* **2017**, *13*, 4063–4078.
- (85) Becke, A. D. Density-functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098.
- (86) Perdew, J. P. Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B* **1986**, *33*, 8822.
- (87) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91*, 146401.
- (88) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. Comparative Assessment of a New Nonempirical Density Functional: Molecules and Hydrogen-Bonded Complexes. *J. Chem. Phys.* **2003**, *119*, 12129–12137.
- (89) Mardirossian, N.; Head-Gordon, M. Mapping the Genome of Meta-Generalized Gradient Approximation Density Functionals: The Search for B97M-V. *J. Chem. Phys.* **2015**, *142*, 074111.
- (90) Kozuch, S.; Martin, J. M. DSD-PBEP86: In Search of the Best Double-Hybrid DFT With Spin-Component Scaled MP2 and Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20104–20107.

- (91) Kozuch, S.; Martin, J. M. Spin-Component-Scaled Double Hybrids: An Extensive Search for the Best Fifth-Rung Functionals Blending DFT and Perturbation Theory. *J. Comput. Chem.* **2013**, *34*, 2327–2344.
- (92) Shee, J.; Loipersberger, M.; Hait, D.; Lee, J.; Head-Gordon, M. Revealing the Nature of Electron Correlation in Transition Metal Complexes With Symmetry-Breaking and Chemical Intuition. *J. Chem. Phys.* **2021**, *154*, 194109.
- (93) 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.
- (94) Furche, F.; Perdew, J. P. The Performance of Semilocal and Hybrid Density Functionals in 3 D Transition-Metal Chemistry. *J. Chem. Phys.* **2006**, *124*, 044103.
- (95) Zhao, Y.; Truhlar, D. G. Comparative Assessment of Density Functional Methods for 3d Transition-Metal Chemistry. *J. Chem. Phys.* **2006**, *124*, 224105.
- (96) Hughes, T. F.; Friesner, R. A. Correcting Systematic Errors in DFT Spin-Splitting Energetics for Transition Metal Complexes. *J. Chem. Theor. Comput.* **2011**, *7*, 19–32.
- (97) Friesner, R. A.; Knoll, E. H.; Cao, Y. A Localized Orbital Analysis of the Thermochemical Errors in Hybrid Density Functional Theory: Achieving Chemical Accuracy via a Simple Empirical Correction Scheme. *J. Chem. Phys.* **2006**, *125*, 124107.
- (98) Friesner, R. A.; Jerome, S. V. Localized Orbital Corrections for Density Functional Calculations on Transition Metal Containing Systems. *Coord. Chem. Rev.* **2017**, *344*, 205–213.
- (99) Jerome, S. V.; Hughes, T. F.; Friesner, R. A. Successful Application of the DBLOC Method to the Hydroxylation of Camphor by Cytochrome p450. *Protein Sci.* **2016**, *25*, 277–285.

- (100) Gagliardi, L.; Truhlar, D. G.; Li Manni, G.; Carlson, R. K.; Hoyer, C. E.; Bao, J. L. Multiconfiguration Pair-Density Functional Theory: A New Way to Treat Strongly Correlated Systems. *Acc. Chem. Res.* **2016**, *50*, 66–73.
- (101) Doblhoff-Dier, K.; Meyer, J.; Hoggan, P. E.; Kroes, G.-J.; Wagner, L. K. Diffusion Monte Carlo for Accurate Dissociation Energies of 3d Transition Metal Containing Molecules. *J. Chem. Theory Comput.* **2016**, *12*, 2583–2597.
- (102) Burton, H. G. A.; Thom, A. J. W. Reaching Full Correlation through Nonorthogonal Configuration Interaction: A Second-Order Perturbative Approach. *J. Chem. Theory Comput.* **2020**, *16*, 5586–5600.
- (103) Cramer, C. J.; Włoch, M.; Piecuch, P.; Puzzarini, C.; Gagliardi, L. Theoretical Models on the Cu₂O₂ Torture Track: Mechanistic Implications for Oxytyrosinase and Small-Molecule Analogues. *J. Phys. Chem. A* **2006**, *110*, 1991–2004.
- (104) Snyder, J. C.; Rupp, M.; Hansen, K.; Müller, K.-R.; Burke, K. Finding Density Functionals With Machine Learning. *Phys. Rev. Lett.* **2012**, *108*, 253002.
- (105) Schmidt, J.; Benavides-Riveros, C. L.; Marques, M. A. Machine Learning the Physical Nonlocal Exchange–Correlation Functional of Density-Functional Theory. *J. Phys. Chem. Lett.* **2019**, *10*, 6425–6431.
- (106) Nagai, R.; Akashi, R.; Sugino, O. Completing Density Functional Theory by Machine Learning Hidden Messages From Molecules. *Npj Comput. Mater.* **2020**, *6*, 1–8.
- (107) Dick, S.; Fernandez-Serra, M. Machine Learning Accurate Exchange and Correlation Functionals of the Electronic Density. *Nat. Commun.* **2020**, *11*, 1–10.
- (108) Meyer, R.; Weichselbaum, M.; Hauser, A. W. Machine Learning Approaches Toward Orbital-Free Density Functional Theory: Simultaneous Training on the Kinetic Energy

- Density Functional and Its Functional Derivative. *J. Chem. Theory Comput.* **2020**, *16*, 5685–5694.
- (109) Kalita, B.; Li, L.; McCarty, R. J.; Burke, K. Learning to Approximate Density Functionals. *Acc. Chem. Res.* **2021**, *54*, 818–826.
- (110) Askerka, M.; Brudvig, G. W.; Batista, V. S. The O₂-Evolving Complex of Photosystem II: Recent Insights from Quantum Mechanics/Molecular Mechanics (QM/MM), Extended X-ray Absorption Fine Structure (EXAFS), and Femtosecond X-ray Crystallography Data. *Acc. Chem. Res.* **2016**, *50*, 41–48.
- (111) Siegbahn, P. E. M. Nucleophilic Water Attack Is Not a Possible Mechanism for O–O Bond Formation in Photosystem II. *Proc. Nat. Acad. Sci.* **2017**, *114*, 4966–4968.
- (112) Ito, T.; Takagi, H.; Ishibashi, S.; Ido, T.; Uchida, S. Normal-State Conductivity Between CuO₂ Planes in Copper Oxide Superconductors. *Nature* **1991**, *350*, 596–598.
- (113) Uemura, Y. J.; Luke, G. M.; Sternlieb, B. J.; Brewer, J. H.; Carolan, J. F.; Hardy, W.; Kadono, R.; Kempton, J. R.; Kiefl, R. F.; Kreitzman, S. R., et al. Universal Correlations Between T_c and $\frac{n_s}{m^*}$ (Carrier Density Over Effective Mass) in High-T_c Cuprate Superconductors. *Phys. Rev. Lett.* **1989**, *62*, 2317.

Graphical TOC Entry