Developments in Auxiliary Field Quantum Monte Carlo for Molecules

John Landstrom Weber

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2023

© 2023

John Landstrom Weber

All Rights Reserved

Abstract

Developments in Auxiliary Field Quantum Monte Carlo for Molecules

John Landstrom Weber

This thesis presents a compilation of recent work on benchmarking, applying, and developing Auxiliary Field Quantum Monte Carlo (AFQMC) for use in *ab initio* simulations of the electronic structure of molecules. With Chapter 1 I begin with a benchmark of phaseless AFQMC versus experiment in obtaining gas phase ligand dissociation energies of a set of tetrahedral and octahedral transition metal complexes. ph-AFQMC is shown to acquire chemical accuracy through the use of correlated sampling (CS) and CASSCF trial wavefunctions selected via a black box procedure. This is followed in Chapter 2 with another gas phase benchmark of ph-AFQMC versus experiment, this time calculating the redox potentials for a set of metallocenes, where we find a mix of correlated sampling and large CASSCF trials necessary to reproduce gas phase experimental values to within 1.7 ± 1.0 kcal/mol. Additionally, the inclusion of QZ ph-AFQMC values, either using UHF or CASSCF trials, was found to be necessary for a few systems, as opposed to using a hybrid approach with alternate methods such as coupled cluster to extrapolate to the basis set limit.

In Chapter 3, having established protocols to obtain decent results on transition metal complexes with known experimental values, I apply ph-AFQMC to successfully predict the activity of a set of new annihilators for optical upconversion. For a set of functionalized anthracene molecules, I report agreement within statistics between ph-AFQMC and a localized approximation to coupled cluster singles doubles and perturbative triples (DLPNO-CCSD(T_0)),

and develop intuitive guidelines for tuning the excited state energies of anthracene. For a single molecule in an additional set of functionalized benzothiadiazole (BTD) molecules, Ph-BTD, ph-AFQMC and DLPNO-CCSD(T) disagree significantly; subsequent experimental testing validates the ph-AFQMC result.

In Chapter 4 I present an approach based on localized orbitals to reduce the scaling with system size from quartic to cubic for the energy evaluation, the functional bottleneck for the majority of AFQMC calculations. Additionally, I describe the practical implementation of such an algorithm to be run on large GPU clusters. This allows AFQMC to be run for both larger systems and trials at a significantly decreased cost, while still reproducing full AFQMC results within the statistics of the method.

With Chapter 5, I conclude with the development and characterization of a novel constraint, linecut (lc-) AFQMC, which exhbits distinct behavior versus the phaseless constraint. We demonstrate benchmarks for a variety of weakly to strongly correlated molecules for which we have the exact total energies, and observe that lc-AFQMC outperforms ph-AFQMC for the majority of systems studied. I conclude with the description of a systematic method to remove the linecut constraint, partially removing the bias and re-introducing the fermionic sign problem while still maintaining a practicable signal to noise ratio. This allows for us to recover the exact energy of FeO with a fraction of the cost of converging the trial wavefunction within ph-AFQMC.

Table of Contents

Acknowledgments xvii

Dedication xix

Preface 1

Chapter 1: Predicting Ligand-Dissociation Energies of 3d Coordination Complexes with Auxiliary-Field Quantum Monte Carlo 2

- 1.1 Introduction 2
- 1.2 Selection of Experimental Data 5
- 1.3 Computational Details 6
- 1.4 Results and Discussion 10
  - 1.4.1 Dihydrogen Complexes 10
  - 1.4.2 Aqua Complexes 12
  - 1.4.3 Ammonia Complexes 14
  - 1.4.4 Carbonyl Complexes 15
  - 1.4.5 Miscellaneous Complexes 17
  - 1.4.6 Performance for the Entire Test Set 18
  - 1.4.7 Discussion 20
- 1.5 Conclusions 24

Chapter 2: Benchmarking Phaseless Auxiliary Field Quantum Monte Carlo (ph-AFQMC) on Transition Metal Complexes 25

2.1 Introduction 25

2.2 II. Experimental Data Sets 32

2.3 III. Computational Details 34

2.4 IV. Results 40

2.4.1 Vertical Ionization Energies 40

2.4.2 Reorganization Energies 45

2.4.3 Adiabatic Ionization Energies 45

2.4.4 Acetylacetonate (acac) Results 48

2.4.5 Reduction Potentials in Solution 53

2.5 V. Implications of the Results for Transition Metal Quantum Chemistry 54

2.6 VI. Conclusions 58

Chapter 3: In silico prediction of annihilators for triplet–triplet annihilation upconversion via auxiliary-field quantum Monte Carlo 60

3.1 Introduction 60

3.2 Methods 64

3.2.1 AFQMC Methodology 64

3.2.2 DFT and DLPNO-CCSD(T) Calculations 65

3.2.3 Experimental Methods 67

3.3 Results 67

3.3.1 Calculating Accurate Triplet Energies for TTA Upconversion Annihilators . 67

3.3.2 Predicting Upconversion Activity 72

3.3.3 Observation of Upconversion 75

3.3.4 Comparison to Low Temperature Phosphorescence 78

3.4 Discussion 79

3.5 Conclusions 81

3.6 Appendices 82

3.6.1 Experimental Methods 82

3.6.2 Additional Calculation Details 86

Chapter 4: A Localized-Orbital Energy Evaluation for Auxiliary-Field Quantum Monte Carlo 95

4.1 Introduction 95

4.2 Theory 100

4.2.1 Overview of ph-AFQMC 100

4.2.2 Half-Rotated ERIs 104

4.2.3 LO-AFQMC Algorithms 105

4.3 Results 109

4.3.1 Metallocenes 111

4.3.2 Polyacenes 112

4.3.3 Platonic Hydrocarbons 116

4.3.4 GPU implementation and timing 117

4.4 Conclusions 120

4.5 Appendices 123

4.5.1 MS Algorithm 123

4.5.2 LO-AFQMC in the MS Algorithm 125

4.5.3 Additional Details of GPU implementation 127
4.5.4 SMW algorithm for the formation of the Green's Function 128
4.5.5 LO error cancellation through imaginary time 130

Chapter 5: The Design of New Practical Constraints in Auxiliary-Field Quantum Monte Carlo 132

5.1 Introduction 132
5.2 Theory 135
5.3 Results 138
5.3.1 Open Shell Atoms 138
5.3.2 The C_2 Dimer 139
5.3.3 Bond Breaking Model Systems 142
5.3.4 Linecut Release 145
5.4 Conclusions 149
5.5 Appendix 151
5.5.1 Details of AFQMC calculations 151
5.5.2 Size Consistency and Stability 152

References 153

List of Figures

1.1 The types of transition metal compounds studied. M can be any $3d$ transition metal from Ti to Cu. 5

1.2 Correlated sampling ph-AFQMC calculations using Summit GPU's; statistical errors from correlated and uncorrelated sampling approaches are compared for the Cu-H_2 bond dissociation energy of the $[Cu(H_2)_4]^+$ molecule. 7

1.3 Schematic of BDE calculations performed in this work. OS abbreviates oxidation state, CS indicates the energy measured by the correlated sampling approach. . . . 8

1.4 Deviations [kcal/mol] of computational methods for the dihydrogen set of bond dissociation reactions where the H_2 that leaves is given at the end of the formula. . 11

1.5 Deviations [kcal/mol] of computational methods for the aqua set of bond dissociation reactions where the H_2O that leaves is given at the end of the formula. 13

1.6 Deviations [kcal/mol] of computational methods for the aqua set of bond dissociation reactions where the NH_3 that leaves is given at the end of the formula. 14

1.7 Deviations [kcal/mol] of computational methods for the carbonyl set of bond dissociation reactions where the CO that leaves is given at the end of the formula. . . . 16

1.8 Deviations [kcal/mol] of computational methods for the other reactions where the ligand that leaves is given at the end of the formula. 17

2.1 The structure of the metal complexes studied. 33

2.2 Schematic of the ionization and reorganization electronic energy calculations performed in this work without consideration of thermal corrections . CS and PC indicate the energy was measured by the CS and PC approaches, respectively. . . . 39

2.3 Performance of ph-AFQMC (CAS trial, CS), DFT functionals (QZ basis), DLPNO-CCSD(T_0) (extrapolated with TZ/QZ basis sets), 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 Co(Cp)$_2$ and Ni(Cp)$_2$ are out of range with deviations of about -20 and 17 kcal/mol respectively. The gray band indicates the uncertainty of the experiments i.e. 2.31 kcal/mol. The errors are given in both units of kcal/mol (lefthand axis) and kJ/mol (righthand axis). 43

2.4 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 $[Mn^{II}Cp_2]^0$. The relative energies (relative E) are given in both units of kcal/mol (lefthand axis) and kJ/mol (righthand axis). 47

2.5 Performance of ph-AFQMC, DFT functionals (QZ basis), and DLPNO-CCSD(T_0) (extrapolated with TZ/QZ basis sets) 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 value for Mn(Cp)$_2$ and the DSD-PBEP86 value for Co(Cp)$_2$ are beyond of the range of the plot with deviations of about -20 and -25 kcal/mol, respectively. The gray band indicates the uncertainty of the experiments i.e. 1.5 kcal/mol. The errors are given in both units of kcal/mol (lefthand axis) and kJ/mol (righthand axis). 49

2.6 Performance of DFT functionals (QZ basis) and DLPNO-CCSD(T_0) (extrapolated with TZ/QZ basis sets) for prediction of experimental adiabatic gas-phase ionization energies for acac complexes. The range is limited to deviations of -20 to 20 kcal/mol. The gray band indicates what is often given as the chemical accuracy for transition metals: 3 kcal/mol. The errors are given in both units of kcal/mol (lefthand axis) and kJ/mol (righthand axis). 53

3.1 A schematic of photon upconversion via triplet-triplet annihilation (TTA). First, the sensitizer is photoexcited to the first excited singlet state (S1), before undergoing rapid intersystem crossing (ISC) to a long-lived triplet state. Collision with an annihilator enables transfer of the triplet state to an annihilator via Dexter triplet-triplet energy transfer (TET). Two annihilators in the T1 state can then undergo TTA in a spin-allowed transition resulting in one S1 and one ground state annihilator, the former of which can then emit a high energy photon via fluorescence. Note that in each step excess energy is lost as heat to the surroundings. 61

3.2 Anthracene derivatives included in this study. 68

3.3 A comparison of T1 values for included TD-DFT functionals, KS-B3LYP, and DLPNO-CCSD(T). Note that KS-B3LYP obtains significantly lower triplet energies when compared to AFQMC for the entire series, whereas DLPNO-CCSD(T) is within one standard deviation from AFQMC for the majority of the compounds involved, with a mean average deviation from AFQMC below the average statistical error of the latter. All DLPNO-CCSD(T) and AFQMC values reflect gas-phase calculations in the cc-pVTZ basis. Numbers for these, as well as CBS/solvation corrections, can be found in the SI. 69

3.4 BTD and BSeD derivatives included in this study. 70

3.5 A comparison of the triplet energies calculated with TD-DFT, KS-B3LYP, and DLPNO-CCSD(T). Once more KS-B3LYP obtains significantly lower triplet energies as compared to AFQMC for nearly all species, except notably for the parent BTD and BSeD compounds, and Ph-BTD. Ph-BTD stands out as an outlier, with the maximum deviation of DLPNO-CCSD(T) from AFQMC at $\simeq 0.3$ eV. Numbers can be seen in the SI. †MeO-BSeD represents AFQMC in the triple-ζ basis (i.e. not CBS limit). 71

3.6 Comparison of the predicted gas phase 2x T1 energetic value, as calculated using AFQMC/U, with S1 values obtained from TD-DFT/CAM-B3LYP for the Anthracene derivatives. Note that the mono-substituted CN-anthracene, Ac-CN1, exhibits the lowest difference between 2x T1 and S1, and therefore the lowest potential energy loss during TTA. Further note the destabilization of the triplet state for the highly inductively-withdrawing CF_3 substituted species, versus e.g. the π system electron withdrawing DCA. 74

3.7 Comparison of the predicted gas-phase 2x T1 energetic value as calculated using AFQMC/CAS, vs S1 using TD-DFT/CAM-B3LYP for the BTD/BSeD derivatives. All molecules exhibit 2xT1 > S1, and therefore are exothermic towards upconversion. Substitution on the phenyl ring of BTD leads to a lowering of both the singlet and triplet excited states in all cases, more notably for the electron donating functional groups. Ph-BTD exhibits the lowest 2xT1 - S1. †MeO-BSeD represents AFQMC in the triple-ζ basis (i.e. not CBS limit). 75

3.8 Predicted AFQMC values versus experimental sensitizer triplet energies for the BTD series. The sensitizer triplet energy is generally required to be above that of the annihilator in order to perform efficient, exergonic triplet-triplet energy transfer, and so we expect Ph-BTD to upconvert when paired with PtOEP. 76

3.9 Photoluminescence confirmation of upconversion (dark blue) by the Ph-BTD/PtOEP system in toluene upon excitation with 532 nm light (green line), absorption of PtOEP (green dashed), absorption (light blue dashed) and photoluminescence (light blue solid) spectra of Ph-BTD. Note that the Ph-BTD/PtOEP upconversion system emits at a higher energy than the excitation wavelength and that the Ph-BTD does not directly absorb light at the excitation wavelength. Visual observation (insert) corroborates this measurement. 77

3.10 Absorption (dashed) spectrum of ZnTPP, and photoluminescence (solid) of the Ph-BTD/ZnTPP pair in toluene with excitation at 532 nm (green line). The emission of the mixture matches that of ZnTPP (Figure 3.12), signifying inefficient TET to Ph-BTD, as predicted by the relative T1 energy levels. 77

3.11 Absorption (dashed) and photoluminescence (solid) of PtOEP in toluene. 83

3.12 Absorption (dashed) and photoluminescence (solid) of ZnTPP in toluene. 83

3.13 Absorption (dashed) spectrum of PtOEP, and photoluminescence (solid) of the BTD/PtOEP pair in toluene with excitation at 532 nm (green line). The emission of the mixture matches that of PtOEP (Figure 3.11), supporting that there is no TET from PtOEP to BTD in solution. 84

3.14 Absorption (dashed) spectrum of ZnTPP, and photoluminescence (solid) of the BTD/ZnTPP pair in toluene with excitation at 532 nm (green line). The emission of the mixture matches that of ZnTPP, (Figure 3.12), supporting that there is no TET from ZnTPP to BTD in solution. 84

3.15 Time-resolved phosphorescence spectra of BTD derivatives in methylcyclohexane (a, b) or methylcyclohexane / iodomethane mixture (2:1, v/v) (c, d) at 77 K. λex = 340 nm (a, b) or λex = 355 nm (c, d). Detection window after light pulse: 10 - 30 ms (a), 1 - 6 ms (b), 100 - 600 μs (c), 20 - 1020 μs (d). 86

3.16 HOMO(left) and LUMO (right) orbital plots for DMA. Note the extension of the π-system into the Methyl groups. 92

3.17 Difference-density plots for the S0-S1 transition for CN and MeO substituted BTD and anthracene. A red surface denotes electron loss upon excitation, while the blue is electron density gained. Note the significant charge transfer in both BTD cases from the functional groups to the sulfur ring, regardless of electron withdrawing (CN) or donating (OMe) character; withdrawal of electrons from CN likely raises the excited state energy in comparison to OMe, even with similar HOMO-LUMO π-extension. In comparison, DCA and OMe-anthracene show the opposite trend, while the difference density plot shows electron withdrawing to the CN and electron donation from the OMe. 93

4.1 Convergence of the LO-AFQMC error with regards to T$_{SVD}$ for dodecane in the cc-pVDZ basis. The error vs regular AFQMC (top, blue) converges to within 1 mHa (red shaded region) around T$_{SVD}$ = 0.0005. Note that values to the left have a tighter threshold, and that the x axis is in a logarithmic scale. The compression efficiency (middle, green) ranges between 82 and 97 %, with an intermediate $T_{SVD} = 10^{-4}$ threshold reducing M by 91.75%. The reduced computational scaling due to the compression is reflected in the speedup vs regular HR-ERI (bottom, red) of the energy evaluation for 16 walkers, which even for the tightest threshold tested (T$_{SVD}$ = 10^{-5}) is over 15x faster than our HR-ERI code, with $T_{SVD} = 10^{-4}$ being ≃ 31 times faster. These calculations were performed on a single NVIDIA GeForce RTX 3090. 110

4.2 A histogram of M$_{SVD}$ for LO-AFQMC in the canonical (orange) and localized (blue) molecular orbital bases for dodecane with T$_{SVD}$ = 10^{-4}. Localization results in an increase in compression efficiency from 48.89% to 91.75%. Note that the full basis size M = 298. 111

4.3 The chemical structure of a generic metallocene, M(Cp)$_2$ 111

4.4 Chemical structure of the polyacenes 1-5 112

4.5 Distribution of LO error over 100,244 walkers after a single block (0.1 Ha^{-1}) of propagation for benzene, fitted to a normal distribution with a mean of 3.6×10^{-5} and standard deviation of 3.75×10^{-6}. 115

4.6 Scaling of the error of LO-AFQMC vs full ph-AFQMC, estimated at a time slice of 280 Ha^{-1}, with respect to the basis size for fixed value of $T_{SVD} = 10^{-4}$. Error bars are not plotted since the error is reported at a single timestep. 116

4.7 Scaling of the error of LO-AFQMC vs full ph-AFQMC, estimated at a time step of 2 Ha^{-1} (a) and 250 Ha^{-1} (b), with respect to the basis size (cc-pVTZ). Error bars are not plotted since the error is evaluated at a single timestep. Note that the error for each time step scales linearly, although there is notable variation as the walker distributions change. 117

4.8 Effect of performing operations for increasing numbers of walkers in parallel on a single GPU, for Fe(Cp)$_2$ in the cc-pVTZ-DK basis (508 basis functions, 200 determinants). The GPU-hours represent the time necessary to propagate 1656 walkers over 200 Ha^{-1}, evaluating the energy 2000 times in total per walker, on the Summit supercomputer at Oak Ridge National Laboratory. 119

4.9 Plot of the mean and statistical error of the LO error of a set of 20 independent walkers for the system C_8H_8. The error exhibits a variance which contributes minimally to the statistical error, and does not grow significantly in imaginary time. The mean error in this case was 0.002 mHa, with a standard deviation of 0.2 mHa, representing a favorable cancellation of error over imaginary time. 131

5.1 Systematic convergence with α for the open shell atoms of AFQMC with the constraint as specified in Eq. 5.12, versus ph-AFQMC. All calculations use CASSCF trials including the valence S and P orbitals in the active space. The grey bar indicates chemical accuracy, or 1 kcal/mol. 138

5.2 Systematic convergence of lc- and ph- AFQMC for C_2 with an increasing number of determinants retained in a heat bath CI trial (active space of 8e16o). In each case the determinants of a specified number with the highest weight in the wave function are kept. 140

5.3 A sample of walker steps from a lc-AFQMC trajectory which cross the negative real axis, resulting in the elimination of the walker by setting the weight to zero. The tails of the arrows represent the overlap of the walker with the trial before the propagation step, while the head represents the overlap after. lc-AFQMC is stable due to the annihilation of walkers whose overlaps with respect to the trial wave function cross the negative real axis, including the origin, from either side. 141

5.4 Distribution of overlaps in the complex plane with respect to the trial wave function for 256 walkers from 100 to 120 Ha^{-1} in a lc-AFQMC calculation. Note the effect of the linecut on the space surrounding the negative real axis. 141

5.5 Distribution of the norm and phase of overlaps with respect to the trial wave function for 256 walkers from 100 to 120 Ha^{-1} in a lc-AFQMC calculation. 142

5.6 Error (kcal/mol) of both phaseless and linecut AFQMC for the dissociation of the H_4 molecule in the STO3G basis using a UHF trial. The grey bar indicates chemical accuracy, or < 1 kcal/mol error. 143

5.7 Error (kcal/mol) of both phaseless and linecut AFQMC for the dissociation of the H_4 molecule in the cc-pVQZ basis using a UHF trial. The grey bar indicates chemical accuracy, or < 1 kcal/mol error. 144

5.8 Error (kcal/mol) of both phaseless and linecut AFQMC for the dissociation of the N_2 molecule in the STO3G basis using a UHF trial. The grey bar indicates chemical accuracy, or < 1 kcal/mol error. 144

5.9 Trajectories of ph-AFQMC and lc-AFQMC using the exact trial for N_2 at 1.5 . ph-AFQMC is in blue, with lc-AFQMC in a semi-transparent red overlaid on top. . 145

5.10 Error of AFQMC trajectories versus exact energy for FeO (def2-SVP, 864 determinants) using linecut and free projection AFQMC. The two are statistically equivalent (although run with different random walks) until the linecut constraint activates at around 5 Ha^{-1}, at which point fp-AFQMC encounters the sign problem and loses signal, and lc-AFQMC incurs a large bias. Both trajectories use 5000 walkers. . . 147

5.11 Error of AFQMC trajectories versus exact energy for FeO (def2-SVP, 864 determinants) for $X = 0$ and $X = 3$, where we observe the minimum of statistical error. By partially releasing the constraint we delay the onset of both the linecut bias and sign problem noise is delayed for long enough in imaginary time to obtain time-averaged results within chemical accuracy of the exact answer. 148

5.12 Correlation energy per N_2 atom for both phaseless and linecut AFQMC (cc-pVDZ) using a 0.05 Ha^{-1} timestep. 152

List of Tables

1.1 Mean absolute errors (MAE), mean signed errors (MSE), and maximum errors (MaxE) [kcal/mol] for dihydrogen complexes. CC refers to DLPNO-CCSD(T). . . 11

1.2 Mean absolute errors (MAE), mean signed errors (MSE), and maximum errors (MaxE) [kcal/mol] for aqua complexes. CC refers to DLPNO-CCSD(T). 13

1.3 Mean absolute errors (MAE), mean signed errors (MSE), and maximum errors (MaxE) [kcal/mol] for ammonia complexes. CC refers to DLPNO-CCSD(T). . . . 15

1.4 Mean absolute errors (MAE), mean signed errors (MSE), and maximum errors (MaxE) [kcal/mol] for carbonyl complexes. CC refers to DLPNO-CCSD(T). 16

1.5 Mean absolute errors (MAE), mean signed errors (MSE), and maximum errors (MaxE) [kcal/mol] for miscellaneous complexes. CC refers to DLPNO-CCSD(T). . 17

1.6 Mean absolute errors (MAE), mean signed errors (MSE), and maximum errors (MaxE) [kcal/mol] of ph-AFQMC, DLPNO-CCSD(T), and DFT results and other methods for the 34 molecule subset shown in Fig. 1.1. The values are sorted by MAE. The ph-AFQMC deviations incorporate both the experimental uncertainty and the statistical uncertainty. 18

2.1 The expected ground state term symbols of the metallocene or acac complexes in III or II oxidation states. 33

2.2 Vertical ionization energies as a function of metallocene type and AFQMC 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. The way the uncertainties for MAE's and MSE's are calculated is described in the SI of Ref. 2. 42

2.3 DFT (QZ basis) and DLPNO-CCSD(T_0) (extrapolated with TZ/QZ basis sets) 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. 44

2.4 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. . 46

2.5 ph-AFQMC 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. . . . 48

2.6 DFT (QZ basis) and DLPNO-CCSD(T_0) (extrapolated with TZ/QZ basis sets) 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. 51

2.7 DFT (QZ basis) and DLPNO-CCSD(T_0) (extrapolated with TZ/QZ basis sets) 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. 52

2.8 Computed reduction potentials (V) using the experimentally derived differential solvation energies (E_{solv}(II) - E_{solv}(III)) (kcal/mol)[116–118] as a function of metallocene and methodology for our best ph-AFQMC method as well as the best performing functional from above for the adibatic ionization energy of metallocenes B3LYP*. The mean absolute errors (MAE), maximum errors (MaxE), root-mean-square deviations (RMSD), and the mean signed errors (MSE) are included. The uncertainty on errors do not incorporate the uncertainty on the experimental differential solvation energy 54

3.1 AFQMC results from various trial wavefunctions for the adiabatic triplet energy of benzonitrile, in eV. Parentheses denote statistical error of AFQMC, i.e. 3.61(6) denotes 3.61 ± 0.06. 68

3.2 A comparison in eV of DLPNO-CCSD(T), KS-DFT, and TD-DFT results for T1 of anthracene derivatives, including mean absolute deviation (MAD), mean signed deviation (MSD), and maximum deviation (Max) versus AFQMC/U. Both DLPNO-CCSD(T) and TD-B3LYP have an MAD below the average statistical error of AFQMC (0.09 eV), although TD-B3LYP exhibits a higher maximum deviation of 0.215 eV vs AFQMC. 69

3.3 A comparison in eV of DLPNO-CCSD(T), KS-DFT, and TD-DFT results for T1 of a set of substituted BTD and BSeD compounds benchmarked against AFQMC/CAS, including mean absolute deviation (MAD), mean signed deviation (MSD), and maximum deviation (Max) versus AFQMC/CAS. All methods have significantly higher maximum deviations from AFQMC than was found for the anthracenes, as is expected given the larger degree of electron correlation observed in these compounds. DLPNO-CCSD(T) and TD-B3LYP again have the lowest and second lowest MADs vs AFQMC, respectively. 71

3.4 CAM-B3LYP TD-DFT results for S1 energies in eV for all substituents, including a subset of CN substituted tetracenes, with available experiments. The MAE was found to be 0.056 eV for all available experiments. Structures for the tetracenes can be found in the SI. 73

3.5 Experimental (phosphorescence) estimations of the 0-0 triplet energy, compared against computational predictions for the monomer (CBS with dielectric solvation corrections); BTD is severely overestimated by all methods, whereas CN-BTD is overestimated by all but KS-B3LYP. In the case of Ph-BTD, AFQMC/C agrees to within statistical error, and indeed exhibits the lowest deviation from experiment, closely followed by KS-B3LYP. 78

3.6 A comparison in eV of AFQMC/U, DLPNO-CCSD(T), KS-DFT, and TD-DFT results for the T1 energies of anthracene derivatives. Parentheses denote statistical error; the value to the right of "/" denotes the value extrapolated to the basis set limit. 87

3.7 A comparison in eV of AFQMC/CAS, DLPNO-CCSD(T), KS-DFT, and TD-DFT results for T1 of a set of substituted BTD and BSeD compounds. Parentheses denote statistical error; the value to the right of "/" denotes the value extrapolated to the basis set limit. 87

3.8 Mean Absolute Deviations for all experimental values (DCA, DPA, BTD, CN-BTD, MeO-BTD, Ph-BTD 88

3.9 Dielectric corrections to the gas-phase energy to account for solvent (toluene) effects, within the CPCM formalism. All corrections were found at the B3LYP/TZ level, using ORCA. Note that tthe solvation correction is minimal in this low-dielectric solvation model, as expected, aside from the MeO-substituted diazoles, which exhibit the strongest charge-transfer due to the electron donating character of MeO; the corrections were -0.15 and -0.13 eV for MeO-BTD and MeO-BSeD, respectively. 89

3.10 Triplet energies including both vibrational and solvent (toluene) corrections within the CPCM formalism. 89

3.11 Electronic bond energies (eV) for the anti-square dimer (at the wB97X-V/cc-pVTZ level of theory), along with a comparison of triplet energies (eV) for the dimer(calculated by DLPNO-CCSD(T)) and monomer (AFQMC) vs experiment. All species exhibit chalcogen bonding, with CN-BTD and BTD exhibiting the highest and second highest bonding strengths, respectively. The phosphorescence for Ph-BTD lacked the fine structure necessary to specify the 0-0 transition energy, and so we report two estimates, one based on the phosphorescence maximum, and one (in parentheses) determined from the 1/2 intensity of the phosphorescence maximum. Triplet energies for the dimers, calculated with DLPNO-CCSD(T), agree with experiment for BTD and CN-BTD, within 0.04 and 0.1 eV, respectively. 90

3.12 Dimerization free energies evaluated using KS-DFT at the B3LYP/cc-pVTZ level. Note that a negative free energy favors the formation of a dimer; all species are monomers at room temperature, but BTD and CN-BTD transition to a favored dimer state at low temperatures. Inspection of the geometries reveal in-plane dimerization for all but Ph-BTD, which is sterically limited to out of plane, likely resulting in the higher observed dimerization free energies. 91

3.13 CAM-B3LYP TD-DFT results for S1 energies in eV. 91

3.14 AFQMC results for both S1 and T1 of Anthracene using Correlated Sampling . . . 94

4.1 List of key intermediates in the LO-AFQMC evaluation of the 2-body energy using SMW, including the memory and computational scaling. The nomenclature $\{X\}_{a,b,c}$ indicates that the scaling listed is for the formation of the set of intermediates X for all valid indices a, b, c. 108

4.2 Errors in total energy for ferrocene (cc-pVTZ-dkh) vs an identical AFQMC calculation with full ERIs for a series of SVD truncation thresholds T_{SVD}. $T_{SVD} = 10^{-4}$ provides a good balance between memory efficiency and accuracy. 112

4.3 Total energies (Ha) for the singlet (S_0) and triplet (T_1) states of the series of polyacenes, as calculated by AFQMC and an identical LO-AFQMC calculation with $T_{SVD} = 10^{-4}$. 114

4.4 ST gaps in kcal/mol for the polyacenes (cc-pVTZ). The experimental values and ph-AFQMC with a UKS trial were obtained from Ref 72; the experimental values are corrected to compare to electronic energies from AFQMC using vibrational free energy corrections from B3LYP as described in Ref 72. 114

4.5 Total Atomization Energies in kcal/mol for the platonic hydrocarbon cages; CCSD(T)/CBS values were taken from Ref. 291. LO Error refers to the difference between the LO-AFQMC result and an identical ph-AFQMC calculation run without compression. 117

4.6 List of key intermediates in the formation of the Green's function using SMW, including the formal memory and computational scaling of forming them. N^a_{occ} and N^a_{unocc} refer to the number of occupied orbitals in the active space and number of unoccupied orbitals in the active space, respectively. 129

5.1 Systematic release of the linecut constraint for C_2 using a trial with the ten highest weight CI determinants in the wave function. All calculations use 1920 walkers and propagate for 300 Ha^{-1}. 146

5.2 Errors (kcal/mol) of phaseless and linecut AFQMC at long imaginary times for FeO (def2-SVP) using different numbers of determinants from a selected CI trial with a full active space. lc-AFQMC performs significantly worse than ph-AFQMC. Note that due to cost we only ran lc-AFQMC at 99% CI weight (26,071 determinants) for 60 Ha^{-1}, hence the higher reported statistical error. 147

5.3 Systematic release of the linecut constraint for FeO using 864 determinants in the trial. All calculations use 5000 walkers and are averaged over 4 to 20 Ha^{-1}. Note the initial decrease in statistical error with X, due to delaying the onset of the linecut bias, followed by an increase at $X = 4$ due to the sign problem. 149

Acknowledgements

I feel very fortunate to have been surrounded by an extraordinary group of scientists and humans over the past five years. First, I owe a lot of my scientific growth to the guidance of the two amazing scientists that have served as my advisors, David Reichman and Richard Friesner. Dave's insight and curiosity into the fundamental questions underlying electronic structure, immaculate knowledge of the literature, and open approach to science is inspiring. Meanwhile, Rich's down to earth practicality and deep intuition has stopped me from going headfirst into innumerable dead ends. I am forever grateful.

Additionally, Shiwei Zhang has been one of the best and most patient sources of knowledge and insight into AFQMC one could hope for. I look back with fondness at all of the times I visited the Flatiron Institute and threw ideas at Shiwei, every time of which I returned with a sharper idea of why it would (or much more likely, why it wouldn't) work. Thank you also to Timothy Berkelbach and Angelo Cacciuto for taking the time to sit on my committee for all these years, and for giving such wonderful advice. To Sandeep Sharma, thank you for sitting on my defense committee!

My Ph.D. started off on the right foot with the help of two of the best co-workers and friends, James Shee and Ben Rudshteyn. Whether it was Ben's two hour long introduction to all things bash that started me off confidently running calculations for the first time, James patiently going through the code he and Evan Arthur wrote so that I could understand where to start, going to the gym, playing friesner-ball in the hallway, or just hanging out, these two made my first two years (and more!) a blast.

PJ and Pierre, my lab-mates in two different groups, have been with me for the entire five years. PJ, thank you for trying to keep me in shape throughout the stresses of graduate school, for letting me borrow your quantum notes every week, for uncountable lunches, and for generally being there for me. Pierre, thank you for the many, many times where you saved me from the evils of computer mumbo jumbo, for introducing me anew to so many fictional worlds, and, also, for generally being there for me. You two are the best of friends.

To the co-workers whom I have had the pleasure to start working with after that first year; Hung, Lisa, Sibali, Joonho, Ankit, and the entire Reichman and Friesner groups; it is because of you that I've been able to get anything done! You're also why I now associate going into the office with having fun.

To my friends at Columbia - PJ, Pierre, Arden, Allyson, Davida - I'm so grateful to have gotten to spend this time with you by my side. Here's to many more years of getting through life, and try not to move too far away!

Finally, thank you to all of my friends and family - the list is long - who have been there for me through the years. I am thankful to say that it's the people in my life that make it worth living, and I love each and every one of you. Most of all, thank you to my best friend and wife, Sara. You've been there for me every step of the way, and I am more grateful than you know.

Dedication

To Sara, for whom the sun doth shine and the dogs doth bark

Preface

This thesis presents a compilation of works done by the author throughout the years 2018-2023, which can additionally be found published in Refs. [1,2,3,4]. We have decided to omit other published work (Refs. [5,6,7,8]) primarily for brevity and cohesiveness.

Chapter 1: Predicting Ligand-Dissociation Energies of 3d Coordination Complexes with Auxiliary-Field Quantum Monte Carlo

Reprinted with permission from B. Rudshteyn *et al.*, *J. Chem. Theory Comput.*, vol. 16, no. 5, 2020. Copyright 2020 American Chemical Society.

1.1 Introduction

The unique electronic structure of transition metals enables a rich variety of chemical reactivity, harnessed in systems ranging from those found in the fields of chemical catalysis[9], biology[10] and materials science[11]. The presence of multiple quantum states within an accessible energy range allows for reaction mechanisms involving sequential redox events and subtle transformations between spin-states, e.g. in clusters of Mn atoms in Photosystem II (PSII) or Fe and Mo atoms in nitrogenases[12–14]. Furthermore, the coordination of small molecules to single metal ions is an important motif in drug design[15], and the correlations exhibited in the copper oxide layers of cuprate materials play a central role in the phenomenon of high-temperature superconductivity[16,17].

Ab initio modeling has the potential to yield essential insights into these transition metal systems. However, exact methods scale exponentially with system size and are thus only applicable to small molecules. Many groups have used density functional theory (DFT) to examine the electronic structure and reaction mechanisms of coordinated transition metal complexes, including the active sites of PSII[13,14] and cytochrome P450,[18,19] catalysts for water oxidation,[20] CO_2 reduction,[21] and sensitizers for optical upconversion.[22] However, there are a number of uncertainties which may cast doubt upon their conclusions, chief among them possible errors due to electron self-interaction and strong correlation. Furthermore, as the majority of parameterized density functionals and dielectric continuum solvation models have been trained on organic compounds (e.g.

the ωB97X-V[23] and ωB97M-V[24] functionals and the SMD solvation model[25]), it is reasonable to suspect the accuracy of the resulting predictions in the domain of transition metal chemistry.

The pronounced lack of reliable and precise gas-phase experimental data for realistic transition metal systems, as illustrated by recent theoretical benchmarking studies, exacerbates these issues.[26–36] This scarcity of experimental measurements is in stark contrast to the large amount of reliable experimental values for organic molecules, which has enabled very accurate parameterizations of DFT functionals and a thorough validation of methods such as CCSD(T), which can readily achieve ~1 kcal/mol accuracy for typical organic molecules[37].

The accuracy of CC methods, most frequently CCSD(T), is often assumed to carry over to transition metal systems, as evidenced by a number of studies that have attempted to draw conclusions about the accuracy of DFT by comparing against reference CC values.[38–42]. However, the reliability of CC methods for transition metal systems, even when multireference effects are approximated, has been the subject of vigorous debate, as illustrated by recent studies on transition metal diatomic-ligand systems[31,43–48]. De Oliveira-Filho and co-workers found that even multireference CCSD(T) could not predict the bond dissociation energies (BDEs) for some diatomics accurately with respect to experimental measurements. A recent study by Head-Gordon and co-workers found that high levels of CC, up to CCSDTQ, are required for chemical accuracy against an exact method known as Adaptive Sampling Configuration Interaction (ASCI) results, albeit in a small basis set.[49] Wilson and co-workers collected a set of 225 heats of formation for compounds with first row transition metal atoms.[31] They found good performance for their composite CC scheme vs. a subset of experimental data with small uncertainties, but the mean absolute error (MAE) of around 3 kcal/mol may be insufficient for many chemical applications. Reiher and co-workers considered transition metal ligand-dissociation energies of very large molecules and showed that a localized variant of CCSD(T) utilizing domain based pair natural orbitals (DLPNO-CCSD(T))[50,51] resulted in pronounced errors, e.g. ~ 9.3 kcal/mol for the cleavage of a Cu complex.[52]

An alternative benchmarking approach involves filtering out strongly correlated cases with multireference diagnostics, and benchmarking DFT against CC methods only for the single-reference subset of molecules. Hansen, Checinski, and co-workers developed the MOR41 test set of organometallic reactions of medium-large size. They removed open-shell, multi-reference cases (with, e.g., FOD and T1 diagnostics). Recently, the properties of a set of transition metal atoms and oxide diatomics, in which strongly multi-reference cases were removed, were predicted by a large number of *ab initio* methods.[48] In our view, this strategy is less than ideal not only because a large subset of relevant chemistry is excluded, but moreover because the utility of affordable multi-reference indicators has increasingly been called into question. Indeed, studies have found mixed success for different kinds of multireference diagnostics[43–46,53] making it hard to judge *a priori* when single-reference methods would be appropriate.

In this work, we assemble a test set of gas-phase ligand-dissociation measurements with low reported experimental uncertainties. On this set we use auxiliary field quantum Monte Carlo with the phaseless constraint (ph-AFQMC),[54,55] accelerated by a correlated sampling technique[56] and our implementation on graphical processing units[57]. We have shown that this method yields robust accuracy for the ionization potential of transition metal atoms[57] and the dissociation energy of transition metal-containing diatomics[47]. The present study marks a large step forward, to more relevant transition metal-containing systems. We demonstrate that ph-AFQMC with correlated sampling yields accurate BDE predictions for various tetrahedral, square planar, and octahedral complexes containing first row transition metal atoms and ligands including dihydrogen, chloride, dinitrogen, aqua, ammonia, carbonyl, and formaldehyde. We then validate the performance of a representative set of DFT functionals and the DLPNO-CCSD(T) method. Consistent with our expectation, we find that single-reference methods such as DFT and the CC hierarchy perform better for coordinated metal compounds compared to the case of diatomic dissociation (as ligand coordination can lower the degree of degeneracy of the metal atomic *d* orbitals). However, we demonstrate that ph-AFQMC still produces a significant improvement in terms of MAE and maximum error (MaxE).

Our results show that ph-AFQMC can consistently produce benchmark-quality results, and with a computational cost which scales as a low polynomial with system size (excluding the cost of obtaining the CASSCF trial wavefunctions). This method will extend accurate reference datasets for future benchmarking studies of approximate methods such as DFT and accurate classical potentials for transition metal ions. In addition, the level of accuracy of the widely-employed quantum-chemical methods included in this study provides a sense of the accuracy to be expected for calculations on similar 4- and 6- coordinated $3d$ metal complexes that are ubiquitous in fields such as biology and catalysis.

1.2 Selection of Experimental Data

We selected gas-phase experimental BDE data with less than or equal to 2.0 kcal/mol uncertainty from the recommended values in the handbook compiled by Luo[58]. Most of the measurements can also be found in the work by Rodgers and Armentrout.[59] For $TiCl_4$, Hildenbrand's updated experimental measurement has been used.[60] The average uncertainty for the molecules included in the present test set is 1.03 kcal/mol. Most of the measurements were performed with the threshold collision-induced dissociation technique except for $[Ni(H_2O)_6]^{2+}$, $TiCl_4$, $CrCO_5H_2$ and $V(H_2O)(H_2)_3$ which were measured with blackbody infrared radiative dissociation, effusion beam mass spectroscopy, transient infrared spectroscopy for kinetic analysis and temperature-dependent equilibrium, respectively. The latter technique was used for all other H_2 complexes as well. The selected compounds are depicted schematically in Fig. 1.1. These experimental data are mostly extrapolated to 0 K, and can therefore be directly compared with quantum-chemical calculations. The two exceptions are $TiCl_4$ and $CrCO_5H_2$, which are measured at 298 K. All the metal complexes have +1 net charge, except for $[Ni(H_2O)_6]^{2+}$, $TiCl_4$, and $CrCO_5H_2$. The full list of reactions is given in the Supporting Information (SI) of Ref. 1.

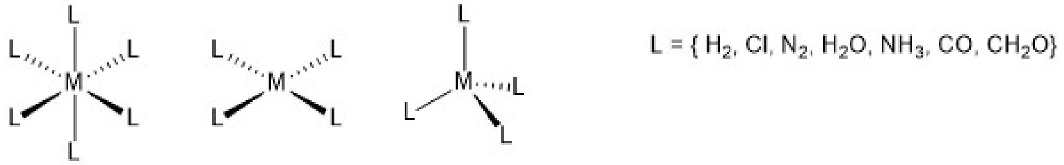


Figure 1.1: The types of transition metal compounds studied. M can be any $3d$ transition metal from Ti to Cu.

1.3 Computational Details

The geometries, reorganization energies (*vide infra*), and enthalpic corrections (just the zero-point energy (ZPE) for cases where the 0 K extrapolated experiment is available, as discussed above) were obtained with DFT calculations with the B3LYP functional[61–63] and cc-pVTZ-dkh[64–67] basis set using the ORCA program package.[68] Details regarding occasional small imaginary frequencies and integration grids are given in Section IV of the SI of Ref. 1.

The DLPNO-CCSD(T) calculations were also done with ORCA using “TightPNO" localization parameters and the cc-pVxZ-dkh basis sets, x=T,Q, and are extrapolated to the complete basis set limit using the procedure built into ORCA,[68] as discussed in the SI of Ref. 1. The DKH2 relativistic correction was used for all DFT and CC calculations.[69]

Integrals for AFQMC were obtained with PySCF[70]. The exact-two-component (x2c) relativistic Hamiltonian[71] was used in place of DKH2. As in our previous work,[47,56,57,72] the imaginary time step for the AFQMC propagation, utilizing single precision floating point arithmetic, was 0.005 Ha^{-1}. The walker orthonormalization, population control, and local energy measurements occurred every 2, 20, and 20 steps, respectively. We utilized a modified Cholesky decomposition of the electron repulsion integrals with a cutoff of 10^{-5}. Walkers were initialized with the RHF/ROHF determinant.

The correlated sampling approach[56] can converge energy differences between similar states by employing a shared set of auxiliary fields for a short projection time, providing accurate results with smaller statistical errors vs uncorrelated AFQMC (the latter would need to run longer projections to reach the same statistical accuracy). This approach performs most efficiently when the ligand being removed is small, as indicated by our previous work in which the reduction in statistical error vs the uncorrelated approach was several times larger for MnH than for MnCl.[47] Similar behavior is found for the transition metal complex systems studied here, as shown in Fig. 1.2 for $[\text{Cu}(\text{H}_2)_4]^+$. In fact, correlated sampling may work better for these complexes than it did for the diatomics since << 50% of the system is being changed. Finally, we note that correlated sampling also can improve the accuracy of the predicted results in certain situations.[47,57]

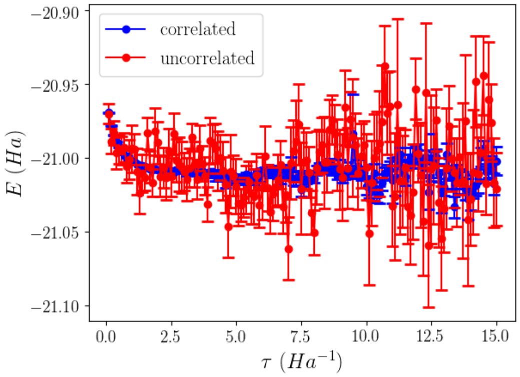


Figure 1.2: Correlated sampling ph-AFQMC calculations using Summit GPU's; statistical errors from correlated and uncorrelated sampling approaches are compared for the Cu-H_2 bond dissociation energy of the $[Cu(H_2)_4]^+$ molecule.

In the context of computing BDEs, our AFQMC calculations used correlated sampling for the difference in energy between the original coordination compound (M-L) and the species missing a ligand (M), i.e. the same geometry but with ghost basis functions centered around the positions of the missing nuclei that comprise the ligand. If the difference in energies was not converged before 15 Ha^{-1}, uncorrelated, separate AFQMC calculations are performed for the optimized structures of both states without ghost basis functions, using a population control scheme in which walkers with large weights are duplicated while those with small weights are randomly destroyed for the optimized structures of both states without ghost basis functions[73]. The isolated ligand (L) was also treated with the population control approach.

The BDE is given as follows:

$$BDE = (H(M) - H(M - L)) + H(L) - \lambda, \tag{1.1}$$

where H are enthalpies including the zero-point corrections and the nuclear repulsion energy. The reorganization energy, λ, is defined as the difference in energy between the product (complex with the ligand dissociated) in its optimal geometry and in the reactant geometry, optimized with the ligand, but with the ligand atoms deleted. λ is computed via DFT. The calculation of BDEs is illustrated in Fig. 1.3.

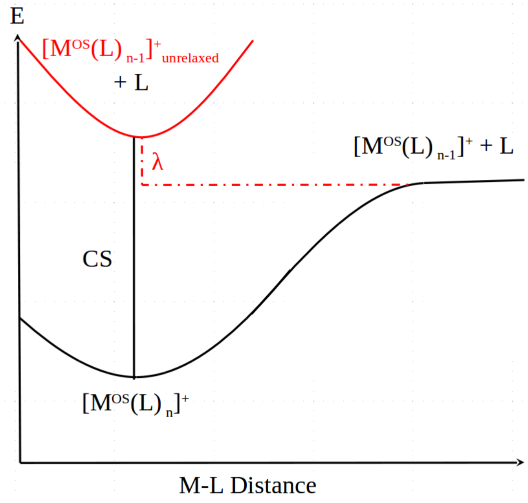


Figure 1.3: Schematic of BDE calculations performed in this work. OS abbreviates oxidation state, CS indicates the energy measured by the correlated sampling approach.

To give a sense of the required computational cost, a correlated sampling ph-AFQMC calculation for $[Fe(N_2)_4]^+$ took about 267 node hours on Summit, using a truncated CASSCF trial wave function containing 1195 determinants. This reflects the use of 20 repeats (i.e. independent trajectories with different random number seeds), each using 20 nodes with 6 GPU's each (each repeat ran for about 42 minutes).

The complete basis set limit for the ph-AFQMC calculations was estimated by extrapolation using DLPNO-CCSD(T) values with the cc-pVxZ-dkh basis sets, x=T,Q, using exponential and $\frac{1}{x^3}$ forms for the mean-field (i.e. UHF) and correlation energies, respectively, as in our previous work[47]. We used the equivalent cc-pVxZ\C auxiliary basis sets for the DLPNO approximations.

If the ph-AFQMC correlation energy with cc-pVTZ-dkh is significantly different from DLPNO-CCSD(T), or if comparison of the extrapolated value with experiment indicates a potential problem (our target accuracy is <3 kcal/mol, which has been referred to as “transition metal chemical accuracy"[74]), then full extrapolation within ph-AFQMC is performed utilizing both cc-pVTZ-dkh and cc-pVQZ-dkh basis sets (for dihydrogen or chloro compounds). In some cases, we instead extrapolate with a UHF trial-based ph-AFQMC procedure, which seems to be a good compromise between speed and accuracy (see Tables S4 and S5 in the supporting information of Ref. 1 for details).

Apart from the basis set extrapolations, the ph-AFQMC calculations utilized CASSCF trial wavefunctions. The size of the CASSCF trial wavefunction for the metal-containing species was automatically selected via the AVAS procedure where only those B3LYP ROKS orbitals that overlap significantly with the 3d and/or 4d atomic orbitals (from the minimal atomic basis set called "MINAO" as used by Knizia[75] or from the Atomic Natural Orbital (ANO-RCC) basis set) of the metal were included (as noted in the SI of Ref. 1)[76]. The single numerical overlap threshold parameter was used to generate sequentially larger active spaces to determine what active space size is needed to reach chemical accuracy.

The active space for the ligand was selected by either using the valence set of electrons and orbitals or using a large number for electrons and orbitals to ensure convergence. Typically >98% of the weight of the CI coefficients was retained. The active spaces were selected so that the active space for the reactant and product metal species were similar (either the same or off by 1 orbital and 2 electrons), which often requires the same AVAS threshold.

We compare ph-AFQMC with the B3LYP, M06[77], and PBE0[78] functionals since they are arguably the most popular, and B97 since this functional performed the best in our previous study.[47] To explore the performance of range-correction and the non-local correlation approach, we include the ωB97X-V functional.[23] We also consider the double hybrid functional, DSD-PBEP86. It is available in ORCA, and has been shown to perform very well,[79–82] accelerated by the resolution of identity (RI) approximation on the MP2 part. In this study, we used the "DSD-PBEP86/2013" functional, which has slightly different parameters than DSD-PBEP86, but refer to it as DSD-PBEP86 throughout the paper.

Since analytical gradients have not yet been implemented in ORCA for all of the functionals in this study, we decided to use B3LYP optimized geometries and performed single-point energy calculations. Grid and density-initialization choices are described in Section IV of the SI of Ref. 1.

For all DFT and HF (the latter is used as a reference wavefunction for DLPNO-CCSD(T)) calculations, we found it essential to perform a stability analysis to ensure that the lowest energy SCF solution was obtained.

1.4 Results and Discussion

The deviations of the computed BDEs from experiment are presented in Figs. 1.4 to 1.8. Values of the BDEs are given explicitly in Tables S1 and S6 in the supporting information of Ref. 1. Tables 1.1 through 1.6 show statistical metrics including Mean Signed Error (MSE), MAE, and MaxE for each ligand type, and ultimately for the entire test set.

1.4.1 Dihydrogen Complexes

In general, as shown in Fig. 1.4 and Table 1.1, the performance of ph-AFQMC is excellent for dihydrogen complexes (where the dihydrogen is the ligand being removed), including $[Ti(H_2)_4]^+$, $[Cu(H_2)_4]^+$, $[V(H_2)_4]^+$, $[V(H_2)_6]^+$, $[Co(H_2)_4]^+$, $[Ni(H_2)_4]^+$, $[Ti(H_2)_6]^+$, $[Co(H_2)_6]^+$, $[Fe(H_2)_6]^+$, $[Fe(H_2)_4]^+$, $[Cr(CO)_5H_2]^+$, $[Cr(H_2)_6]^+$, $[VH_2O(H_2)_3]^+$, and $[Cr(H_2)_4]^+$.

Table 1.1: Mean absolute errors (MAE), mean signed errors (MSE), and maximum errors (MaxE) [kcal/mol] for dihydrogen complexes. CC refers to DLPNO-CCSD(T).

	ph-AFQMC	CC	B3LYP	B97	M06	PBE0	ωB97X-V	DSD-PBEP86
MAE	0.85 ± 0.21	1.82	1.43	0.93	2.50	0.75	1.43	1.09
MSE	0.09 ± 0.21	1.75	-1.36	-0.67	1.94	0.33	1.08	1.04
MaxE	1.51 ± 1.36	7.54	3.29	2.05	4.68	2.91	8.08	8.49

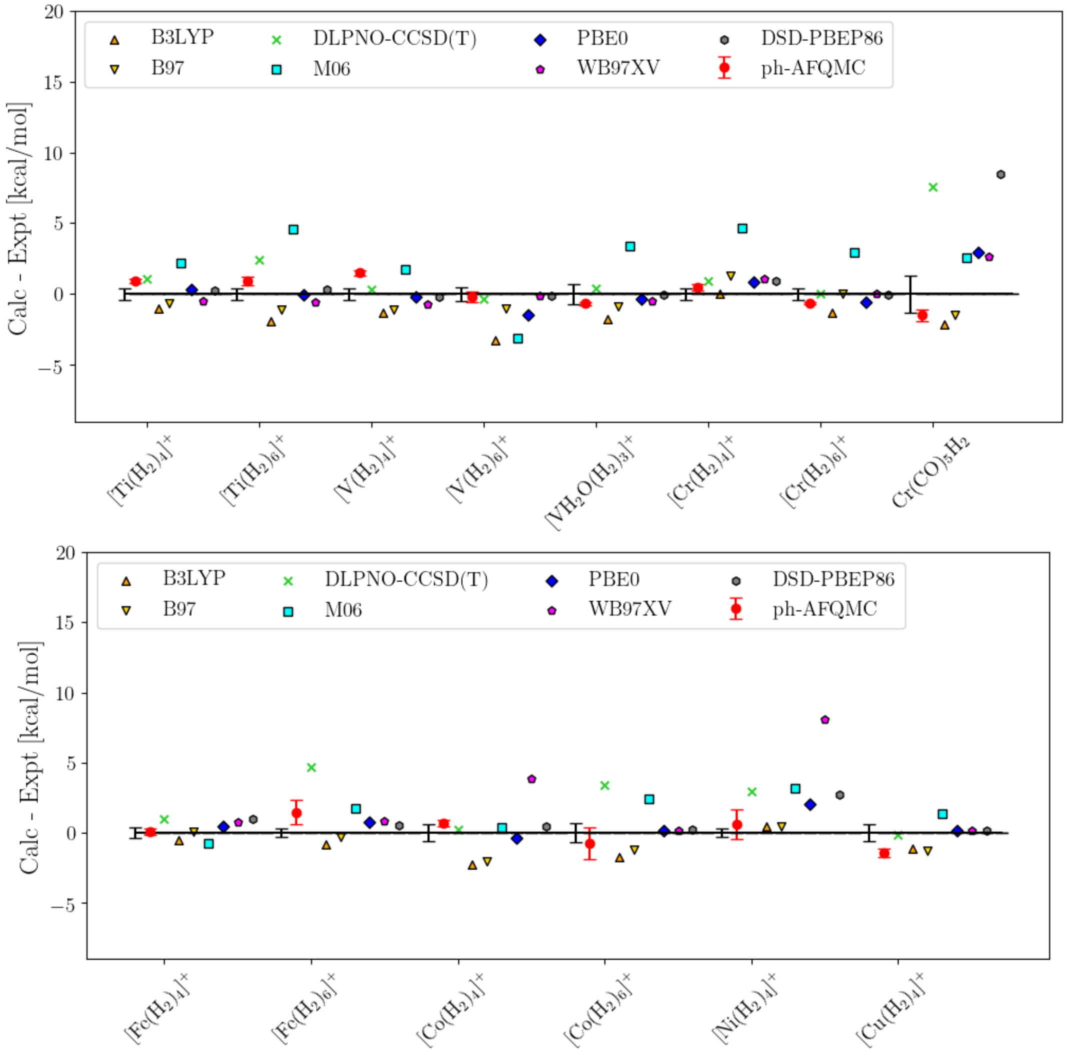


Figure 1.4: Deviations [kcal/mol] of computational methods for the dihydrogen set of bond dissociation reactions where the H_2 that leaves is given at the end of the formula.

The relatively small system sizes of these dihydrogen complexes renders the ph-AFQMC calculations affordable even with the QZ basis set. Therefore, for $[Ni(H_2)_4]^+$, which showed deviations > 2 kcal/mol (see the supporting information of Ref. 1), we opted to do the full TZ/QZ extrapolation entirely within ph-AFQMC, and found better agreement. In contrast, the scaling factor, i.e. the ratio between the correlation energies computed by ph-AFQMC and DLPNO-CCSD(T) at the TZ level was close to or more than 1.3 for $[Co(H_2)_6]^+$ and $[Fe(H_2)_6]^+$, a metric found in our previous work,[47] so we also did TZ/QZ extrapolation entirely within ph-AFQMC in these cases, leading to good agreement. In the SI of Ref. 1, we show that using ph-AFQMC/UHF to extrapolate gives similar results to the full treatment for the dihydrogen species.

M06 yields the largest MAE (2.5 kcal/mol) while B97, PBE0, and ph-AFQMC have MAEs less than 1 kcal/mol. While ph-AFQMC and most density functionals (DFs) perform reasonably well for $Cr(CO)_5H_2$, especially given the relatively large experimental uncertainty, DSD-PBEP86 and DLPNO-CCSD(T) are off by 6-8 kcal/mol. We note that in the next section DSD-PBEP86 is seen to over-stabilize all carbonyl complexes. ωB97X-V drastically overestimates the BDE of the $[Ni(H_2)_4]^+$ complex, with a deviation of 8.08 kcal/mol. Indeed, as will be shown, this functional over-stabilizes all Ni complexes.

1.4.2 Aqua Complexes

As shown in Fig. 1.5 and Table 1.2, ph-AFQMC also yields accurate results for the hexaaqua complex $[Ni(H_2O)_6]^{2+}$ and the tetraaqua complexes $[Cr(H_2O)_4]^+$, $[Ni(H_2O)_4]^+$, $[Ti(H_2O)_4]^+$, $[V(H_2O)_4]^+$, and $[Fe(H_2O)_4]^+$. While all other methods seem to overbind these complexes, as can be seen by large and positive MSEs, ph-AFQMC appears to predict the BDEs in a relatively balanced manner.

Table 1.2: Mean absolute errors (MAE), mean signed errors (MSE), and maximum errors (MaxE) [kcal/mol] for aqua complexes. CC refers to DLPNO-CCSD(T).

	ph-AFQMC	CC	B3LYP	B97	M06	PBE0	ωB97X-V	DSD-PBEP86
MAE	1.61 ± 0.84	3.70	2.72	2.61	5.65	3.20	4.25	3.81
MSE	0.89 ± 0.84	1.60	1.99	1.91	5.65	2.81	4.25	3.40
MaxE	2.96 ± 1.71	7.24	5.26	5.54	9.49	5.98	8.48	7.99

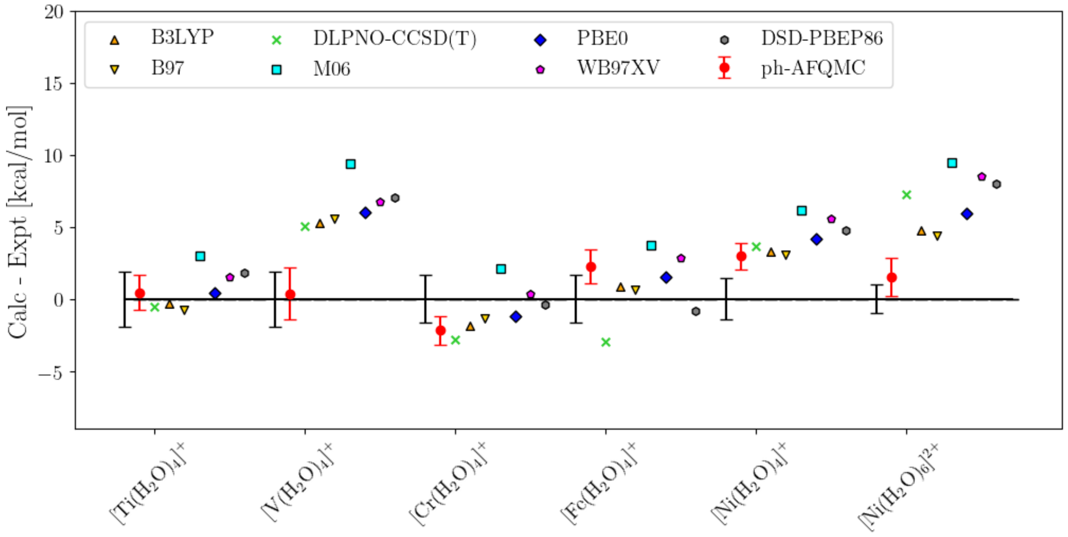


Figure 1.5: Deviations [kcal/mol] of computational methods for the aqua set of bond dissociation reactions where the H_2O that leaves is given at the end of the formula.

In the case of $[Ni(H_2O)_6]^{2+}$, the scaling factor was below 0.6, which indicates a poor match between the correlation energies of ph-AFQMC and DLPNO-CCSD(T). As full TZ/QZ extrapolation within ph-AFQMC was unaffordable in the present version of our code implementation due to prohibitively high required device memory, we opted to do the extrapolation with a single-determinant (UHF) trial based QMC in place of DLPNO-CCSD(T) and found good results.

Notably, all other methods overestimate the BDE for this molecule by at least 5 kcal/mol, well outside the reported experimental uncertainty. All DFs and DLPNO-CCSD(T) give errors in excess of 5 kcal/mol for this molecule. Similarly, we performed the extrapolation with ph-AFQMC/UHF for $[V(H_2O)_4]^+$, on the basis of disagreement of experiment rather than the scaling factor, and found that the deviation went from 4.03 ± 1.95 kcal/mol with the DLPNO-CCSD(T) extrapolation to 0.35 ± 2.63 kcal/mol with the ph-AFQMC/UHF extrapolation. The other methods have errors around 5-9 kcal/mol for this molecule. These findings suggest that these two species exhibit significant multireference character.

On average, as seen in Table 1.2, the accuracy of CC and DFT methods for metal-aqua complexes is similar with MAE's between 2.61 (B97) and 5.65 (M06) kcal/mol. The MAE of ph-AFQMC is 1.61 ± 0.84 kcal/mol, with a MaxE of 2.96 ± 1.71 kcal/mol found for the $[Ni(H_2O)_4]^+$ species. We note that all methods overestimate the BDE of this molecule, although not by a huge amount, especially in light of the experimental error bars. It is thus possible that the experimental value for this case should be reinvestigated.

1.4.3 Ammonia Complexes

Fig. 1.6 and Table 1.3 summarize the performance of the computational methods for the tetraammonia complexes: $[Co(NH_3)_4]^+$, $[Ni(NH_3)_4]^+$, $[Mn(NH_3)_4]^+$, $[Cu(NH_3)_4]^+$, and $[Fe(NH_3)_4]^+$.

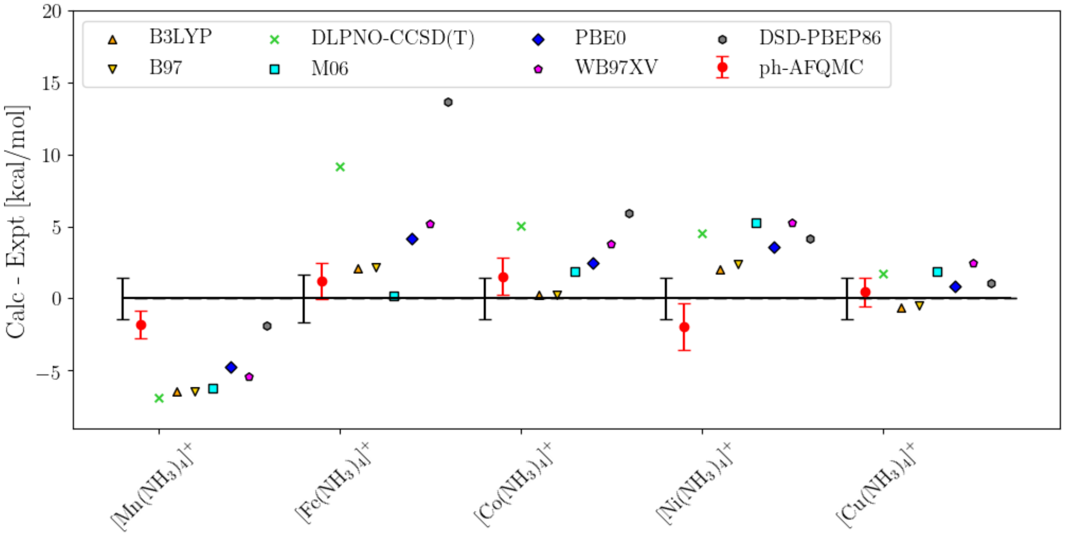


Figure 1.6: Deviations [kcal/mol] of computational methods for the aqua set of bond dissociation reactions where the NH_3 that leaves is given at the end of the formula.

$[Mn(NH_3)_4]^+$ is a difficult case for all methods. DSD-PBEP86 and ph-AFQMC, with deviations of ∼2 kcal/mol, performed better compared to other methods which showed errors of ∼6 kcal/mol. This reaction involves the only 2 molecules (i.e. $[Mn(NH_3)_4]^+$ and $[Mn(NH_3)_3]^+$) where we had to run separate ph-AFQMC calculations with population control because the imaginary

Table 1.3: Mean absolute errors (MAE), mean signed errors (MSE), and maximum errors (MaxE) [kcal/mol] for ammonia complexes. CC refers to DLPNO-CCSD(T).

	ph-AFQMC	CC	B3LYP	B97	M06	PBE0	ωB97X-V	DSD-PBEP86
MAE	1.39 ± 0.87	5.46	2.29	2.36	3.09	3.15	4.44	5.36
MSE	-0.12 ± 0.87	2.71	-0.55	-0.42	0.60	1.25	2.26	4.61
MaxE	1.95 ± 2.16	9.15	6.48	6.45	6.22	4.74	5.45	13.69

trajectories were not convincingly equilibrated by 15 β. Additionally, there were many CAS convergence issues that prevented us from running larger CASSCF active spaces to check the convergence. Further investigation will be required. DLPNO-CCSD(T) and the remaining DFs perform particularly poorly for this molecule with errors around or above 5 kcal/mol.

We note that $[Ni(NH_3)_4]^+$ is another case for which basis set extrapolation with ph-AFQMC/UHF reduced the deviation from experiment. As before, this may indicate multireference character, which causes all other methods to significantly overbind the ammonia ligand.

Overall, ph-AFQMC, B3LYP, B97, and M06 have notably small MSEs. ph-AFQMC is outstanding here with respect to MAE (1.39 ± 0.87 kcal/mol) and MaxE (1.95 ± 2.16 kcal/mol) while other methods show a MaxE around 6-14 kcal/mol for these complexes. DLPNO-CCSD(T) and DSD-PBEP86 showed the largest deviations with MAEs of 5.46 and 5.36 kcal/mol, respectively. They show extreme errors for $[Fe(NH_3)_4]^+$ in particular, with MaxEs of 9-14 kcal/mol.

1.4.4 Carbonyl Complexes

As shown in Fig. 1.7 and Table 1.4, ph-AFQMC also performed well for the species with all carbonyl ligands: $[Ti(CO)_6]^+$, $[Ni(CO)_4]^+$, $[Cu(CO)_4]^+$, $[Ti(CO)_4]^+$, $[Fe(CO)_4]^+$, and $[V(CO)_6]^+$. In particular, ph-AFQMC is the only method to predict a BDE close to the experimental value for $[Ti(CO)_6]^+$ (although B3LYP is just outside the AFQMC statistical error bars).

Table 1.4: Mean absolute errors (MAE), mean signed errors (MSE), and maximum errors (MaxE) [kcal/mol] for carbonyl complexes. CC refers to DLPNO-CCSD(T).

	ph-AFQMC	CC	B3LYP	B97	M06	PBE0	ωB97X-V	DSD-PBEP86
MAE	0.87 ± 0.72	2.65	0.83	1.71	4.99	3.43	2.35	7.99
MSE	0.85 ± 0.72	2.18	0.52	1.71	4.99	3.43	2.35	7.99
MaxE	2.39 ± 1.46	6.07	2.64	3.80	10.02	5.90	4.88	12.68

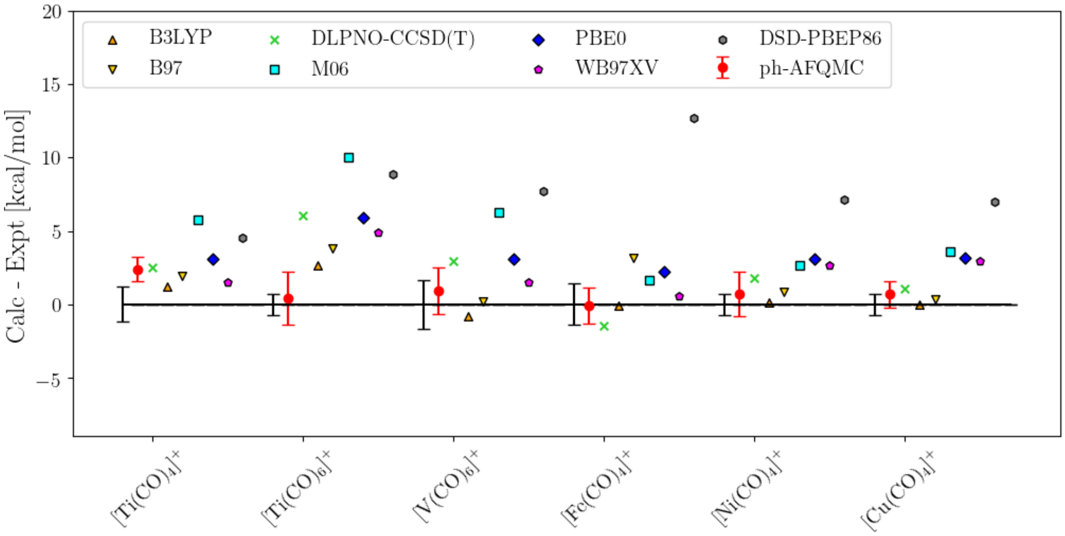


Figure 1.7: Deviations [kcal/mol] of computational methods for the carbonyl set of bond dissociation reactions where the CO that leaves is given at the end of the formula.

DSD-PBEP86 gives an extremely large deviation of 12.68 kcal/mol for $[Fe(CO)_4]^+$, and in fact overpredicts all carbonyl species in this set, with an MAE and MSE of ~ 7.99 kcal/mol. M06 has the second largest MAE (4.99 kcal/mol) and MaxE (10.02 kcal/mol for $[Ti(CO)_6]^+$) among all methods. For these carbonyl complexes, both ph-AFQMC and B3LYP showed outstanding performance with balanced predictions (low MSEs), MAEs of < 1 kcal/mol, and MaxEs of ~ 2.5 kcal/mol.

In the case of $[Ti(CO)_4]^+$, all methods predict BDEs above the experimental measurement. We therefore suggest, for a future study, that the experimental value be examined carefully.

Table 1.5: Mean absolute errors (MAE), mean signed errors (MSE), and maximum errors (MaxE) [kcal/mol] for miscellaneous complexes. CC refers to DLPNO-CCSD(T).

	ph-AFQMC	CC	B3LYP	B97	M06	PBE0	ωB97X-V	DSD-PBEP86
MAE	1.07 ± 1.19	2.45	2.37	0.89	5.72	1.56	3.80	4.66
MSE	-0.37 ± 1.19	2.45	-1.35	0.27	5.72	0.17	-2.01	2.76
MaxE	2.16 ± 2.36	4.29	4.12	1.54	10.18	2.59	5.15	6.37

1.4.5 Miscellaneous Complexes

As can be seen in Fig. 1.8, ph-AFQMC continues to predict consistently accurate BDEs for these three complexes. While a statistical analysis of three compounds is likely not rigorously meaningful, we nonetheless provide a summary in Table 1.5, for completeness.

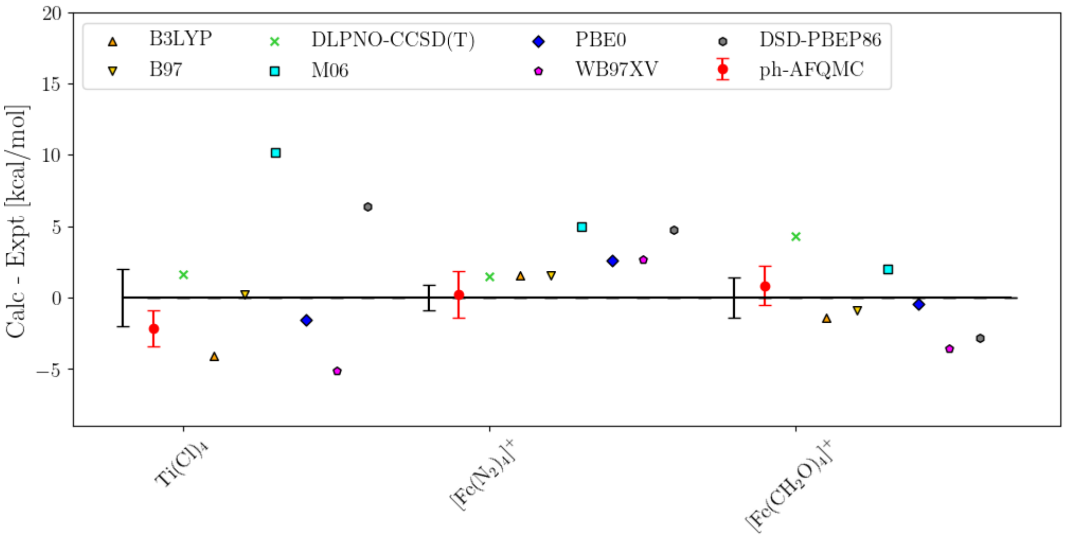


Figure 1.8: Deviations [kcal/mol] of computational methods for the other reactions where the ligand that leaves is given at the end of the formula.

The experimental uncertainty corresponding to the measured Ti(Cl)$_4$ BDE is the highest among the molecules included in this study, at 2 kcal/mol. Most of the methods give reasonable performance except DSD-PBEP86, M06 and ωB97X-V. The first two overestimated the BDE by ~ 6-10 kcal/mol while the latter underestimated it by 5.15 kcal/mol.

We note that all DFT methods overestimate the BDE of [Fe(N$_2$)$_4$]$^+$, with M06 and DSD-PBEP86 yielding deviations of around 5 kcal/mol.

The formaldehyde ligands make $[Fe(CH_2O)_4]^+$ the largest molecule studied in this work. ωB97X-V and DSD-PBEP86 yield deviations of ~ -3 kcal/mol while DLPNO-CCSD(T) yields of a deviation around ~ 3 kcal/mol.

1.4.6 Performance for the Entire Test Set

The statistical performance of each computational method over all ligand types is summarized in Table 1.6. We note that the average experimental uncertainty is 1.03 kcal/mol.

Table 1.6: Mean absolute errors (MAE), mean signed errors (MSE), and maximum errors (MaxE) [kcal/mol] of ph-AFQMC, DLPNO-CCSD(T), and DFT results and other methods for the 34 molecule subset shown in Fig. 1.1. The values are sorted by MAE. The ph-AFQMC deviations incorporate both the experimental uncertainty and the statistical uncertainty.

	MAE	MSE	MaxE
ph-AFQMC	1.09 ± 0.27	0.30 ± 0.27	2.96 ± 1.71
B97	1.57	0.33	-6.45
B3LYP	1.76	-0.32	-6.48
PBE0	2.08	1.43	5.98
ωB97X-V	2.74	1.77	8.48
DLPNO-CCSD(T)	2.89	2.00	9.15
DSD-PBEP86/2013	3.73	3.36	13.69
M06	3.87	3.27	10.18

ph-AFQMC, B97, and B3LYP have near-zero MSEs, while all other methods systematically overestimate the BDEs. ph-AFQMC outperforms all DFT functionals and DLPNO-CCSD(T), with an MAE of 1.09 ± 0.27 kcal/mol and MaxE of 2.96 ± 1.71 kcal/mol. DLPNO-CCSD(T) performs worse than most of the hybrid functionals in the study, with MAE and MaxE of 2.89 and 9.15 kcal/mol, respectively. In light of the average uncertainty in the experimental measurements reported above, the B97 and B3LYP functionals arguably yield, on average, comparable accuracy to ph-AFQMC, with MAEs of 1.57 and 1.76 kcal/mol, respectively. Yet the MaxE's of 6.45 and 6.48 kcal/mol are more than twice as large as that from ph-AFQMC, and would be considered much too large for many predictive applications. ωB97X-V achieved a similar accuracy as DLPNO-CCSD(T), with MAE and MaxE of 2.74 and 8.48 kcal/mol, respectively. This performance is rather satisfactory given that there were no transition metals in the training set used to fit the 10 empirical parameters in the functional.[23] In contrast, the Minnesota functional, M06, is heavily parameterized and results in the largest MAE of 3.87 kcal/mol. The poor performance of M06 for transition-metal complexes was also mentioned in our group's previous paper[83] and in the work of Grimme and co-workers[39]. In contrast to the high accuracy achieved by double-hybrid functionals for organic molecules[81,82], the DSD-PBEP86 functional for this dataset yielded an MAE of 3.73 kcal/mol and MaxE of 13.69 kcal/mol.

According to Grimme and co-workers, DFs with a smaller amount of HF exchange tend to perform better than those with larger percentages.[39] We see a similar trend that B97 (19.43% HF exchange) gives the best performance for this dataset while M06 (27% HF exchange) and DSD-PBEP86 (~ 70% HF exchange) perform the worst. PBE0 with an MAE of 2.17 kcal/mol is slightly worse than B3LYP and B97; however, it yields good results for dihydrogen complexes.

We attempted to correlate a number of multireference diagnostics, such as the fractional occupation number weighted electron density (FOD)[84,85] and the square of the leading CI coefficient in the CASSCF calculation,[53] with errors from DLPNO-CCSD(T). However, no significant correlation was found. This is consistent with previous studies reporting similar inefficacy for transition metal systems.[43–46,53] We emphasize the need for further investigation and development of multireference diagnostics that can reliably identify the presence of strong correlation effects and thus signal caution to users of single-reference methods such as DFT and CCSD(T). One promising approach involves examining the deviation of $\langle S^2_{UHF} \rangle$ from spin-pure values, in conjunction with the use of an orbital-optimized method, e.g. MPn, to rule out artificial symmetry breaking[86].

For reactions involving Sc, Ti, V, and Cr centers, our ph-AFQMC results are typically in good agreement with experiment even when relatively small active spaces are employed in the trial wavefunction. Such calculations need only use the MINAO basis set to specify the 3*d* orbitals as inputs for the AVAS procedure for selecting the active space. For the remaining metals, larger active spaces (i.e. including higher-lying virtual orbitals) are required, and we therefore used the ANO-RCC basis for AVAS, specifying both the 3*d* and 4*d* atomic orbitals to account for the double-shell effect.[57,87,88]

As a number of functionals were trained utilizing larger basis sets than the one employed in this work, we note that the results may change slightly if such optimal basis sets had been employed. We did investigate the basis set dependence for the double-hybrid functional, as the MP2-like part is known to perform better with a basis larger than TZ to more closely approach the complete basis set limit.[89,90] We found for the largest outliers for DSD-PBEP86 that using a QZ basis set for the single-point energy calculations did not significantly change the results. For example, the calculated BDEs of $[Fe(NH_3)_4]^+$ in TZ and QZ deviate from experiment by 13.69 and 13.89 kcal/mol, respectively.

1.4.7 Discussion

The results we have obtained lead to interesting observations concerning all three classes of approaches considered in this paper: AFQMC, DLPNO-CCSD(T), and DFT. These observations have implications that go beyond the current data set. Our previous AFQMC study on transition metal containing dimers[47] could be viewed as addressing a very special subset of unusual and difficult molecules from an electronic structure point of view. In particular, these systems are coordinatively unsaturated, with nearly degenerate electronic states in a number of cases, and of a form rarely present in important chemical systems relevant to practical applications in biology and materials science. In contrast, the present data set contains many typical bonding motifs, namely four and six coordinated metal-ligand complexes, although the oxidation states are lower than is usually found in condensed phase systems. Arguably, a system such as the water splitting complex in Photosystem II poses a much more difficult quantum chemistry problem than the molecules considered here. A method that displays a significant number of outliers in our present data set would be difficult to trust as reliable if applied to a strongly interacting, multi-metal complex with a large number of low lying electronic states.

The AFQMC results satisfy all of the criteria one could reasonably expect (given the uncertainties in the experimental data) for true benchmark performance. The largest deviation from experiment is less than 3 kcal/mol, often cited as the target for ‘transition metal chemical accuracy’,[74] and close to being within the cited experimental error bars. For most of the ligands studied, the maximum deviation is closer to 2 kcal/mol and well within experimental error. Results reliably improve (sometimes considerably) as the quality of the calculation is increased, e.g. via an upgrade in the basis set extrapolation method. In fact, the error for the $[\mathrm{Ni(H_2O)_4}]^+$ molecule, which represents the MaxE of ph-AFQMC in Table 1.6, can be reduced to less than 1 kcal/mol when utilizing QMC/UHF rather than DLPNO-CCSD(T) for the basis set extrapolation (we indicate in Tables S4 and S5 in the supporting information of Ref. 1 that extrapolating with QMC/UHF will produce equally good if not better final BDEs for a representative selection of molecules, suggesting that such extrapolation is to be preferred, if computationally feasible, in future studies). With this update the MaxE of ph-AFQMC would be lowered to 2.39 ± 1.46 kcal/mol, for $[\mathrm{Ti(CO)_4}]^+$, which is a rather outstanding result in light of the experimental uncertainty. The overall mean unsigned deviation from experiment of 1.1 kcal/mol is highly satisfactory. It is in fact not obvious how much of this deviation is due to errors in the theory and how much to errors in the experiment. In our transition metal dimer publication, it is noteworthy that when new (and more reliable) experiments were released after the calculations were completed (but prior to publication), agreement of AFQMC with these results was significantly better than with older values. In the absence of significantly more accurate experiments, it is hard to imagine a better performance from a tractable theoretical approach.

The DLPNO-CCSD(T) results, in contrast, reveal a large number of major outliers (with a maximum outlier of 9.15 kcal/mol) across every single ligand series (maximum deviations for the individual series range from 4.29 kcal/mol to 9.15 kcal/mol). The DLPNO approximations are likely *not* the most significant sources of error, given that we use the tightest possible cutoff parameters, and in light of the results in Ref. 91. In addition, due to the relatively small size of the dissociating ligand, it is reasonable to expect some degree of cancellation in the localization errors. It is most likely that excitations of higher order than (T) are required for consistently high accuracy, though we note that it would be a useful future investigation to probe the effects of utilizing orbitals from, e.g., an unrestricted DF calculation. Regardless of the source of the errors, the implication is that much more expensive (and poorly scaling) variants of coupled cluster will be needed to converge this approach to chemical accuracy for transition metal containing systems. Now that benchmark values are available (via our AFQMC results) for both transition metal containing dimers and small four and six coordinated complexes (comprising roughly 80 systems in all), we look forward to alternative CC approximations being rigorously evaluated using this data. At that point, assuming that comparable benchmark quality can be achieved, it will be interesting to compare the computational requirements, and scaling with system size, of both methods.

The DFT results shown here are far from a comprehensive survey of the various flavors of functionals currently available, but do contain a number of qualitatively different functionals as well as several of the most widely used approaches. A striking observation is that the three best performing functionals- by a considerable margin- were published more than 20 years ago. Despite the use of considerably more sophisticated functional forms, the performance of the three more recent functionals (wB97X-V, DSD-PBDP86, and M06) have substantially worse average errors, and larger and more frequent outliers, than the older approaches. It should also be noted that the best performing DFT approaches work substantially better than DLPNO-CCSD(T). This observation is in accordance with the proposition put forth along these lines by Truhlar and coworkers several years ago, which has been the subject of considerable controversy in the literature.[43,46,47] While one could ultimately converge coupled cluster–based methods to a benchmark level of accuracy by including higher (and considerably more expensive) levels of theory, what is going to be necessary and sufficient to accomplish that convergence is apparently more demanding than some of the earlier papers in this debate have suggested.

Our results cast doubt as to whether the newer DFT models use a functional form that is an actual improvement from the point of view of transition metal chemistry, as the incorporation of asymptotically correct exchange, non-local correlation, MP2 contributions, kinetic energy density-dependence and/or a greater number of parameters appears not to yield improved accuracy over simpler hybrid GGA forms. As in the case of typical machine learning problems, consideration of additional parameters generally leads to better performance when the test cases are similar to the molecules in the training set, i.e. when direct interpolation is performed. Extrapolation outside of the training set, however, is a very different proposition. The lack of confidence in the experimental values for transition metal energetics has deterred extensive incorporation of data of the type we have studied here into the process of fitting DFT functionals. Our benchmark level of agreement with experiment should enable new efforts, incorporating the data we have validated here, to proceed with more confidence. And it is of course possible that one of the many DFT functionals that we have not tested in this paper would improve upon any of the results presented above. Again, data is now available to rigorously interrogate such a proposition.

The performance of the two best performing methods, B3LYP and B97, is quite remarkable considering their vintage and relatively small number of fitting parameters (3 and 10, respectively). It is interesting that whereas B97 was clearly superior for the transition metal dimer data set, the results for the present data set are much closer in average and maximum error. For calculations of large, transition metal–containing systems, we would view either of these alternatives as the best currently available, particularly given the extensive experience with them over the past several decades (although not of benchmark quality, in view of the presence of a significant number of outliers in the 3-7 kcal/mol error range). If the AFQMC calculations can be scaled up to address systems with 50-100 atoms, perhaps by using localized orbital techniques, a combination of AFQMC benchmarks followed by B97 or B3LYP modeling of a larger set of conformations (including environmental effects such as solvation), could provide a path towards calculations of high enough quality to understand reaction mechanisms, identify intermediates, and contribute to molecular design efforts.

1.5 Conclusions

Our ph-AFQMC approach has produced reliable theoretical values for BDEs in $3d$ transition metal coordination complexes. Our results demonstrate that future, predictive benchmarking should employ CAS trial wavefunctions in the TZ basis with QMC/UHF for CBS extrapolation. The MAEs of the DFs considered in this study are in general quite satisfactory, but the occasional presence of large, unsystematic errors leaves cause for concern. The performance of methods by MAE from best to worst is ph-AFQMC, B97, B3LYP, PBE0 DLPNO-CCSD(T), ωB97X-V, DSD-PBEP86, and M06, respectively.

We envision that this dataset of gas-phase BDEs may prove useful for the development of new approximate methods, and new DFs. The reliability of the ph-AFQMC method, namely its ability to compute accurate gas-phase energetics in a reasonable amount of wall-time, will enable the development of accurate force-fields for metal ion interactions with various ligands. The method will also help in a forthcoming investigation of DFT's ability to predict solution-phase properties. For instance, we are now in a position to answer the question: are errors found in recent studies of aqueous pK$_a$'s[19] and redox potentials[83] due inherently to deficiencies in the quantum-chemical electronic structure description or in the implicit solvent models employed, or both?

For the systems in this work, we were generally able to converge the BDEs with respect to active space size of the trial wavefunctions. However, moving on to larger systems, perhaps containing multiple metals or bulky ligands, we anticipate that the relevant active space sizes will overcome conventional CASSCF algorithms and available computing resources. Investigations along these lines are currently underway, as are efforts to implement a localized orbital approach to ph-AFQMC.

Chapter 2: Benchmarking Phaseless Auxiliary Field Quantum Monte Carlo (ph-AFQMC) on Transition Metal Complexes

Reprinted with permission from B. Rudshteyn *et al.*, *J. Chem. Theory Comput.*, vol. 18, no. 5, 2022. Copyright 2021 American Chemical Society.

2.1 Introduction

Quantum chemical methodology has made tremendous progress in both accuracy and computational efficiency during the past three decades[92]. 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).[93] 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.[94]

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,[95] B3LYP,[61–63] and related models.[78,96–99]

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[100,101] and Neese[50,51,91,102–104]) 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,[105] and corrected (if necessary) by CCSD(T) cluster calculations on the reactive core of the system.[106]

Assessing the accuracy of both DFT and CCSD(T) for transition metals has been more difficult than for typical organic systems. CCSD(T) yields highly precise results for transition metal atoms (for example for ionization energies),[57] but the chemical accuracy of this approach as well as DFT for diatomic molecular bond dissociation energies has remained somewhat less clear.[31,43–48]. In an impressive recent investigation, Hait et al. examined the convergence of a variety of coupled cluster approaches for a series of 69 3d transition metal oxides, sulfides, carbides, and nitrides within a small basis set.[49] They found that CCSD(T) generally produces highly accurate results for systems other than carbides and nitrides, and reasonable but less accurate results for these polyvalent systems. Still, outliers exist, for example the molecule NiO, for which CCSD(T) produces a large (~ 10 kcal/mol) error. Hait et al. point out that different metrics for assessing multi-reference character can conflict, and in general the errors produced by CCSD(T) were not tightly correlated with multireference character. Taken together, the results of Hait et al. suggest that CCSD(T) should indeed be a method of choice for transition metal-containing systems. However given the cost of the approach, systems of the size of the molecules we treat in the present work cannot be treated with full CCSD(T), especially if full basis set extrapolations are to be carried out. Thus less expensive and unfortunately less accurate versions of coupled-cluster must be investigated.

Unlike the case for systematically improvable wave function methods, DFT results can 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,[26–36] 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[83,107]), or reference reactions, which can be used to take advantage of error cancellation,[108,109] 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.[110]

In a series of recent publications, we have made 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,[54,55] 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 planewaves[54]) or N^4 (with Gaussian-type orbitals[73]) [vs. N^7 for full CCSD(T)[37]], a non-perturbative and multi-reference nature, and the ability to utilize a multiconfigurational SCF trial wavefunction. Unlike other QMC methods, it does not involve real-space sampling, but rather, sampling in the space of Slater determinants. The ability to use a sophisticated multi-determinant trial is crucial for the treatment of many 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,[57,111] utilization of correlated sampling (CS) in the Monte Carlo protocol to directly compute energy differences,[56] and implementation on GPU hardware,[57,112] has made it feasible to treat significantly larger systems.[1,3,113] Ref. 111 is recommended as a good introduction to the theory of AFQMC.

These advances have allowed systematic studies of three classes of small transition metal containing species (atoms, diatomic molecules, and 4-6 coordinate complexes containing simple small molecule ligands) with highly encouraging results. Atomic ionization potentials,[57] diatomic bond energies,[47] and complex ligand dissociation energies,[1] have all been computed with a MAE of less than 1.5 kcal/mol across relatively large experimental gas phase data sets. These results were found to outperform the best DFT functionals and feasible variants of coupled cluster theory. 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.[1] Improved agreement with precise, state-of-the-art experimental data,[114] some of which were measured after the calculations were carried out, further validated the robustness of the AFQMC approach.

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,[46,48,49] 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. 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_0) 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.[115] Finally, adiabatic gas phase experimental data, measured with electron transfer equilibrium (ETE) by Richardson and co-workers[116–118] with relatively low experimental uncertainties, as well as vertical gas phase experimental data, as measured with photoelectron spectroscopy by Green and co-workers,[110,119] exist for this series, which can form the basis to evaluate the accuracy of the various quantum chemical approaches for prototypical organometallic species, though this is not an exhaustive set. Metallocenes on their own are an important class of organometallic compounds, given their importance in alkene polymerization and electrochemistry.[120] Much of the previous literature of correlated calculations has focused on either predicting the spin splitting of metallocenes in solution,[121–123] which presents complications due to the solvent environment, and/or the bond dissociation energy of the M-Cp bond[124] rather than the ionization energy, which has a direct equivalent in solution and has both adiabatic and vertical variants. When the ionization energy is studied, it is usually just for ferrocene[121,125].

A number of significant conclusions emerge from the metallocene calculations presented here. The localized coupled cluster approach that we have employed — DLPNO-CCSD(T_0) with the particular thresholds described in the methods section — displays a number of large outliers and an overall MAE that is comparable to those typically obtained from DFT functionals. To further probe the source of the errors, we have examined one of the more challenging metallocene systems, $Mn(Cp)_2$, at higher levels of coupled cluster theory, tightening the cutoffs and replacing (T_0) with (T_1) in the DLPNO approach, and carrying out full CCSD(T) calculations in a small basis set. These calculations show considerable differences from our default DLPNO-CCSD(T_0) results, moving the computed ionization energies in the direction of the experimental value. Rigorously converging CCSD(T) (or even higher levels of excitation) to the CBS limit would be computationally very expensive, and hence is beyond the scope of the present paper. However, it is clear that further effort to test and develop scalable coupled cluster based methods for treating transition metal containing systems should be a high priority of the community, and is likely to yield fruitful results.

The gradient corrected, hybrid, and range-separated hybrid DFT functionals display MAEs between 3.5 and 5.5 kcal/mol for both the vertical and adiabatic ionization energies, with one or more individual errors greater than 7 kcal/mol; no functional performs at the lower end of this range for both data sets. The double hybrid functional DSD-PBEP86 displays significantly worse average errors and outliers than those seen in other functionals, in agreement with our prior results on other transition metal containing test sets, and consistent with discussions in the literature with regard to difficulties experienced by the current generation of these functionals for many transition metal containing systems.[126,127] Attempts to improve double hybrid performance for metal containing systems are at present ongoing in a number of research groups.

Motivated by these results, we also tested these DFT functionals and DLPNO-CCSD(T_0) on a second set of gas phase ionization energies measured by the same experimental group,[128,129] the tri-acetylacetonate (acac) systems (V through Co). These acac complexes are also an important set of coordination compounds with organic scaffolds, widely studied as models for other tris-β-diketonate complexes and as sources of transition metals in chemical vapor deposition processes.[129] Here, we find two molecules, namely $[\mathrm{Cr(acac)_3}]^{1-}$ and $[\mathrm{Mn(acac)_3}]^{1-}$, for which very large outliers are obtained, confirming the initial picture that one can often obtain quite accurate results, but that major failures can occur as well.

For the vertical and adiabatic metallocene ionization energies, we were able to develop a systematic AFQMC protocol that achieved accuracy within the experimental noise limitations. This was achieved by overcoming significant challenges arising from the greater size and complexity of the metallocene series (as compared to molecules in earlier publications).[1,3,47,57,72] For example, prior work[1] has demonstrated that it is essential to employ an appropriate multiconfigurational trial wavefunction, since calculations based on Hartree-Fock (HF) trial wavefunctions did not reliably lead to chemically accurate results for these systems. Deployment of an appropriate multiconfigurational trial wavefunction in the application of the AFQMC approach to transition metal containing systems remains essential if one is aiming at robust, benchmark quality results. Another key component of our protocol is to use correlated sampling (CS) to compute energy differences whenever the two (or more) systems can be effectively correlated. In this work all vertical ionization energies (including the vertical step of an adiabatic ionization) are computed by CS. We then validate geometry reorganization energies predicted by lower-level theories with separate AFQMC calculations using standard branching and population control (PC).

In the latter form of AFQMC calculations, MC sampling is carried out independently for two different states and, at intervals, walkers with large weights are duplicated while those with small weights are destroyed with appropriate probability via a "comb" algorithm[73]. This is needed to mitigate the weight fluctuations in the branching random walk to maintain Monte Carlo sampling efficiency. A CS simulation typically can be carried out for a much shorter duration, during which the need for PC of AFQMC walkers is minimized and the accruement of phaseless constraint error is sometimes reduced. 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 some cases.[1,47] The use of CS AFQMC to measure energy differences for vertical transitions, and PC AFQMC, rather than DFT, to measure differences between two geometries on the same surface appears to be a very promising approach for all cases in this work. 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_0) and DFT methods employed. Section IV presents results of AFQMC, various DFT functionals, and DLPNO-CCSD(T_0) calculations for the metallocene series, as well as DFT and DLPNO-CCSD(T_0) 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_0), and AFQMC in addressing transition metal chemistry. Finally, in Section VI we conclude with a summary of our results and outline future directions.

2.2 II. Experimental Data Sets

The gas-phase ionization energy experiments that we investigate below center on first row transition metals in the II or III oxidation state, with either anionic cyclopentadienyl (Cp) ligands (metallocenes) or acetylacetonate ligands (acac series), depicted schematically in Figs. 2.1a and 2.1b, respectively.

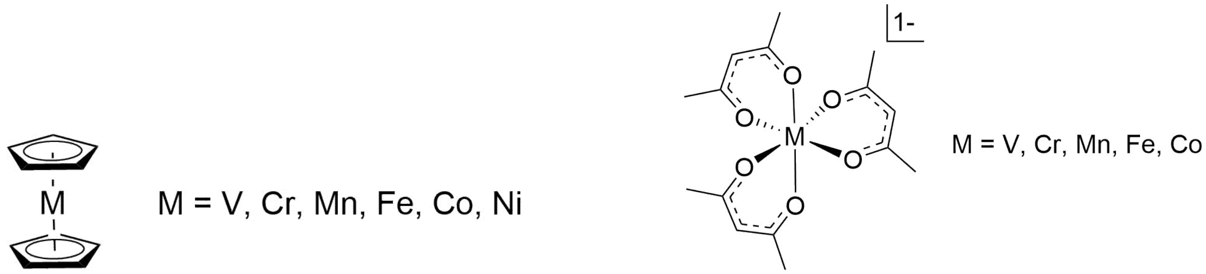
The metallocenes we investigate are vanadocene, chromacene, manganocene, ferrocene, cobaltocene, and nickelocene. The ionization processes we study are for the II oxidation state (charge = 0) to the III oxidation state (charge = +1). All of the metallocene molecules in these two oxidation states are 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). Note that in the original work by Richardson and co-workers, what is actually reported is the “attachment" energy,[128,129] meaning the reduction energy from oxidation states III to II, but we reverse the sign here to facilitate comparison to the ionization of metallocenes. 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/term symbols of all species are given in Table 2.1.[107,117–119,124,130–135] The expected ground state term symbols from the acac complexes come from standard Tanabe–Sugano diagrams; we do not explicitly constrain the geometric symmetry or term symbols of our calculations, merely the multiplicity and charge.

The adiabatic experimental values come from the electron transfer equilibrium (ETE) measurements of Richardson and co-workers[116–118]. These experiments utilize Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) to determine adiabatic ionization energetics

Table 2.1: The expected ground state term symbols 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	$^{3}\mathrm{A}_{2}$	$^{4}\mathrm{A}_{2}$	$^{3}\mathrm{T}_{1}$	$^{4}\mathrm{A}_{2}$
Cr	$^{4}\mathrm{A}_{2}$	$^{3}\mathrm{E}_{2}$	$^{4}\mathrm{A}_{2}$	$^{5}\mathrm{E}$
Mn	$^{3}\mathrm{E}_{2}$ ($^{5}\mathrm{E}_{1}$)	$^{6}\mathrm{A}_{1}$	$^{5}\mathrm{E}$	$^{6}\mathrm{A}_{1}$
Fe	$^{2}\mathrm{E}_{2}$	$^{1}\mathrm{A}_{1}$	$^{6}\mathrm{A}_{1}$	$^{5}\mathrm{T}_{2}$
Co	$^{1}\mathrm{A}_{1}$	$^{2}\mathrm{E}_{1}$	$^{1}\mathrm{A}_{1}$	$^{4}\mathrm{T}_{1}$
Ni	$^{2}\mathrm{E}_{1}$	$^{3}\mathrm{A}_{2}$	-	-

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.[136] Because these experiments were performed at 350 K, we use this temperature value in our calculated free energy corrections.



(a) Metallocene compounds (M = V to Ni).

(b) acac compounds (M = V to Co).

Figure 2.1: The structure of the metal complexes studied.

The vertical experimental values come from the photoelectron spectroscopy (PES) measurements of Green et al.,[110,119] 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. Alternate experimental values are discussed in the SI of Ref. 2. They are mostly similar to the results from Green et al. suggesting the older results are reliable. It should be acknowledged though that vertical excitations may not have well defined band shapes and true values can deviate from the maxima by a few kcal/mol. A potentially similar effect in photoaborption spectroscopy appears to be on average 2.54 ± 1.85 kcal/mol,[137] which is in line with 2.31 kcal/mol uncertainties provided by Green et al.[119]

2.3 III. Computational Details

The geometries, certain reorganization energies (*vide infra*), and ideal gas free energy corrections[138] were obtained with the B3LYP functional[61–63] utilizing the cc-pVTZ-DKH[64–67] basis set and DKH2 relativistic corrections[139–142] using the one-center approximation (as implemented in ORCA) and without symmetry constraints. Geometries were confirmed to be minima using normal-mode analysis. Structures with unpaired electrons were confirmed to display Jahn-Teller distortions. These calculations were performed using the ORCA program package.[68] In the cc-pVTZ-DKH basis set, metallocenes typically have around 508 basis functions and 95 electrons. In the same basis set, 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 of Ref. 2.

We investigated the addition of D4[143] dispersion to the geometry optimization of $MnCp_2$. Consistent with previous results,[124] we find the geometry is relatively similar to that without these dispersion corrections, as discussed in the SI of Ref. 2. We also have evaluated the B3LYP-D4 energetics for the ionization energies of both the metallocenes and acac complexes and find similar results to those found below, as discussed in the SI of Ref. 2.

Utilizing diffuse functions in the basis set for both metallocenes (aug-cc-pVTZ on all atoms) and acac complexes (aug-cc-pVTZ on the O atoms), which one might argue could be significant because of the anionic ligands, yields very similar results, as is discussed in the SI of Ref. 2. Similar calculations for the metallocenes, but with the diffuse functions on the C using B3LYP yield similar MAE's.

We were able to perform all-electron AFQMC calculations, as opposed to invoking the frozen-core approximation. We have tested the effect of our use of the relatively compact basis sets optimized for frozen-core calculations (cc-pVXZ-DKH) by comparing, via the use of an ECP and the cc-pVTZ-pp basis set[144,145]. The results in the SI of Ref. 2 show that this effect would not change the results noticeably.

Unrestricted DLPNO-CCSD(T_0) (T_0 refers to the semi-canonical approximation to the perturbative triples correction[50]) calculations were performed with quasi-restricted orbitals (QROs) generated from unrestricted B3LYP reference orbitals and “NormalPNO" (moderate energy cut-off criteria for correlation between localized orbitals) localization parameters using ORCA. These calculations utilized cc-pVxZ-DKH and auxiliary cc-pVxZ\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,[68] as discussed in the SI of Ref. 2. TZ/QZ extrapolation schemes have been used successfully for AFQMC[1,47,56,57,72,146]. Such extrapolation (typically with TZ/QZ basis sets) or at least evaluation at QZ has been shown to be important in various applications of DLPNO-CCSD(T).[41,104,147–150] 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,[151] The keyword “NoFrozenCore" was used so that no electrons would be frozen. The SI of Ref. 2 shows that not including this keyword would not significantly change the results, similar to the AFQMC results. For these calculations, the one-center approximation was not used for the relativistic corrections.

As discussed in the SI of Ref. 2, we investigated the convergence of the DLPNO-CCSD(T_0) results with respect to the PNO cut-off values, including extrapolating the TCutPNO parameter to 0[150], and the treatment of the (T) correction for the case of $MnCp_2$, which turns out to be the biggest outlier for DLPNO-CCSD(T_0) for our adiabatic results discussed below. We also compare to full CCSD(T) in the DZ basis set. The results do improve significantly (from an error of about 10 kcal/mol to about 1.5 kcal/mol from experiment) with the use of increasingly tight PNO cut-off criteria and iterative T_1 corrections. Preliminary results suggest that the CBS limit would not be much different. Such an error reduction is beyond those seen from the use of NormalPNO by others.[104,149,152] htPNO" cut-offs, and extrapolation of the TCutPNO approximation are used. Nevertheless, we limit our interpretation of the coupled cluster results to the version of DLPNO-CCSD(T_0) that we used and propose that further investigation of these systems with coupled cluster variants is warranted, given the expense of running the most rigorous implementation.

Electron repulsion integrals and the Hamiltonian for ph-AFQMC were obtained with PySCF[70]. The exact-two-component (x2c) relativistic Hamiltonian[71,153–157] was used in place of DKH2, as the latter is not implemented in PySCF. Both methods are discussed in the SI of Ref. 2.

While ph-AFQMC can be extended to excited states[158], the implementation we use is limited to studying the ground state of a given combination of charge, multiplicity, and geometry. 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 in these systems such that errors from the Trotter decomposition are negligible given our target statistical accuracy.[56] We utilize single precision (sp) rather than double precision (dp) floating point arithmetic, as discussed in the SI of Ref. 2. 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.[159] Additional details are given in the SI of Ref. 2.

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 (~ 10%) with the 3d and 4d orbitals (as defined from the Atomic Natural Orbital (ANO-RCC) basis set)[160] of the metal ion were included.[76] 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. We typically retain 98% of the CI weight (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 of Ref. 2.

The CBS limit for the ph-AFQMC calculations was estimated using an approach similar to that described in previous work[1]. Briefly, we extrapolate the ionization energy computed with ph-AFQMC PC with the cc-pVxZ-DKH basis sets (x=3,4 for T,Q) using an inexpensive trial wavefunction such as UHF or a CASSCF wavefunction with a small active space. 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_0), 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,

$$\rho = \frac{\Delta E^{corr}(\text{TZ, high-level})}{\Delta E^{corr}(\text{TZ, low-level})}. \tag{2.1}$$

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.

Our CS approach[56] enables fast convergence of vertical energy differences 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 absence of PC results in CS requiring more initial walkers, but the reduced statistical fluctuations from correlated samples allow 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. For this approach to be justified, it is necessary that the simulations 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. Cross checks with the corresponding independent AFQMC runs can help to validate convergence.

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.[1,47,57] For $MnCp_2$, 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 of Ref. 2). Results across three different types of data sets[1,47,56] suggest that the CS methodology achieves a greater reliability than PC approaches in comparing different electronic surfaces to within near chemical accuracy.

Since CS calculations are most effective if the geometry of the two surfaces is held constant, the adiabatic ionization free energy (IE), is computed by

$$IE = E(III) - E(II) - \lambda + \Delta(III) - \Delta(II), \tag{2.2}$$

where E is the electronic energy obtained from AFQMC. Δ is the thermal correction that account for temperature effects obtained from ORCA.

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, as in our previous work,[1] or the PC ph-AFQMC with either a UHF or CAS trial wavefunction, as described in the Results section. The approximation of using B3LYP for the reorganization 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 calculations for the two geometries remedies this problem, and yields highly satisfactory agreement with experiment for all cases. The calculation of IEs is illustrated in Fig. 2.2. λ combined with the energy difference between the minima $E(III) - E(II)$ gives the ionization energy of the reaction at 0 K ignoring zero point energy. The thermal correction difference for the reaction, $\Delta(III) - \Delta(II)$, is then added to bring the energy to the value at 350 K.

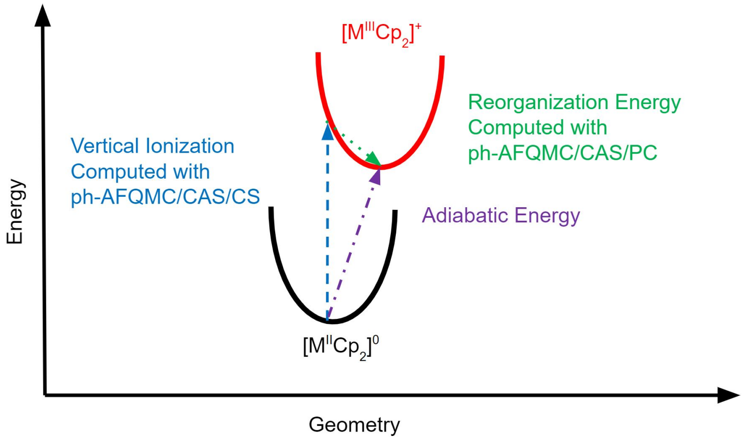


Figure 2.2: Schematic of the ionization and reorganization electronic energy calculations performed in this work without consideration of thermal corrections . CS and PC indicate the energy was measured by the CS and PC approaches, respectively.

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 the Summit supercomputer at the Oak Ridge Leadership Computing Facility (< 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[161,162], hybrids B3LYP[61], B3LYP*,[163] PBE0,[78] and B97,[96,97] the meta-GGA hybrids M06[77] and TPSSh,[164,165] the semilocal meta-GGA (non-hybrid) B97M-V,[166] the range-separated hybrid meta-GGA ωB97M-V,[24] the range-separated hybrid ωB97X-V functionals[23], and the double-hybrid DSD-PBEB86.[79,80] These calculations were done in ORCA. We perform single-point energy calculations, without the one-center approximation for the relativistic corrections. The MP2 part of the double-hybrid calculations used the frozen core approximation (10 electrons ($1s^22s^22p^6$) for 3d transition metals and 2 electrons for C and O ($1s^2$)). B3LYP geometries were obtained using the large “Grid7" option in ORCA (see SI of Ref. 2).

2.4 IV. Results

2.4.1 Vertical Ionization Energies

The vertical IE results using ph-AFQMC methods are given in Table 2.2. AFQMC PC/UHF has numerous large outlier cases, such as $CrCp_2$ and $NiCp_2$. 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 ^{5}Mn(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.[126] Indeed, the calculated M-Cp lengths (Table S29 of Ref. 2) and experimental gas-phase homolytic M-Cp dissociation energies (Table S30 of Ref. 2)[117,118] 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 H_2 or N_2 are stretched, an analogous situation has previously been reported for a weakly bound tetramine Mn cation in the gas phase.[126] 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 $Mn^{II}Cp_2$, $^5Mn^{III}Cp_2$, 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 of Ref. 2).

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 and DLPNO-CCSD(T_0) results for vertical ionization energies, using the various functionals enumerated in section III are compared to ph-AFQMC in Fig. 2.3 and are enumerated in Table 2.3.

The hybrid functionals B3LYP B3LYP*, M06, B97, PBE0, TPSSh, ωB97X-V and ωB97M-

Table 2.2: Vertical ionization energies as a function of metallocene type and AFQMC 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. The way the uncertainties for MAE's and MSE's are calculated is described in the SI of Ref. 2.

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

V, as well as the meta-GGA B97M-V all have MAE's between 4.5 and 5.7 kcal/mol indicating 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. 126. 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 2.6 below. Among different functionals, the maximum error often occurs at different metallocenes, indicating a lack of predictable reliability.

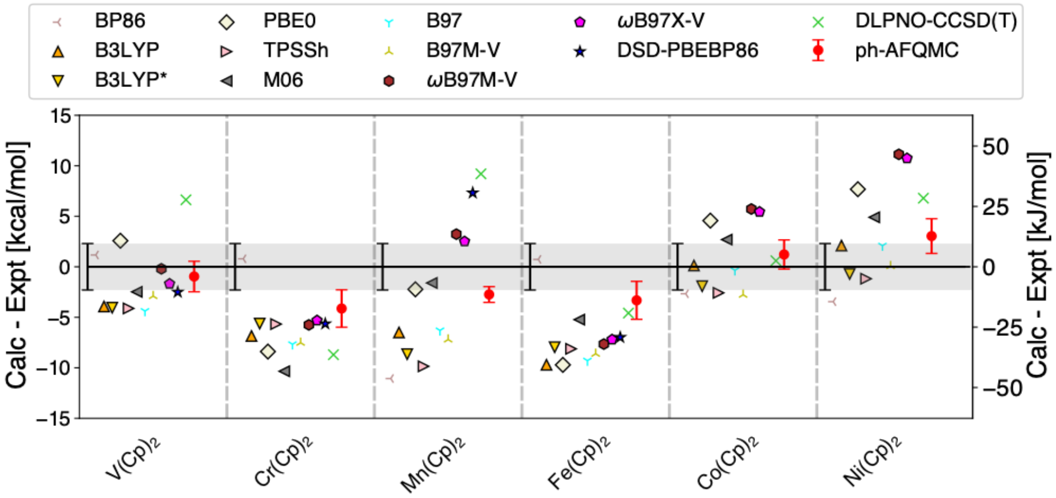


Figure 2.3: Performance of ph-AFQMC (CAS trial, CS), DFT functionals (QZ basis), DLPNO-CCSD(T_0) (extrapolated with TZ/QZ basis sets), and experiment for prediction of experimental vertical gas-phase ionization energies for metallocenes. The range is limited to deviations of -15 to 15 kcal/mol. The DSD-PBEP86 result for Co(Cp)$_2$ and Ni(Cp)$_2$ are out of range with deviations of about -20 and 17 kcal/mol respectively. The gray band indicates the uncertainty of the experiments i.e. 2.31 kcal/mol. The errors are given in both units of kcal/mol (lefthand axis) and kJ/mol (righthand axis).

Table 2.3: DFT (QZ basis) and DLPNO-CCSD(T_0) (extrapolated with TZ/QZ basis sets) 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	ωB97X-V	ωB97M-V	DSD-PBEP86	DLPNO-CCSD(T)
$V(Cp)_2$	156.3 ± 2.3	157.47	152.37	152.25	158.88	152.18	153.82	151.92	153.40	154.62	156.10	153.80	162.93
$Cr(Cp)_2$	131.4 ± 2.3	132.19	124.55	125.78	123.01	125.73	121.06	123.73	123.88	126.09	125.65	125.76	122.70
$Mn(Cp)_2$	159.3 ± 2.3	148.24	152.80	150.65	157.07	149.45	157.70	153.00	152.10	161.81	162.53	166.62	168.50
$Fe(Cp)_2$	158.7 ± 2.3	159.43	149.00	150.74	148.99	150.58	153.47	149.41	150.09	151.50	151.05	151.72	154.12
$Co(Cp)_2$	128 ± 2.3	125.33	128.13	126.08	132.57	125.42	130.69	127.65	125.23	133.44	133.72	108.35	128.60
$Ni(Cp)_2$	150.1 ± 2.3	146.65	152.20	149.39	157.78	148.92	155.00	152.19	150.17	160.84	161.23	167.29	156.89
MAE		3.31	4.87	4.82	5.86	5.25	4.54	5.01	4.84	5.48	5.61	9.88	6.08
$\|MaxE\|$		11.06	9.70	8.65	9.71	9.85	10.34	9.29	8.61	10.74	11.13	19.65	9.20
RMSD		4.90	5.82	5.63	6.53	6.05	5.39	5.89	5.75	6.25	6.57	11.70	6.73
MSE		-2.42	-4.12	-4.82	-0.92	-5.25	-2.01	-4.32	-4.82	0.75	1.08	-1.71	1.66

2.4.2 Reorganization Energies

Table 2.4 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[118]. In particular, the B3LYP M-Cp ring centroid distance decreased from 2.08 Å to 1.78 Å. To explore this further, we systematically 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 2.4 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. While the AFQMC curve is flatter and gives the appearance of a double minimum which is probably due to statistical noise, the minimum still overlaps significantly with that of the B3LYP curve. We use different AFQMC reorganization energies to calculate adiabatic ionization energies and compare the results to experiment in the next section.

2.4.3 Adiabatic Ionization Energies

The adiabatic ionization energies using different AFQMC methods are compared to experiment in Table 2.5. With AFQMC CS, the adiabatic ionization energy is computed with the two-step procedure illustrated in Fig. 2.2, while with AFQMC PC, it is computed directly as a two-point energy difference (purple line). 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

Table 2.4: 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
$V(Cp)_2$	0.37	1.73 ± 0.58	1.76 ± 0.52
$Cr(Cp)_2$	3.04	0.14 ± 0.5	-0.19 ± 0.52
$Mn(Cp)_2$	28.46	52.38 ± 0.7	41.5 ± 0.68
$Fe(Cp)_2$	3.17	0.82 ± 0.71	0.87 ± 0.69
$Co(Cp)_2$	6.57	12.45 ± 0.79	10.75 ± 0.66
$Ni(Cp)_2$	6.30	11.34 ± 0.72	11.04 ± 0.57

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.

DFT results for the metallocenes using a variety of functionals are given in Table 2.6 using the QZ basis set, as is recommended in Ref. 167. DLPNO-CCSD(T_0) results are also given. The MAEs, and average errors, of all of these methods (other than the double hybrid functional DSD-PB86, which again displays the worst performance) are in a range similar to that observed for the vertical ionization energies, between 3.5 and 5.5 kcal/mol, with a maximum error greater than 7 kcal/mol. None of these results are overall of benchmark quality, although the best performing functionals do obtain good results for individual cases. This data can be contrasted with the sig-

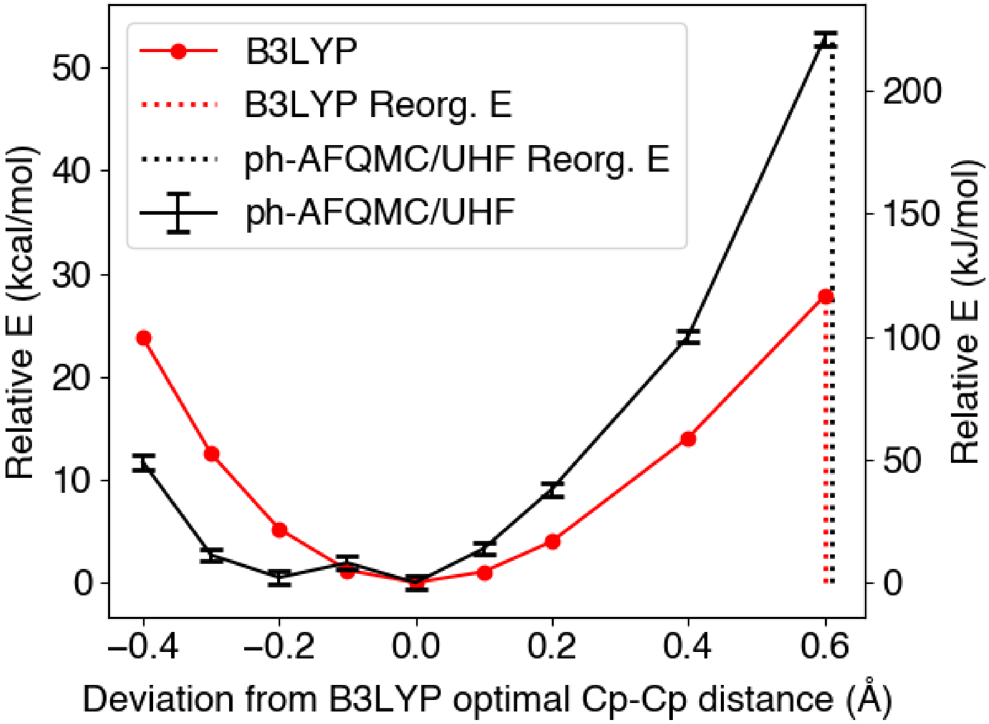


Figure 2.4: Rigid scans of the Cp-Cp distance in $[Mn^{III}Cp_2]^{1+}$ where red = B3LYP and black = ph-AFQMC/UHF. The x-axis is the deviation from the B3LYP optimal Cp to Cp ring distance of 3.55 Å. The dashed vertical lines indicate what the reorganization energy would be at the optimal ring distance of $[Mn^{II}Cp_2]^0$. The relative energies (relative E) are given in both units of kcal/mol (lefthand axis) and kJ/mol (righthand axis).

nificantly lower MAE and maximum error obtained from the best AFQMC protocol. In fact, the performance is consistent with the prior results that we have reported for transition metal diatomics (bond dissociation energies)[47] and small coordination complexes (ligand removal energies).[1] The best methods are in the ballpark for many cases, but predictions are lacking in robustness.

An important question to ask at this point with regard to the DLPNO-CCSD(T_0) results is to what extent the errors are due to the specific approximations (both localized orbital cutoffs and the use of the less rigorous (T_0) representation of triple excitations), as opposed to an intrinsic limitation of CCSD(T) itself. This is a challenging issue to explore, because the use of more computationally expensive approximations becomes problematic for systems as large as the metallocenes (let alone grand challenge problems in biology and materials science, such as the water

Table 2.5: ph-AFQMC 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	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	157.64 ± 1.55	160.25 ± 2.15	155.44 ± 1.61	153.7 ± 1.52
$Cr(Cp)_2$	127.5 ± 1.5	127.04 ± 1.43	133.54 ± 1.96	127.63 ± 1.81	131.18 ± 2
$Mn(Cp)_2$	142.5 ± 1.5	156.26 ± 1.54	151.69 ± 2	158.53 ± 1.97	146.09 ± 2.37
$Fe(Cp)_2$	153.1 ± 1.5	158.41 ± 1.84	158.37 ± 2.32	150.34 ± 1.75	152.78 ± 2.02
$Co(Cp)_2$	123.5 ± 1.5	119.06 ± 1.6	122.18 ± 1.85	128.1 ± 1.52	124.35 ± 1.83
$Ni(Cp)_2$	143.8 ± 1.5	144.75 ± 1.42	143.49 ± 1.64	148.25 ± 1.7	144.25 ± 2.15
MAE		4.68 ± 0.89	4.65 ± 1.02	4.82 ± 0.93	1.61 ± 1.02
$\|MaxE\|$		13.76 ± 2.15	9.19 ± 2.5	16.03 ± 2.47	3.68 ± 2.5
RMSD		6.43	5.53	7.15	2.16
MSE		3.04 ± 0.89	4.1 ± 1.02	3.9 ± 0.93	1.24 ± 1.02

splitting cluster in Photosystem II). Nevertheless, we have made an initial effort to address this issue, varying the localization cutoffs and triples implementation (replacing T_0 with T_1) as detailed in the SI of Ref. 2, for the $Mn(Cp)_2$ system. The use of tight cutoffs and T_1 brings the results to within 3-4 kcal/mol of experiment; using a DZ basis set to evaluate full CCSD(T) (all that we could afford) and extrapolating with DLPNO results, one would appear to come quite close to experiment (although one would have to be concerned about the accuracy of this protocol, given the large differences between the various coupled cluster approaches).

2.4.4 Acetylacetonate (acac) Results

To provide a further assessment of preliminary DFT and DLPNO-CCSD(T_0) (using our default cutoff settings), we decided to carry out calculations for the acac series of coordination 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. Results for our entire suite of DFT functionals, along with those for DLPNO-CCSD(T_0), are pre-

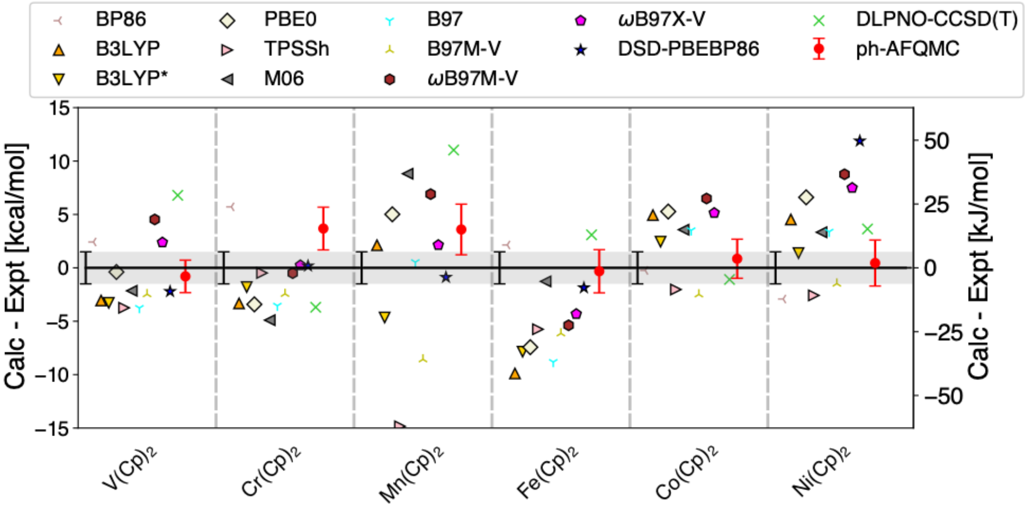


Figure 2.5: Performance of ph-AFQMC, DFT functionals (QZ basis), and DLPNO-CCSD(T_0) (extrapolated with TZ/QZ basis sets) for prediction of experimental adiabatic gas-phase ionization energies for metallocenes. The range is limited to deviations of -15 to 15 kcal/mol. The BP86 value for Mn($Cp)_2$ and the DSD-PBEP86 value for Co($Cp)_2$ are beyond of the range of the plot with deviations of about -20 and -25 kcal/mol, respectively. The gray band indicates the uncertainty of the experiments i.e. 1.5 kcal/mol. The errors are given in both units of kcal/mol (lefthand axis) and kJ/mol (righthand axis).

sented in Table 2.7 and Fig. 2.6 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.

A number of striking features of the data are immediately apparent. Firstly, none of the DFT methods perform as well as the best performers do for the metallocene series; the MAE and RMSD are in all cases well above (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.[83,107,168]

The B97, B3LYP, and ωB97M-V functionals yield the best results, at 4.56, 4.82, 4.94 kcal/mol

MAE respectively, but it must be noted that ωB97M-V was one of the worst performing functionals for the metallocenes, with an MAE of 5.43 kcal/mol. Interestingly the performance of the double hybrid functional DSD-PBEP86 is relatively similar between metallocenes and acac complexes. Across both data sets, all of the DFT functionals exhibit several failures with quite large errors. Secondly, the DLPNO-CCSD(T_0) 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 $[Cr(acac)_3]^{1-}$ species, for which most of the DFT functionals yield an adiabatic ionization potential that is ~10 kcal/mol smaller than experiment. Assuming that the acac experiments 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 the version of DLPNO-CCSD(T_0) we use 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 2.6: DFT (QZ basis) and DLPNO-CCSD(T_0) (extrapolated with TZ/QZ basis sets) 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	ωB97X-V	ωB97M-V	DSD-PBEP86	DLPNO-CCSD(T_0)
$V(Cp)_2$	154.5 ± 1.5	156.92	151.44	151.22	154.11	150.76	152.35	150.74	152.03	156.88	159.03	152.29	161.29
$Cr(Cp)_2$	127.5 ± 1.5	133.21	124.17	125.70	124.08	127.02	122.60	123.95	125.06	127.74	127.00	127.68	123.81
$Mn(Cp)_2$	142.5 ± 1.5	122.75	144.62	137.85	147.51	127.66	151.31	143.03	133.93	144.64	149.40	141.62	153.53
$Fe(Cp)_2$	153.1 ± 1.5	155.24	143.21	145.25	145.67	147.35	151.83	144.28	146.96	148.78	147.72	151.24	156.19
$Co(Cp)_2$	123.5 ± 1.5	123.26	128.44	125.95	128.77	121.47	127.07	127.01	121.01	128.64	129.99	98.62	122.41
$Ni(Cp)_2$	143.8 ± 1.5	140.88	148.33	145.18	150.40	141.22	147.13	147.18	142.32	151.29	152.56	155.68	147.44
MAE		5.53	4.64	3.57	4.69	4.90	4.00	3.93	3.93	3.62	5.43	6.98	4.89
$\lvert MaxE \rvert$		19.75	9.89	7.85	7.43	14.84	8.81	8.82	8.57	7.49	8.76	24.88	11.03
RMSD		8.58	5.29	4.19	5.22	6.81	4.68	4.63	4.68	4.31	6.00	11.33	5.85
MSE		-2.11	-0.78	-2.29	0.94	-4.90	1.23	-1.45	-3.93	2.18	3.47	-2.96	3.30

Table 2.7: DFT (QZ basis) and DLPNO-CCSD(T_0) (extrapolated with TZ/QZ basis sets) 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	ωB97X-V	ωB97M-V	DSD-PBEP86	DLPNO-CCSD(T_0)
$[V(acac)_3]^{1-}$	24.9 ± 2.4	29.11	24.30	24.28	24.21	23.11	28.42	24.67	26.37	20.66	22.60	24.60	22.50
$[Cr(acac)_3]^{1-}$	20 ± 3	7.46	10.60	8.35	10.21	2.91	15.53	11.88	1.41	8.98	10.61	14.57	7.79
$[Mn(acac)_3]^{1-}$	59 ± 5	40.15	50.31	46.30	52.09	41.48	54.84	51.75	43.78	51.87	54.55	62.67	57.13
$[Fe(acac)_3]^{1-}$	43 ± 2.4	43.11	42.86	41.74	40.85	39.75	36.30	42.11	36.62	44.25	44.50	47.04	34.05
$[Co(acac)_3]^{1-}$	47 ± 4	30.44	52.28	44.95	56.80	39.44	54.05	53.33	33.74	53.59	54.05	66.28	56.45
MAE		10.46	4.82	5.66	5.87	9.44	5.18	4.56	10.99	6.05	4.94	6.54	6.97
$\lvert MaxE \rvert$		18.85	9.40	12.70	9.80	17.52	7.05	8.12	18.59	11.02	9.39	19.28	12.21
RMSD		12.69	6.20	7.79	7.00	11.58	5.37	5.64	12.62	6.86	5.75	9.28	8.10
MSE		-8.73	-2.71	-5.66	-1.95	-9.44	-0.95	-2.03	-10.40	-2.91	-1.52	4.25	-3.20

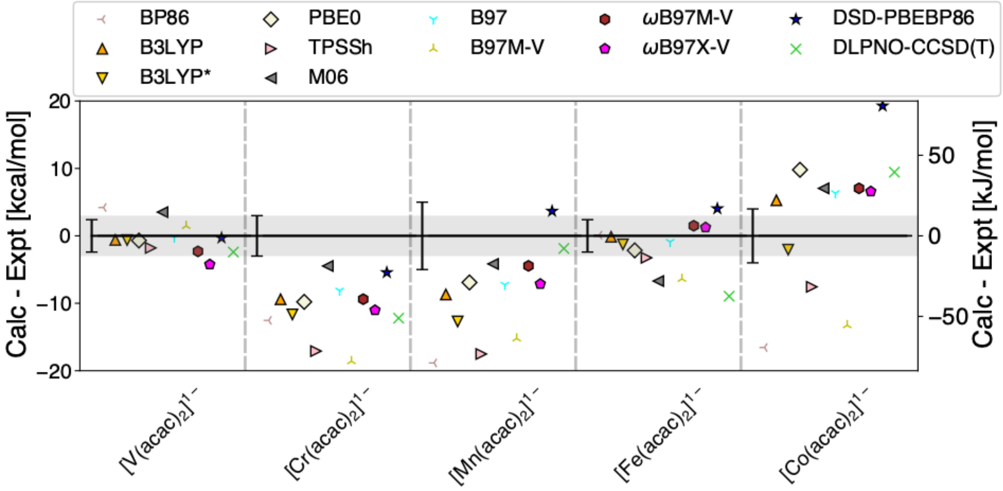


Figure 2.6: Performance of DFT functionals (QZ basis) and DLPNO-CCSD(T_0) (extrapolated with TZ/QZ basis sets) for prediction of experimental adiabatic gas-phase ionization energies for acac complexes. The range is limited to deviations of -20 to 20 kcal/mol. The gray band indicates what is often given as the chemical accuracy for transition metals: 3 kcal/mol. The errors are given in both units of kcal/mol (lefthand axis) and kJ/mol (righthand axis).

2.4.5 Reduction Potentials in Solution

Richardson and co-workers[117,118] derive THF/acetonitrile differential solvation energies for the metallocenes 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. 2.8. Details on how the potentials are derived using a thermodynamic cycle is given in the SI of Ref. 2. Table 2.8 gives the resulting potentials using our best ph-AFQMC method as well as the best performing (in terms of MAE) DFT methods for the adiabatic ionization of metallocenes, namely B3LYP*. 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 poten-

Table 2.8: Computed reduction potentials (V) using the experimentally derived differential solvation energies (E_{solv}(II) - E_{solv}(III)) (kcal/mol)[116–118] as a function of metallocene and methodology for our best ph-AFQMC method as well as the best performing functional from above for the adibatic ionization energy of metallocenes B3LYP*. The mean absolute errors (MAE), maximum errors (MaxE), root-mean-square deviations (RMSD), and the mean signed errors (MSE) are included. The uncertainty on errors do not incorporate the uncertainty on the experimental differential solvation energy

	Expt. Differential Solvation Energy	Expt. Potential	AFQMC/CAS CS CAS Trial	B3LYP*
Units	kcal/mol	V	V	V
$V(Cp)_2$	60 ± 4	-0.31	-0.38 ± 0.19	-0.48
$Cr(Cp)_2$	36 ± 4	-0.43	-0.31 ± 0.19	-0.55
$Mn(Cp)_2$	38 ± 4	0.11	0.25 ± 0.2	-0.11
$Fe(Cp)_2$	35 ± 4	0.65	0.67 ± 0.19	0.34
$Co(Cp)_2$	38 ± 4	-0.60	-0.7 ± 0.19	-0.63
$Ni(Cp)_2$	38 ± 4	0.25	0.17 ± 0.2	0.21
MAE			0.09 ± 0.19	0.15
$\|MaxE\|$			0.14 ± 0.26	0.31
RMSD			0.09	0.18
MSE			0 ± 0.19	-0.15

tials can be obtained by properly describing both the gas-phase and solvation parts. Alternatively, good results can otherwise be obtained using empirical corrections[19,83,169,170] or error cancellation schemes.[109] 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.

2.5 V. Implications of the Results for Transition Metal Quantum Chemistry

The primary feature of the present paper is its evaluation of a significant number of quantum chemistry methods, both high level wavefunction based approaches and DFT functionals, via comparison with experimental gas phase ionization energies for two challenging series of transition metal containing molecules. All of the methods display some limitations – DFT and DLPNO-CCSD(T_0) in accuracy, and AFQMC and more accurate versions of coupled cluster (including full CCSD(T)) in the ability to scale up to larger systems such as the acac complexes.

Nevertheless, we consider the results to be very promising in a number of dimensions. The best of the DFT methods are within striking distance of achieving 'transition metal' thermochemical accuracy (2-3 kcal/mol MAEs) for the systems under study, and there is every reason to believe that progress towards this goal in the general case can be made if a larger and more relevant set of training data for transition metal containing systems is supplied to DFT developers. The effort towards enablement of AFQMC towards benchmark accuracy for large systems (i.e., ~ 2000 basis functions in a TZ basis) is well underway, although nontrivial problems remain (most prominently the ability to reliably generate sufficient trial wavefunctions).

Optimization of localized CCSD(T) methods such as DLPNO specifically for transition metal problems has reasonable prospects of ultimately enabling a scale up of coupled cluster based methods as well at the benchmark level of accuracy, although significant difficulties remain. See for example the recent work of Harvey and co-workers on non-heme iron complexes.[152,171] In the first work, they found that canonical CCSD(T) was in poor agreement with higher orders of coupled cluster and DMRG-CASPT2[171]. They also found that DLPNO-CCSD(T), even with tight PNO cut-offs, was not in agreement with canonical CCSD(T). In the second work, they found that these results hold for a larger system including the T_1 corrections on the DLPNO-CCSD(T) calculations.[152]. Our very preliminary results for $Mn(Cp)_2$ are in fact more encouraging than these conclusions.

Evaluation of full CCSD(T) for the systems in our present work, extrapolated to the basis set limit, is a highly computationally challenging calculation, which we did not attempt in the present paper. Our preliminary results for one metallocene, $Mn(Cp)_2$, do show significant movement towards the experimental data as compared to the initial DLPNO calculations as the quality of the triples correction is enhanced from (T_0) to (T), the PNO thresholds are tightened, and the basis sets size is increased. If a substantial amount of computational effort is committed to the problem, it should be possible to rigorously evaluate the performance of fully converged CCSD(T) for metallocenes ionization potentials, and we believe it is important to do so. Given the limited amount of relevant experimental data, and possibility of more noise in the experiments than is estimated in the experimental papers (always a concern for gas phase experiments on transition metal containing molecules), the convergence of multiple benchmark methods to similar results would be the best way to further validate the hypothesis (which we have proposed in the present work) that the right answer is being obtained by our AFQMC calculations for the right reason.

We are optimistic that, via parallel studies of increasingly accurate AFQMC and CCSD(T) methodologies, converged results can be obtained across a sufficient number of relevant cases to validate the creation of accurate benchmark data sets for transition metal containing species, at the very least for systems without a high degree of multireference character. As noted above, rigorous data from two high level methods provides significantly more confidence in estimated error bars than that from a single approach. Once a benchmark approach is validated, it can be used to develop new DFT functionals, and also approximate CCSD(T) protocols, with higher accuracy and reliability while retaining much lower computational costs and acceptable scaling with system size. Such developments are critical to enable the treatment of larger and more complex systems, where the scaling of the calculations with system size (and the prefactor) determine how many different states of the system (if any) can be investigated (and sometimes, for example in the case of the photosystem II water splitting cluster, a minimum of hundreds of states needs to be considered in the investigation of the catalytic mechanism).

We see the results to date obtained by our AFQMC implementation for a series of increasingly challenging, and diverse, transition metal test cases, as illuminating a path towards both improved high level approaches and optimized DFT functionals. 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),[172] other types of QMC such as diffusion Monte Carlo (DMC),[173] nonorthogonal configuration interaction with second order perturbation theory (NOCI+PT2),[174] and so on. When doing so, a balanced assessment of both accuracy and computational efficiency will be necessary.

The DFT results relative to our AFQMC calculations that we have obtained so far are consistent with our prior work. Semi-local GGA functionals appear to be incapable of obtaining reliable results for organometallic and coordination complexes of the type we have studied to date, although performance may be better for other metal-containing systems.[175] 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. There are benchmark data sets such as the MOR41 and ROST61 sets of (single-reference) closed-shell and open-shell organometallic reactions[41,176] which use DLPNO-CCSD(T) as the benchmark data set, the TMC151 set of diatomic dissociation energies as well as reaction energies and barriers for typical transition metal reactions which uses a mix of experimental and CCSD(T) reference values,[42] the MOBH35 database of 35 transition metal complex reaction barrier heights computed with DLPNO-CCSD(T) in a Weizmann-1 scheme.[177] Similarly, the bond dissociation energies of various transition metal fluoride complexes have also been studied with CCSD(T) as the theoretical benchmark.[178] The spin transition properties of several iron spin-crossover complexes have also been studied.[104] However, given the variable reliability of CCSD(T) for transition metals, especially open-shell systems, more robust benchmarks are desirable, even if more costly.

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,[167] 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,[179–184] may also prove to be useful in the optimization process.

2.6 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. For the same data sets, the results for various DFT functionals, and for DLPNO-CCSD(T_0) applied with loose cutoffs, display significantly higher MAEs as well as a number of large outliers (5-10 kcal/mol deviations from experiment) which are outside of the experimental error bars.

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,[13,14] or the CuO planes in high T_c superconductors.[185,186]

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.

An important question going forward will be the relative costs and scaling with system size of the best practices converged AFQMC and CCSD(T) approaches. Our most recent AFQMC results, in which we utilize a localized orbital approach to achieve cubic scaling, with a small prefactor, require only ~67 GPU hours for a TZ calculation on a metallocenes, a ~20x cost reduction as compared to the canonical orbital calculations cited above. We expect that both approaches will undergo significant algorithmic advances over the next few years, and this, combined with the usual acceleration from more powerful computational hardware, will play a major role in enabling the creation of large and diverse benchmark data sets suggested above.

Chapter 3: In silico prediction of annihilators for triplet–triplet annihilation upconversion via auxiliary-field quantum Monte Carlo

Reproduced from J. L. Weber *et al.*, *Chemical Science*, vol. 12, no. 3, 2021 with permission from the Royal Society of Chemistry.

3.1 Introduction

The relative energetic landscape involving states of different spin multiplicities is of essential importance in photoredox catalysis[187–189], the design of light emitting diodes[190], and optical processes such as singlet fission[191], thermally activated delayed fluorescence (TADF)[192], and upconversion[190,193]. In particular, for a system with a singlet ground state (S0), the most relevant quantities for these applications are typically the energies of the first excited singlet state (S1) and the lowest-lying triplet state (T1). Triplet-triplet annihilation (TTA) upconversion is a process which enables a system to emit photons of an energy higher than the energy of absorbed photons. This phenomenon has been used to increase the theoretical efficiency of photovoltaics[190,194], and to perform optogenetic manipulations and photocatalytic reactions with visible light in media (e.g. biological tissue) accessible only by photons of lower energy[187,195]. A schematic of TTA upconversion is shown in Figure 3.1. Following photoexcitation of a sensitizer to the S1 state, intersystem crossing (ISC) populates a relatively long-lived triplet state, T1. The sensitizer then undergoes Dexter triplet-triplet energy transfer (TET) to excite a separate molecular species, known as the annihilator, into a T1 state. Two annihilators excited to their T1 states can then undergo TTA to yield one annihilator in the S1 state and the other reverted to the ground S0 state[196]. Thus far, there are few families of annihilators capable of emitting high energy blue to near-UV light.[190,197]. These include 9,10 substituted anthracenes[190,197–200], *para*-terphenyl[201], pyrene,[202] and 2,5-diphenyloxazole.[203,204] Enlarging the chemical space of high energy upconverting annihilators would therefore represent a significant advancement towards the widespread use of photon upconversion for a variety of applications.

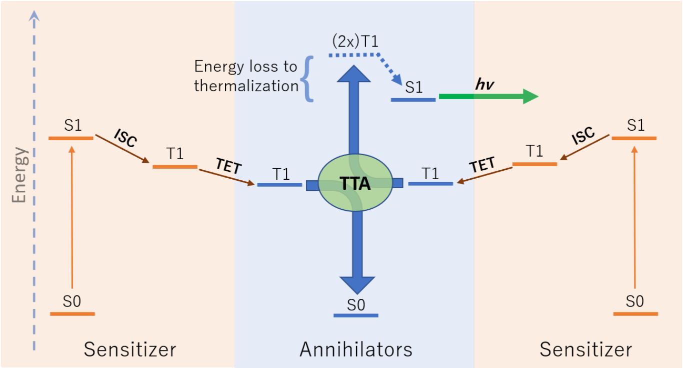


Figure 3.1: A schematic of photon upconversion via triplet-triplet annihilation (TTA). First, the sensitizer is photoexcited to the first excited singlet state (S1), before undergoing rapid intersystem crossing (ISC) to a long-lived triplet state. Collision with an annihilator enables transfer of the triplet state to an annihilator via Dexter triplet-triplet energy transfer (TET). Two annihilators in the T1 state can then undergo TTA in a spin-allowed transition resulting in one S1 and one ground state annihilator, the former of which can then emit a high energy photon via fluorescence. Note that in each step excess energy is lost as heat to the surroundings.

Thermodynamically, upconversion requires that a) the sensitizer T1 energy be higher than that of the annihilator for TET, and b) twice the annihilator T1 energy exceed the annihilator S1 energy for TTA.[205,206] However, the degree of exothermicity for both of these processes translates directly to the amount of thermal energy lost to heat during TET and TTA, respectively. When designing optimal sensitizer/annihilator pairs to minimize energetic losses, it is important to know the relative energy levels of these excited states. For example, in addition to high fluorescence yields, TTA annihilators should exhibit a minimally positive gap between twice the T1 and S1 to reduce energy loss to thermalization. While S1 energies can be extracted from experimental spectra (e.g. via estimation of the energetic location of the zero-phonon line), the triplet energy can be challenging to obtain experimentally[207–209]. The minimal (or lack of) phosphorescence is largely due to competing non-radiative pathways.

The inability to experimentally measure triplet energies has created a need which, in principle, can be met by predictions from *ab initio* computational methods. However, the development of a theoretical approach which is both accurate and feasible (with respect to computational costs) is far from trivial. The emergence of open-shell singlet ground states in large, conjugated aromatic systems reflects significant biradical, and even polyradical, character.[210] In addition, the excited states of cyclic aromatic molecules are known to be anti-aromatic[211] and thus similarly challenging for single-reference computational methods. These manifestations of strong electron correlation, in addition to potentially relevant phenomena such as excitations characterized by two-electron correlations and charge transfer, are well known to render commonly used computational techniques such as Kohn-Sham (KS-) or time-dependent (TD-) Density Functional Theory (DFT) unreliable[212,213].

Several methods have been shown to be promising for the description of spin gaps of potentially biradicaloid molecules, such as spin-projected orbital-optimized MP2[214], spin-flip methods[215,216], multi-configurational pair-DFT[172,217], various configuration interaction approaches[218,219], and optimized DFT functionals[220]. Recent efforts to reduce the scaling of Coupled Cluster (CC) methods, notably CC with singles, doubles, and perturbative triplets (CCSD(T)), have resulted in promising approaches based on domain-based localized pair natural orbital (DLPNO) approximations. Yet while these have extended the reach of CCSD(T) to larger systems,[103,221,222] the potential inadequacy of the underlying theory for strong correlation still remains.[223] Although higher order CC theories should in principle provide an increasingly accurate description, their application to relevant photoactive molecules is simply infeasible due to prohibitively high scaling with respect to system size.

We have observed that most computational results applied in the experimental literature of upconversion processes rely on TD-DFT for T1 and S1 excitation energies[206,224–226], despite known instabilities regarding the calculation of the T1 energy in the presence of spin-symmetry breaking[227,228]. In this work we survey three DFT functionals prevalent in the experimental literature, including the hybrid functional B3LYP, its range-separated counterpart CAM-B3LYP[229], and the highly parameterized meta-GGA M06-2X[77]. These have been shown to perform well within the Tamm-Dancoff approximation (TDA)[230] when benchmarked against MS-CASPT2[231]. Indeed tuning the extent of exact exchange included in hybrid DFT functionals such as these can lead to favorable cancellation of error in systems with similar charge transfer character[231]. However, the performance of such functionals is highly variable between different families of molecules[231–233], complicating efforts to predict novel TTA annihilators for upconversion *a priori*.

Phaseless auxiliary-field quantum Monte Carlo (hereafter referred to as AFQMC)[73,234] is a systematically improvable stochastic electronic structure method which scales modestly with the fourth power of the system size in our current implementation. It has recently been shown to produce accurate triplet energies for all linear polyacenes with experimentally reported T1 energies (naphthalene through pentacene) as well as for biradicals.[72] Recent algorithmic advances[56,57,235,236] have greatly reduced the computational costs of this methodology, enabling its use in the accurate prediction of novel chromophores, even those which may be strongly-correlated.

In this work we use AFQMC to compute T1 energies for a series of potential TTA annihilators. Anthracenes with two methyl substituents (DMA) or two phenyl substituents (DPA) are known TTA annihilators in optical upconversion schemes[200,237]. In Section 3.3.1 we generate candidate compounds by replacing the 9,10 substituents with various functional groups that are synthetically feasible, and probe the effects, if any, on the triplet energies. We then compute the triplet energies for a series of cyano-substituted anthracenes. In Section 3.3.1 we examine derivatives of benzothiadiazole (BTD), a compound widely used in donor-acceptor paradigms typically in the context of polymers[238,239]. It is known to have a fluorescent S1 state with an energy in the UV range (> 3 eV),[240,241] making this molecule and its derivatives potentially useful targets for TTA upconversion. We also investigate benzoselenodiazole (BSeD), which contains a selenium atom in place of sulfur. In Section 3.3.2, we validate the use of TD-DFT to predict adiabatic S1 energies by comparing with available experimental measurements. With an accurate computational protocol to predict both S1 and T1, we then assess the thermodynamic viability of upconversion for all molecules considered in this work by comparing twice T1 with S1. In Section 3.3.3 we present experimental upconversion outcomes for the phenyl-substituted BTD when coupled with platinum octaethylporphyrin (PtOEP) and zinc tetraphenylporphyrin (ZnTPP) sensitizers. This not only enables us to validate our AFQMC prediction for the triplet energy of Ph-BTD, but also provides our first example of the design of a novel, successful upconverting system informed by *ab initio* predictions. In section 3.3.4 we report phosphorescence measurements of the triplet energy for the BTD series, further validating the accuracy of AFQMC for Ph-BTD and the series as a whole.

3.2 Methods

3.2.1 AFQMC Methodology

AFQMC[54,242] utilizes imaginary-time propagation to stochastically sample properties associated with a given Hamiltonian via a random walk within the complex manifold of Slater determinants. The exponentially growing noise that would otherwise be incurred while averaging observables in imaginary-time is controlled by the use of a trial wavefunction to implement the phaseless constraint, at the expense of a bias which can be systematically reduced via improvement of the trial wavefunction. The lowest-energy state of each irreducible representation of the symmetry group of the Hamiltonian can be computed by AFQMC in the same manner as the ground state (which is a special example of such a state). Our singlet calculations have $N_\alpha = N_\beta$ and triplet calculations have $N_\alpha = N_\beta$+2. Properties of low-lying excited states belonging to the same irreducible representation can be obtained from the AFQMC methodology via the use of a trial wavefunction chosen such that it is orthogonal to eigenstates of lower-energy.[158,243] In practice, of course, the exact targeted eigenstate is unknown beforehand, necessitating the use of approximate wavefunctions obtained from other quantum chemical methods, which are typically nearly orthogonal to the ground-state. A spin filtration technique[159] allows us to preserve the total spin ($< S^2 >= 0$ and 2 for singlets and triplets, respectively) in the AFQMC projection. The use of trial wave functions which preserve or better approximate symmetries helps to improve results, as further discussed below.

The use of unrestricted single determinant trials has been shown to yield sub-kcal/mol accuracy for the triplet energies of polyacenes with closed-shell ground-states, and many biradicaloid molecules with open-shell singlet states that can be qualitatively described by two determinants.[72,113] However, some highly multi-reference systems such as transition metal compounds require the use of non-orthogonal determinant expansions[244] or truncated CASSCF trial wavefunctions[47,48,57] to yield high accuracy. In this work, all AFQMC calculations implement unrestricted single-determinant trial wavefunctions selected according to the AFQMC/U protocol,[72] except for those on the BTD and BSeD derivatives, which were found to exhibit signs of strong correlation (*vide infra*) and thus required truncated CASSCF trials. As the lowest excited states for such conjugated molecules are π to π^* transitions[245], we use active spaces spanned by all valence π-orbitals. In the case of phenyl-substituted BTD/BSeD, the resulting active spaces were intractable, and so the three highest and lowest virtual and occupied orbitals, respectively, were neglected in active space optimization. Trial wavefunctions and all required integrals for AFQMC calculations were obtained using PySCF.[70] Extrapolations to the the complete basis set (CBS) limit were performed using DLPNO-CCSD(T) values in TZ and QZ dunning basis sets[64,246] (see Ref. 47 for details of this protocol), and dielectric solvation corrections were computed using a simple conductor-like polarizable continuum model (CPCM) at the B3LYP/TZ level. Further details regarding the AFQMC calculations can be found in the SI.

3.2.2 DFT and DLPNO-CCSD(T) Calculations

KS-B3LYP, TD-DFT, and DLPNO-CCSD(T)[50,51] calculations were performed with the ORCA quantum chemistry program[68]. S0 and T1 geometries were optimized at the KS-B3LYP/cc-pVTZ level of theory. The reference wavefunctions for DLPNO-CCSD(T) calculations were chosen as follows. As large deviations from the exact S^2 values were found for both S0 and T1 states of the anthracene derivatives at the UHF level, inconsistent with the stable closed-shell nature of acenes of this length,[247] we utilize restricted orbitals for the anthracene derivatives (RHF/ROHF for S0/T1). For the BTD series, we use UHF reference wavefunctions. The semi-canonical approximation to the triples correction, DLPNO-CCSD(T_0), was used[50]. Henceforth, DLPNO-CCSD(T) will be refer to DLPNO-CCSD(T_0). The “NormalPNO" cutoff was used for all DLPNO-CCSD(T) calculations[50].

For TD-DFT calculations of adiabatic S1 energies, we correct the vertical excitation energy (with respect to S0 geometries) with a relaxation term, obtained from geometries which reflect the minimum energy of the target excited state within the TDA approximation. Subsequent single-point excitation energies were then computed without the TDA approximation. Regarding T1 calculations via TD-DFT, it has been found that triplet instabilities can lead to an unphysical underestimation of T1 energies especially when using functionals with a significant percentage of exact exchange, and that employing TDA can help to ameliorate this error[227]. Since in the anthracene set all molecules exhibit notable spin contamination in the singlet state, we utilize the TDA approximation when calculating the triplet energy using TD-DFT, specifically when calculating the vertical excitation energy corresponding to the optimized geometry of S0. To report adiabatic T1 energies, we correct the vertical excitation energy with T1 geometry relaxation energies, obtained by adding the difference in total KS-DFT/B3LYP T1 energies between the optimized S0 and T1 geometries to the vertical excitation energies.

For a subset of molecules in Section 3.3.1 we investigated the importance of supplementing gas-phase electronic energy gaps with vibrational and solvation effects. We found (Table 3.10) that inclusion of the above effects did not change the calculated triplet energies by an amount larger than the statistical error bars of AFQMC, and thus while our results for the anthracene derivatives in this paper reflect gas-phase electronic gaps, we expect these to be close to what would be realistically measured in toluene solvent. For the BTD series, in particular MeO-BTD, which exhibits strong charge transfer characteristics, the dielectric solvation corrections were not negligible, and so our calculated values reflect a correction term obtained from separate calculations employing the CPCM continuum solvation model. All calculations for the anthracenes use the cc-pVTZ basis set, as it was found in every case to be near the complete basis set (CBS) limit. The calculated triplet energies for the BTD and BSeD series have been extrapolated to the CBS limit (using X=T,Q basis sets, and a $1/X^3$ form for the correlation energy). For the selenium complexes we use the cc-pVXZ-dkh basis set and the x1c formalism to include scalar relativistic effects. We refer the reader to the SI for further information.

3.2.3 Experimental Methods

Details for the synthesis of Ph-BTD can be found in Section S1 of the SI. All starting materials were obtained from commercial sources, including Fisher Scientific, TCI Chemical, and Strem Chemicals. BTD (ACROS Organics), ZnTPP (Fisher Scientific), and PtOEP (Sigma-Aldrich) were purchased and used without further purification.

NMR spectra were collected on a Bruker 500 MHz spectrometer at ambient temperature. UV-Vis absorption spectra were collected by a Technologies Cary 60 UV-Vis spectrophotometer. Steady-state photoluminescence spectra were collected by an Ocean Optics QEPro spectrometer.

Solution concentrations for photon upconversion studies were prepared as 1x10^{-5} M sensitizer and 1x10^{-3} M annihilator in degassed anhydrous toluene. Solutions for each sensitizer-annihilator pair were made in a nitrogen glovebox, sealed, and removed from the glovebox for upconversion photoluminescence study.

Phosphorescence measurements were taken at 77K in a frozen solution of methylcyclohexane (BTD/CN-BTD) and methylcyclohexane/iodomethane (2:1 v/v) (MeO-BTD, Ph-BTD) (details in SI).

3.3 Results

3.3.1 Calculating Accurate Triplet Energies for TTA Upconversion Annihilators

Anthracene 9,10 Functionalization

As a preliminary test, to investigate the accuracy of unrestricted single-determinant trials for substituted acenes, we compared AFQMC/UHF and AFQMC/UB3LYP with AFQMC/CAS for benzonitrile, a small but representative system for which large CASSCF trial wavefuctions (and thus near exact AFQMC energies) can readily be obtained. In previous studies we have shown that using such trial wavefunctions can largely eliminate the bias from the phaseless constraint such that the resulting predictions agree well with experimental measurements.[1,47,57,72] For benzonitrile we use an active space of 8 electrons in 16 orbitals (8e16o), representing the full π system plus

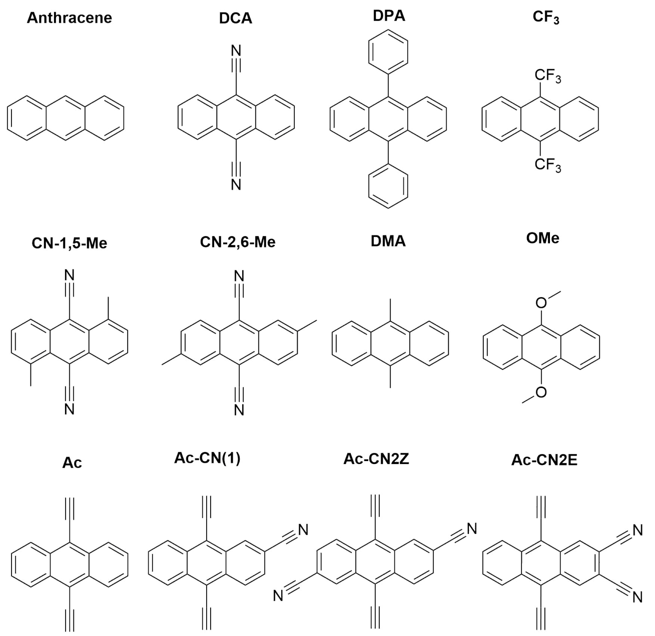


Figure 3.2: Anthracene derivatives included in this study.

a second set of virtual orbitals. The results are shown in Table 3.1. While all methods produce triplet energies above the lower bound from experiment, the result from the KS-B3LYP trial is within 0.01 eV of that from the CASSCF trial. This is consistent with our previous validation of the AFQMC/U protocol for small-molecule biradicals and unsubstituted acenes, in which UHF is used as a trial unless there is significant spin contamination (in the case of benzonitrile singlet, where $< S^2 >= 0.59$), in which case an unrestricted Kohn-Sham (UKS) trial is used.[72] A similar protocol has been shown to improve the accuracy of CC methods[248].

Table 3.1: AFQMC results from various trial wavefunctions for the adiabatic triplet energy of benzonitrile, in eV. Parentheses denote statistical error of AFQMC, i.e. 3.61(6) denotes 3.61 ± 0.06.

AFQMC/UHF	AFQMC/UKS	AFQMC/CAS	Expt.[249]
3.86(9)	3.62(8)	3.61(7)	>3.35

Figure 3.3 presents adiabatic triplet energies obtained from KS-DFT, TD-DFT with three different representative functionals, DLPNO-CCSD(T), and AFQMC/U for the functionalized anthracenes shown in Figure 3.2, along with mean absolute deviations (MADs) between each method and AFQMC shown in Table 3.2. In nearly every case, the DLPNO-CCSD(T) and AFQMC/U re-

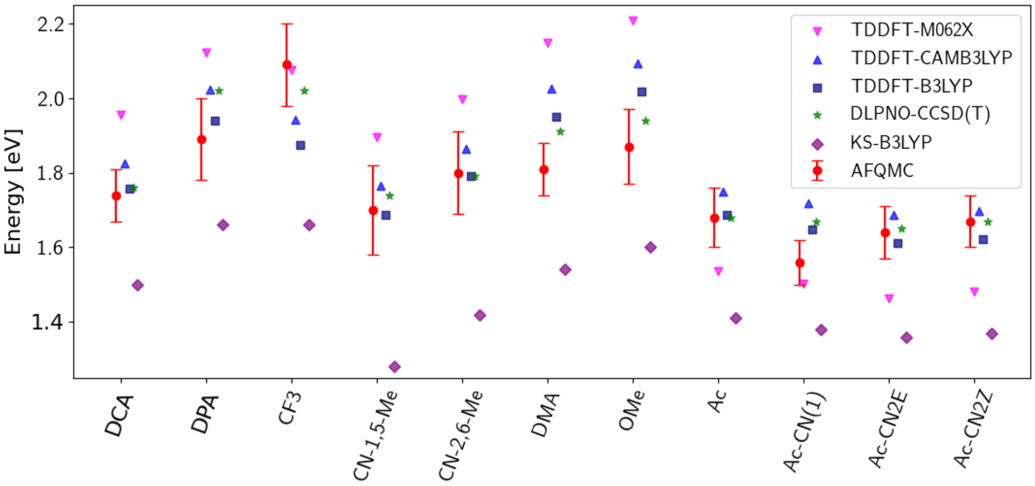


Figure 3.3: A comparison of T1 values for included TD-DFT functionals, KS-B3LYP, and DLPNO-CCSD(T). Note that KS-B3LYP obtains significantly lower triplet energies when compared to AFQMC for the entire series, whereas DLPNO-CCSD(T) is within one standard deviation from AFQMC for the majority of the compounds involved, with a mean average deviation from AFQMC below the average statistical error of the latter. All DLPNO-CCSD(T) and AFQMC values reflect gas-phase calculations in the cc-pVTZ basis. Numbers for these, as well as CBS/solvation corrections, can be found in the SI.

Table 3.2: A comparison in eV of DLPNO-CCSD(T), KS-DFT, and TD-DFT results for T1 of anthracene derivatives, including mean absolute deviation (MAD), mean signed deviation (MSD), and maximum deviation (Max) versus AFQMC/U. Both DLPNO-CCSD(T) and TD-B3LYP have an MAD below the average statistical error of AFQMC (0.09 eV), although TD-B3LYP exhibits a higher maximum deviation of 0.215 eV vs AFQMC.

	KS-B3LYP	TD-B3LYP	TD-CAM-B3LYP	TD-M062X	DLPNO-CCSD(T)
MAD vs AFQMC	0.297	0.070	0.112	0.191	0.051
MSD vs AFQMC	-0.297	0.013	0.085	0.085	0.037
Max vs AFQMC	0.430	0.215	0.223	0.338	0.135

sults agree to within the statistical error bars of the latter, with the MAD between AFQMC/U and DLPNO-CCSD(T) (0.05 eV) being less than the mean statistical error from AFQMC/U (0.09 eV).

We are aware of only one direct experimental measurement of the triplet energy in a comparable solvent for this set of molecules, namely for DCA in toluene, which has a value of 1.8 eV[250]. Both DLPNO-CCSD(T) and AFQMC/U are in good agreement with this value, whereas KS-DFT with the B3LYP functional systematically underestimates the gap. In addition, we previously reported an AFQMC/U value for anthracene within 0.04 ± 0.05 eV of a gas phase experimental measurement[72,247]. Recently, DPA was reported to have a triplet energy of about ≃ 1.75 eV, measured in a polymer host matrix consisting of poly(4-bromostyrene) and benzophenone[251]. Neglecting the experimental uncertainty (which was not reported), this is slightly outside of the error bars of our AFQMC calculation (1.89 ± 0.11 eV). We postulate that this possible, small discrepancy is due to the environment of the experiment. We also note that, among the theoretical methods considered (barring KS-DFT/B3LYP, which underestimates the triplet energy of DCA by ≃ 0.3 eV), AFQMC yields the closest value to experiment. These available comparisons suggest that AFQMC provides reliable predictive power for this class of anthracene derivatives.

BTD/BSeD Based TTA Annihilators

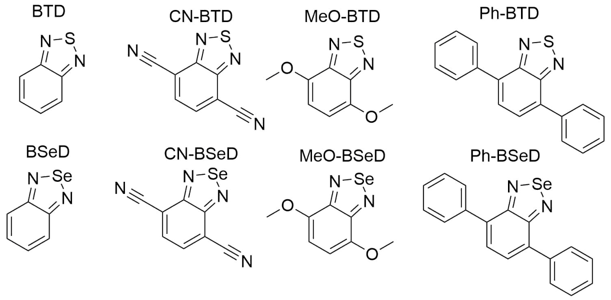


Figure 3.4: BTD and BSeD derivatives included in this study.

In this section we investigate the triplet energies of a set of synthetically-feasible derivatives of benzothiadiazole (BTD) and benzoselenodiazole (BSeD), shown in Figure 3.4. We find that these molecules exhibit a substantial degree of electron correlation, e.g. the CASSCF wavefunctions for S0 and T1 of Ph-BTD contain roughly 40k and 60k determinants, respectively (representing 99.5% of the sum of squares of CI coefficients). In this regime, AFQMC/UKS is no longer expected to produce accurate results (indeed, AFQMC/UKS and AFQMC/CAS produced results differing by 0.26 ± 0.08 eV for BTD); we therefore use AFQMC/CAS. It is known that initializing CASSCF active spaces with the full π system, as identified visually at the restricted HF level, is necessary for quantitative results in conjugated aromatics[245]. We follow this protocol for all systems except those with phenyl groups, in which case we had to exclude the lowest three occupied orbitals and highest three virtuals from the active space due to computational limitations. Due to the large computational cost of these calculations, only the first ~500 determinants were maintained in the CASSCF trials, which still represented over 94% of the CI weights for each molecule.

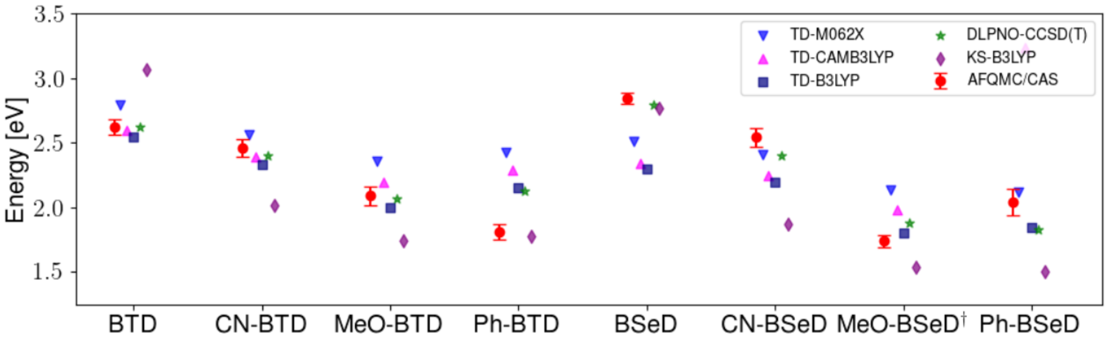


Figure 3.5: A comparison of the triplet energies calculated with TD-DFT, KS-B3LYP, and DLPNO-CCSD(T). Once more KS-B3LYP obtains significantly lower triplet energies as compared to AFQMC for nearly all species, except notably for the parent BTD and BSeD compounds, and Ph-BTD. Ph-BTD stands out as an outlier, with the maximum deviation of DLPNO-CCSD(T) from AFQMC at ≃ 0.3 eV. Numbers can be seen in the SI. †MeO-BSeD represents AFQMC in the triple-ζ basis (i.e. not CBS limit).

Table 3.3: A comparison in eV of DLPNO-CCSD(T), KS-DFT, and TD-DFT results for T1 of a set of substituted BTD and BSeD compounds benchmarked against AFQMC/CAS, including mean absolute deviation (MAD), mean signed deviation (MSD), and maximum deviation (Max) versus AFQMC/CAS. All methods have significantly higher maximum deviations from AFQMC than was found for the anthracenes, as is expected given the larger degree of electron correlation observed in these compounds. DLPNO-CCSD(T) and TD-B3LYP again have the lowest and second lowest MADs vs AFQMC, respectively.

Species	KS-B3LYP	B3LYP	CAM-B3LYP	M062X	DLPNO-CCSD(T)
MAD vs AFQMC/CAS	0.35	0.22	0.36	0.26	0.12
MSD vs AFQMC/CAS	-0.23	-0.12	0.14	0.15	0.00
Max vs AFQMC/CAS	0.67	0.55	1.20	0.62	0.32

Whereas the anthracenes exhibit negligible basis set incompleteness effects (Table 3.6), this is not the case for the BTDs and BSeDs, and so all AFQMC and DLPNO-CCSD(T) numbers for these molecules reflect an extrapolation to the CBS limit. Additionally, the relatively more substantial charge-transfer character in some cases, vs the anthracenes, can lead to a significant solvent correction, e.g. a shift of -0.15 eV for MeO-BTD. For consistency, we therefore include the correction from the dielectric continuum model for all BTD and BSeD derivatives.

While for single-reference systems, i.e. those that can be well-described by one orbital-occupancy configuration, DFT and CCSD(T) methods are capable of producing robust accuracy, we can be less confident that these methods will produce accurate T1 energies for the BTD and BSeD derivatives. Interestingly, we find good agreement between DLPNO-CCSD(T) and AFQMC, excepting the case of Ph-BTD. Screening for spin contamination in the stable UHF references revealed minimal spin contamination for all triplet species (except Ph-BTD), and significant deviations from the exact value (0) for all singlets. This implies that states of different spin-multiplicities (e.g., singlet, triplet, quintet) are sufficiently close in energy that they “mix" to lower the energy at the mean-field level (at the expense of spin symmetry breaking). The determinant constructed from unrestricted Kohn-Sham (UKS) orbitals removed the spin contamination, and using this as a reference wavefunction resulted in a nearly equivalent DLPNO-CCSD(T)/cc-pVTZ result for all species, suggesting that the use of spin-contaminated reference orbitals cannot account for the deviation from the AFQMC result. In Section 3.3.3 we will show experimental evidence which suggests that the triplet energy as predicted by AFQMC/CAS is accurate. TD-DFT with the B3LYP functional performs best (with respect to AFQMC) among the DFT methods investigated, while inclusion of long-range HF exchange with the CAM-B3LYP functional worsens the MAD by more than a factor of two. A plot of the calculated triplet energies for AFQMC/CAS and alternate methods can be seen in Figure 3.5. It should be noted that all methods follow the same general trend, where the triplet energy of BTD > CN-BTD > MeO-BTD > Ph-BTD, consistent with S1 calculations that are presented and rationalized based on π-system extension and donor-acceptor paradigms in Table 3.13.

3.3.2 Predicting Upconversion Activity

S1 Energies

In contrast to T1, S1 can readily be measured experimentally. However, in order to make predictions about the thermodynamics of new potentially upconverting systems, it is necessary to accurately calculate S1. Previous studies have shown that the CAM-B3LYP functional yields S1 energies of extended polyaromatics that are very close to experimental measurements.[254229227] In

Table 3.4: CAM-B3LYP TD-DFT results for S1 energies in eV for all substituents, including a subset of CN substituted tetracenes, with available experiments. The MAE was found to be 0.056 eV for all available experiments. Structures for the tetracenes can be found in the SI.

Species	S1 (TD-DFT)	Expt	Difference
Tetracenes:			
CN0	2.30	2.30	0.00
CN1	2.28	2.26	0.02
CN2T	2.25	2.21	0.04
CN2E cis	2.25	2.23	0.02
CN2H	2.19	2.16	0.03
CN3	2.23	2.20	0.03
CN4	2.21	2.19	0.02
Anthracenes:			
DPA	3.14		
DMA	3.10		
OMe	3.06		
CF3	3.03		
CN-2,6-Me	2.93		
DCA	2.92	2.90[250]	0.02
CN-1,5-Me	2.79		
Ac	2.88	2.80	0.08
Ac-CN1	2.82	2.71	0.11
Ac-CN2E	2.78	2.66	0.12
Ac-CN2Z	2.77	2.65	0.12
BTD derivatives:			
BTD	3.86	<3.97[241]	<0.11
CN-BTD	3.64		
MeO-BTD	3.05	3.13[252]	0.08
Ph-BTD	3.04	3.08[253]	0.04
BSeD	3.57		
CN-BSeD	3.41		
MeO-BSeD	2.78		
Ph-BSeD	2.85		

Table 3.4 we have collected a set of conjugated potential annihilators for which experimental S1 energies are available. This set supplements the molecules in this study with 7 tetracene derivatives containing 0-4 cyano substituents. The mean absolute error (MAE) with respect to experiment is 0.056 eV, giving us confidence that this functional can be used to compute S1 energies for these molecules with sufficient accuracy (i.e. comparable to the statistical error bars on our AFQMC T1 calculations). We note that it is possible to obtain S1 energies with AFQMC via an appropriately imposed symmetry constraint, and present an example computing S1 and T1 for anthracene in the SI. Given the demonstrated accuracy of TD-DFT methods, we leave this for future work.

Energetic Efficiency of TTA Upconverting Candidates

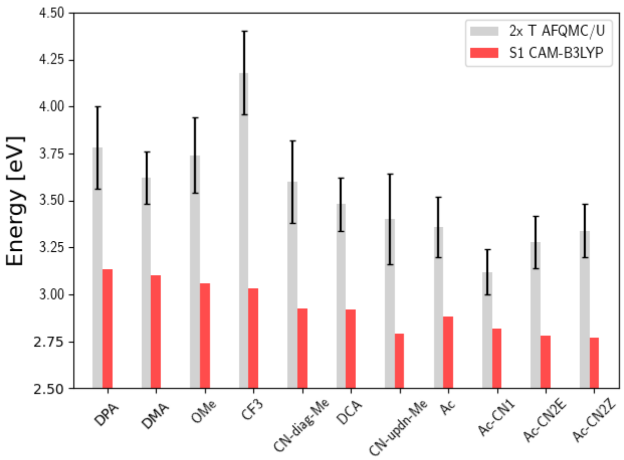


Figure 3.6: Comparison of the predicted gas phase 2x T1 energetic value, as calculated using AFQMC/U, with S1 values obtained from TD-DFT/CAM-B3LYP for the Anthracene derivatives. Note that the mono-substituted CN-anthracene, Ac-CN1, exhibits the lowest difference between 2x T1 and S1, and therefore the lowest potential energy loss during TTA. Further note the destabilization of the triplet state for the highly inductively-withdrawing CF_3 substituted species, versus e.g. the π system electron withdrawing DCA.

While the inequality 2*T1 > S1 is a thermodynamic prerequisite for upconversion, achieving efficiencies necessary for practical applications may require additional considerations. For example, it is often preferable to minimize the energy loss during TTA by engineering 2*T1 - S1 to be minimally positive[198]. In Figures 3.6 and 3.7 we compare 2*T1, as predicted via AFQMC, with S1, as predicted from TD-DFT/CAM-B3LYP. Among the anthracene derivatives, Ac-CN1 is pre-

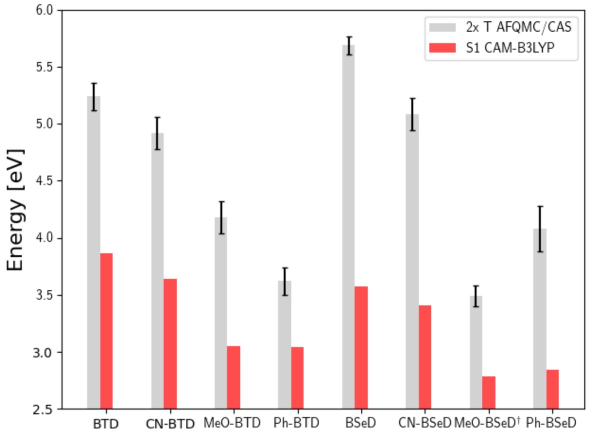


Figure 3.7: Comparison of the predicted gas-phase 2x T1 energetic value as calculated using AFQMC/CAS, vs S1 using TD-DFT/CAM-B3LYP for the BTD/BSeD derivatives. All molecules exhibit 2xT1 > S1, and therefore are exothermic towards upconversion. Substitution on the phenyl ring of BTD leads to a lowering of both the singlet and triplet excited states in all cases, more notably for the electron donating functional groups. Ph-BTD exhibits the lowest 2xT1 - S1. †MeO-BSeD represents AFQMC in the triple-ζ basis (i.e. not CBS limit).

dicted to be the most efficient annhilator by this metric. Among the BTD and BSeD compounds, 2*T1 - S1 is smallest for Ph-BTD.

We note in passing that the trends in the S1 gaps for the molecules shown in Figures 3.6 and 3.7 can be qualitatively predicted by simple models that describe extended conjugated molecules, e.g. particle-in-a-box π-extension and donor-acceptor (charge transfer) paradigms. A discussion rationalizing S1 energies in these molecules is presented in the SI. Similar trends in triplet energies are found, albeit with notable outliers, such as CF_3-anthracene and Ac-CN1. These observations, particularly the discrepancies between trends in S1 and T1, further emphasize the need for quantitatively accurate *ab initio* electronic structure methods for the calculation of triplet energies.

3.3.3 Observation of Upconversion

Two observations motivated us to experimentally investigate Ph-BTD. First, the predicted triplet energies for Ph-BTD via all TD-DFT methods and DLPNO-CCSD(T) are significantly larger than that predicted by AFQMC/CAS, by 0.3-0.67 eV, representing a significant discrepancy between traditional electronic structure methods and AFQMC. Second, of the BTD and BSeD series, Ph-BTD is predicted (by AFQMC) to have the smallest energetic loss from TTA. We therefore decided to experimentally test for upconversion activity by coupling the Ph-BTD annihilator with two different sensitizers, platinum octaethylporphyrin (PtOEP) and zinc tetraphenylporphyrin (ZnTPP), with known experimental triplet energies of 1.91 eV[255] and 1.61 eV[256], respectively.

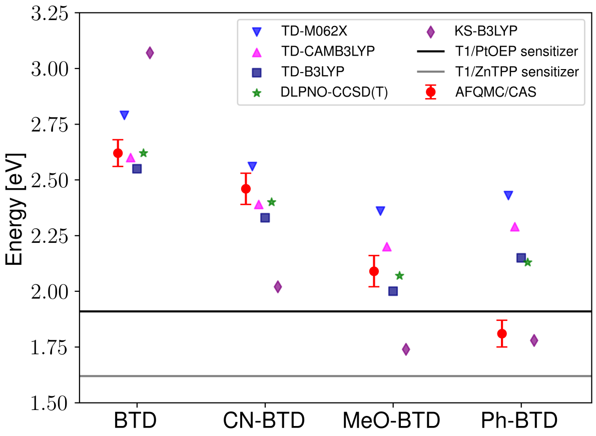


Figure 3.8: Predicted AFQMC values versus experimental sensitizer triplet energies for the BTD series. The sensitizer triplet energy is generally required to be above that of the annihilator in order to perform efficient, exergonic triplet-triplet energy transfer, and so we expect Ph-BTD to upconvert when paired with PtOEP.

In upconverting systems, the initial population of a sensitizer's S1 state via photoexcitation is followed by ISC to the sensitizer's T1 state, and then by TET, in which the energy of the T1 state of the sensitizer is transferred to form the T1 state of the annihilator (Figure 3.1). TET is thermodynamically allowed when the triplet energy of the annihilator is downhill from that of the sensitizer. As can be seen in Figure 3.8, the PtOEP and ZnTPP sensitizer triplet energies effectively sandwich our AFQMC-predicted triplet energy for the Ph-BTD annihilator, 1.77(6) eV. We can thus expect that if our AFQMC prediction is correct, the PtOEP/Ph-BTD system should be able to upconvert, whereas the ZnTPP/Ph-BTD system should not.

Indeed, Ph-BTD exhibits the ability to upconvert when coupled to a PtOEP sensitizer, with an anti-Stokes shift of approximately 0.2 eV from the excitation energy to the peak emission of the system, as seen in Figure 3.9. This provides evidence for the triplet energy of Ph-BTD being below 1.91 eV, consistent with our AFQMC predictions. Note that none of the TD-DFT results

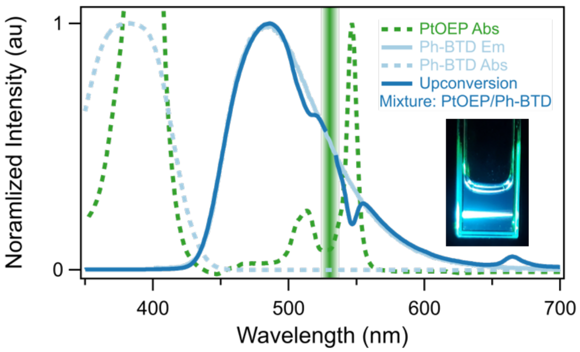


Figure 3.9: Photoluminescence confirmation of upconversion (dark blue) by the Ph-BTD/PtOEP system in toluene upon excitation with 532 nm light (green line), absorption of PtOEP (green dashed), absorption (light blue dashed) and photoluminescence (light blue solid) spectra of Ph-BTD. Note that the Ph-BTD/PtOEP upconversion system emits at a higher energy than the excitation wavelength and that the Ph-BTD does not directly absorb light at the excitation wavelength. Visual observation (insert) corroborates this measurement.

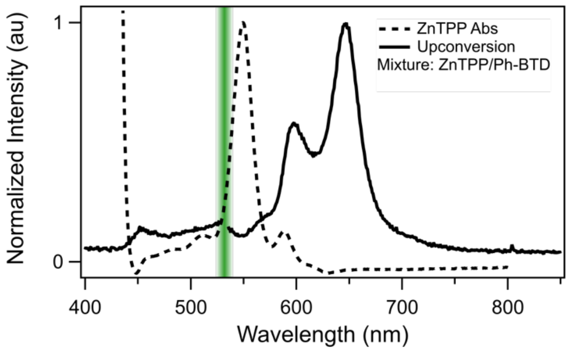


Figure 3.10: Absorption (dashed) spectrum of ZnTPP, and photoluminescence (solid) of the Ph-BTD/ZnTPP pair in toluene with excitation at 532 nm (green line). The emission of the mixture matches that of ZnTPP (Figure 3.12), signifying inefficient TET to Ph-BTD, as predicted by the relative T1 energy levels.

are consistent with this observation, and neither is DLPNO-CCSD(T). KS-DFT with the B3LYP functional is consistent with this observation, but as it underestimates triplet energies for most compounds it is most probable that this agreement is fortuitous. On the other hand, the mixture of ZnTPP and Ph-BTD shows phosphorescence of the sensitizer (Figure 3.10), indicating that prominent upconversion does not occur, thus supporting our prediction that the triplet energy of Ph-BTD is too large for effective TET from ZnTPP. These two experimental observations imply that 1.61 eV < Ph-BTD(T1) < 1.91 eV, consistent with our AFQMC/CAS prediction.

We comment that endothermic entropically-driven, endothermic TET has previously been reported in the literature[195,257,258]. In these representative instances, TET is found to be up-hill by <0.1 eV, which is a significantly smaller number than the smallest TDDFT-predicted endothermicity of PtOEP/Ph-BTD, which is 0.25 eV (B3LYP). This very extreme endothermicity, in light of literature precedent, most likely would not allow forward TET. This strengthens the claim that these TDDFT and DLPNO-CCSD(T) methods are unreliable in predicting the triplet energy of this PtOEP/Ph-BTD system.

3.3.4 Comparison to Low Temperature Phosphorescence

We performed low-temperature phosphorescence for the series of functionalized BTD compounds at 77K frozen in methyl-cyclohexane. The resulting spectra can be seen in the SI, and the estimated 0-0 triplet energies are shown in Table 3.5. Importantly, AFQMC/C calculations in the CBS limit and with implicit solvent corrections agree with experimental triplet energies to within 0.11 ± 0.07 eV for MeO-BTD and to within 0.04 ± 0.06 eV for Ph-BTD. The experimental triplet energies of BTD and CN-BTD are reported to be ≃0.28 and ≃0.38 eV below that of the AFQMC-predicted values; in general, none of the electronic structure predictions were consistent with these experiments, though KS-B3LYP was within 0.1 eV for CN-BTD (it should be noted that KS-B3LYP was off by ≃0.7 eV for BTD).

Species	Experiment	AFQMC/C	KS-B3LYP	B3LYP	CAM-B3LYP	M062X	DLPNO-CCSD(T)
BTD	2.33	2.62(6)	3.07	2.55	2.60	2.79	2.62
CN-BTD	2.08	2.46(7)	2.02	2.33	2.39	2.56	2.40
MeO-BTD	1.98	2.09(7)	1.74	2.00	2.20	2.36	2.07
Ph-BTD	1.85	1.81(6)	1.78	2.15	2.29	2.43	2.13

Table 3.5: Experimental (phosphorescence) estimations of the 0-0 triplet energy, compared against computational predictions for the monomer (CBS with dielectric solvation corrections); BTD is severely overestimated by all methods, whereas CN-BTD is overestimated by all but KS-B3LYP. In the case of Ph-BTD, AFQMC/C agrees to within statistical error, and indeed exhibits the lowest deviation from experiment, closely followed by KS-B3LYP.

Due to the significant discrepancy between all computational methods and experiment for BTD and CN-BTD, we attempted to pinpoint global sources of computational error. A literature search revealed that there is some precedence for favorable anti-square dimerization of benzothiadiazoles[259]. To explore this, we estimated dimerization free energies at the ωB97X-V/cc-pVTZ-DK level of theory, which suggest that both BTD and CN-BTD exist as gas phase dimers at 77K. Additionally, DLPNO-CCSD(T)/cc-pVTZ calculations of the dimers effectively reduce the error for BTD and CN-BTD from 0.29 eV and 0.32 eV to 0.09 eV and 0.21 eV (see SI for further details).

3.4 Discussion

The results of this study serve as a caution to practitioners relying on DFT methods to predict triplet energies of various types of molecules, especially in the absence of careful, system-specific benchmarking, despite the convenience resulting from the speed, black-box nature, and frequent accuracy of such calculations. The data suggest that, of the DFT-based approaches, TD-B3LYP shows the highest level of accuracy with respect to AFQMC reference values, and on average its predictions lie within the statistical error bars of the AFQMC calculations for the weakly correlated anthracene derivatives. This is consistent with Ref. 233, which found similar accuracy for a set of annihilators including diphenyl anthracene. The so-called “gold standard" of traditional electronic structure theory, CCSD(T), here represented by the DLPNO-CCSD(T) variant, also shows outstanding accuracy for the anthracene series, with a maximum error of just 0.135 eV. A judicious choice of trial wavefunction for AFQMC, based on the AFQMC/U formalism described in Ref. 72, is shown to be a promising tool for fast triplet energy screening, with all such calculations taking $\simeq$2 hours of wall time on the Summit supercomputer.

However, when extending the data set to the BTD and BSeD series, which exhibit charge transfer characteristics and significant electron correlation effects, the accuracy of all TD-DFT functionals notably deteriorates as compared to the AFQMC reference values, with maximum deviations between 0.4 and 1.25 eV. DLPNO-CCSD(T) exhibits good agreement with AFQMC except for the case of Ph-BTD, where DLPNO-CCSD(T) overestimates the triplet energy by around 0.3 eV. While this discrepancy might be an artifact of unsuitable localization thresholds utilized in the default DLPNO implementation, our effort to use a more mild approximation proved intractable, highlighting the computational cost of the underlying CCSD(T) method. In the outlier case of Ph-BTD, the accuracy of our AFQMC prediction is experimentally validated by pairing with two sensitizers of known triplet energies, which provides further evidence that AFQMC can reliably produce quantitatively accurate relative spin state energetics for a wide variety of medium-sized organic molecules at an affordable computational cost. Additional low temperature phosphorescence measurements of the triplet energies of Ph-BTD provides further evidence for the enhanced predictive accuracy of AFQMC for the spin gaps of these organic systems. It is notable that in some cases Kohn-Sham B3LYP is surprisingly accurate, while in others it is wildly inaccurate; e.g., for Ph-BTD, only KS-B3LYP correctly predicts exothermic TTA (along with AFQMC), but is off by 0.7 eV for the unsubstituted BTD.

Interestingly, a notable deviation between experiment and computational predictions was found for BTD and CN-BTD. This discrepancy would be ameliorated somewhat by including considerations of weak dimerization at low temperatures[see SI]. We note, however, that the possibility of dimerization is in conflict with both the very low concentration of BTD ($\simeq \mu$M) in the phosphorescence experiments, and the speed at which the solution is cooled, and we do not consider this further. As the discrepancy for all electronic structure methods screened is additionally not due to basis set errors, and both AFQMC/C and DLPNO-CCSD(T) agree, these two cases warrant further investigation. Even with these notable outliers, AFQMC exhibits the lowest deviation (0.17 ± 0.07 eV) from available experiments (for DCA, DPA, and the BTD series) out of the computational methods screened, statistically equivalent to TD-B3LYP. But we note that in the important case of Ph-BTD, the newly discovered annihilator for TTA upconversion, TD-B3LYP overestimates the experimental triplet energy by some 0.3 eV.

A few comments are now in order, regarding the significance of our present discovery of the PtOEP/Ph-BTD upconverting system. While the reported anti-Stokes shift is not particularly remarkable compared with those of some existing blue or near UV emitting annihilators,[197,203,260–262] our calculations suggest that Ph-BTD can achieve notably high energetic efficiency, i.e. minimal energetic loss during TTA upconversion. Moreover, we have demonstrated that derivatives of the

BTD core are a new class of aromatic molecules that can participate in TTA upconversion, expanding the growing library of high energy annihilator structures. With appropriate sensitizer pairings, the BTD and BSeD derivatives investigated here are predicted to satisfy the thermodynamic requirements for photon upconversion, and to emit in the range of 2.8-3.9 eV. The computational methods validated in this work provide a platform for the rational design of novel upconverting systems, which can both screen for energetic efficiency and provide a link between chemical functionalization and tunable photophysical properties. With these tools as a guide, further investigations into unexplored corners of chemical space for the BTD and BSeD series are under way.

3.5 Conclusions

We have found that AFQMC is an *ab initio* methodology that is accurate in its predictions of triplet energies and is scalable to realistic systems relevant to photophysical processes such as upconversion. We provide predictions for a variety of known and potential annihilators designed by adding substituent groups to anthracene, BTD, and BSeD frameworks. We find that triplet energies calculated from DFT and DLPNO-CCSD(T) methods show minimal deviations from the AFQMC values in the case of the anthracenes, with the B3LYP functional in the context of TD-DFT providing accuracy comparable to DLPNO-CCSD(T). Investigation of the BTD and BSeD series led to similar agreement among the theoretical approaches, with the notable exception of Ph-BTD, for which DLPNO-CCSD(T) and all TD-DFT methods overestimated the triplet energy by $\simeq 0.35$ to 0.60 eV compared to AFQMC. The AFQMC predictions are supported by experimental evidence of the occurrence of TET when Ph-BTD is coupled to a sensitizer with a larger triplet energy (PtOEP), but not when coupled to one with a smaller triplet energy (ZnTPP). Additionally, low-temperature phosphorescence measurements of Ph-BTD agree to within 0.04 ± 0.06 eV of AFQMC. Large deviations from phosphorescence values for BTD and CN-BTD for all methods were found, though the possibility of dimerization due to weak chalcogen bonding deserves further investigation.

Together with calculated S1 energies from the CAM-B3LYP/TD-DFT, which were shown to accurately predict a set of experimental measurements, the AFQMC triplet energies were used to investigate the energetic efficiency of TTA for all molecules. This led to the discovery of a novel annihilator, Ph-BTD, which when coupled to PtOEP emits upconverted blue light. This system exhibits an encouragingly small energy difference between twice T1 and S1, which results in less energetic loss through TTA, and thus high theoretical efficiency. More broadly, we have introduced a new class of upconverting annihilators which can be tuned via chemical functionalization to emit in the violet-UV regime.

This work echoes a previous study[263] in highlighting the utility of computer simulations in the screening of TTA upconversion emitters for the rational design of upconverting materials. Yet crucially, the TD-DFT and DLPNO-CCSD(T) methods examined in this study would have led us to overlook the Ph-BTD/PtOEP pair, underscoring the importance of predictive accuracy on the level of around a tenth of an eV or less. In contrast to the other computational methods investigated here and, e.g, in Ref. 263, AFQMC is capable of providing this resolution for triplet energies, and thus will be a powerful tool for the design of upconverting annihilators.

3.6 Appendices

3.6.1 Experimental Methods

All starting materials were obtained from commercial chemical sources, including Fisher Scientific, TCI Chemical, and Strem Chemicals. BTD (ACROS Organics), ZnTPP (Fisher Scientific), and PtOEP (Sigma-Aldrich) were purchased and used without further purification. Materials Ph-BTD[264], MeO-BTD[265,266], and CN-BTD[267] were synthesized according to previously reported procedures.

NMR spectra were collected on a Bruker 500 MHz spectrometer at ambient temperature. UV-Vis absorption spectra were collected by a Technologies Cary 60 UV-Vis Spectrophotometer. Steady-state photoluminescence spectra were collected by an Ocean Optics QEPro Spectrometer.

TTA Upconversion Analysis

Solutions were made with 1x10^{-}5 M sensitizer and 1x10^{-}3 M annihilator in degassed anhydrous toluene. Solutions for each sensitizer-annihilator pair were made in a nitrogen glovebox, sealed, and removed from the glovebox for upconversion photoluminescence study.

Normalized emission and absorption spectra of sensitizers PtOEP and ZnTPP are seen in figures 3.11 and 3.12, where dashed lines denote absorption and solid lines denote emission upon excitation at 365 nm.

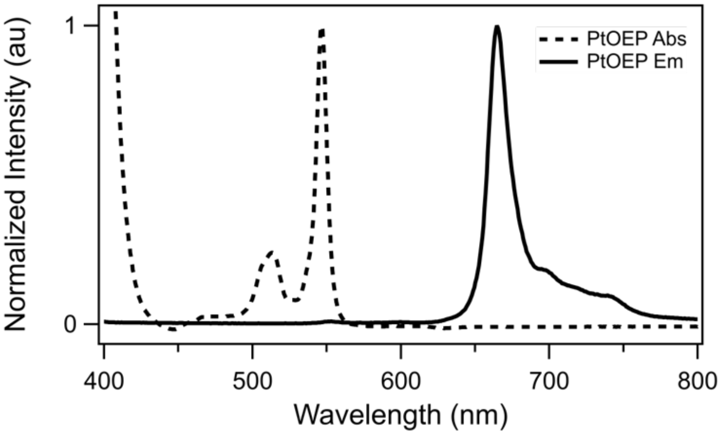


Figure 3.11: Absorption (dashed) and photoluminescence (solid) of PtOEP in toluene.

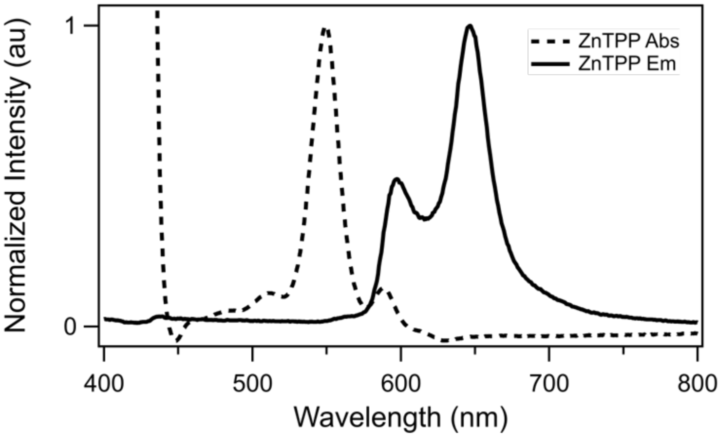


Figure 3.12: Absorption (dashed) and photoluminescence (solid) of ZnTPP in toluene.

Normalized absorption spectra of sensitizers PtOEP and ZnTPP and emission of solution mixture of the sensitizers paired with BTD are seen in figures 3.13 and 3.14, where dashed lines denote the absorption of the sensitizer and solid lines denote the emission upon excitation at 532 nm.

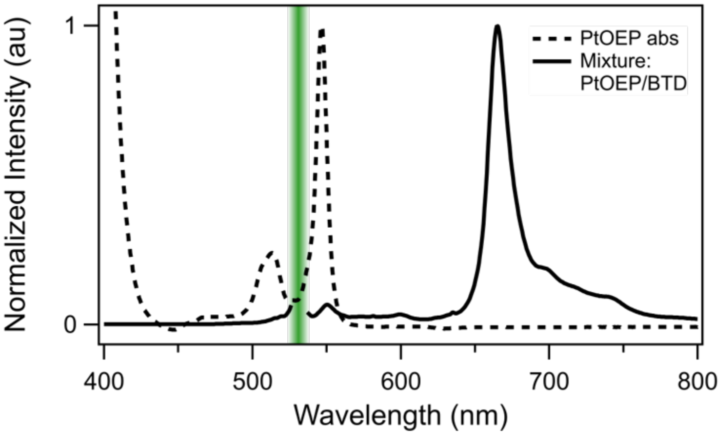


Figure 3.13: Absorption (dashed) spectrum of PtOEP, and photoluminescence (solid) of the BTD/PtOEP pair in toluene with excitation at 532 nm (green line). The emission of the mixture matches that of PtOEP (Figure 3.11), supporting that there is no TET from PtOEP to BTD in solution.

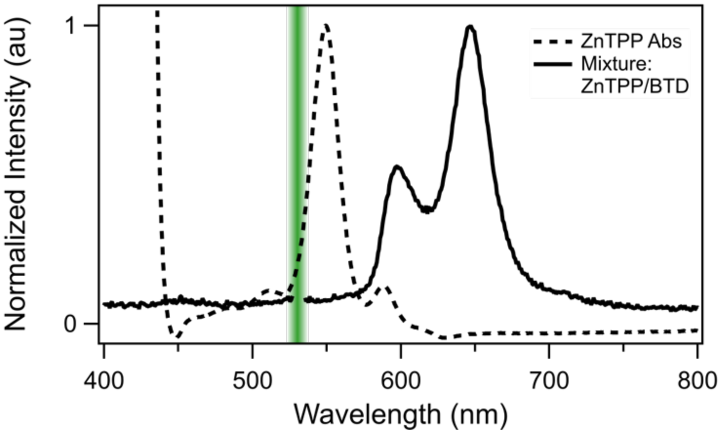


Figure 3.14: Absorption (dashed) spectrum of ZnTPP, and photoluminescence (solid) of the BTD/ZnTPP pair in toluene with excitation at 532 nm (green line). The emission of the mixture matches that of ZnTPP, (Figure 3.12), supporting that there is no TET from ZnTPP to BTD in solution.

Phosphorescence Measurements

Time-resolved phosphorescence measurements for BTD and CN-BTD were recorded on a Fluorolog-3P fluorometer (HORIBA Jobin Yvon). Sample solutions in 3-mm quartz tubes (inner diameter) were frozen in a quartz liquid nitrogen Dewar and excited with a pulsed xenon lamp. Luminescence signal detection was delayed 10 ms (for BTD) or 1 ms (for CN-BTD) after the light pulse and collected for 20 ms (BTD) or 1 ms (CN-BTD). Phosphorescence lifetimes (T_p) at 77 K for BTD and CN-BTD were measured by multichannel scaling on an OB920 spectrometer (Edinburgh Analytical Instruments) in conjunction with a pulsed xenon lamp. Time-resolved phosphorescence measurements of MeO-BTD and Ph-BTD were measured in frozen methylcyclohexane/iodomethane (2:1, v/v) in 3 mm quartz tubes (inner diameter) at 77 K in a optical liquid N_2 quartz dewar. Iodomethane was added to increase intersystem crossing into the triplet state. The frozen samples inside the quartz dewar were excited with a pulsed Spectra Physics GCR-150-30 Nd:YAG laser (355 nm, ca. 1 mJ per pulse, 5 ns pulse length). The time-resolved phosphorescence spectra at 100 μs (for MeO-BTD) or 20 μs (for Ph-BTD) after pulsed excitation and a gate width of 500 μs (MeO-BTD) or 1 ms (Ph-BTD) were recorded on an Acton Spectrograph (SpectraPro-2150) in conjunction with an intensified CCD detector (PI-MAX from Princeton Instruments) with fiber optics attachment. The triplet energies (E_T) were determined from the highest energy peaks of the phosphorescence spectra of BTD, CN-BTD and MeO-BTD. Because the phosphorescence spectrum of Ph-BTD did not show resolved vibrational peaks, the triplet energy was estimated to have a value between 1.85 eV (phosphorescence maximum) and 1.99 eV (wavelength of $\frac{1}{2}$ intensity of the phosphorescence peak).

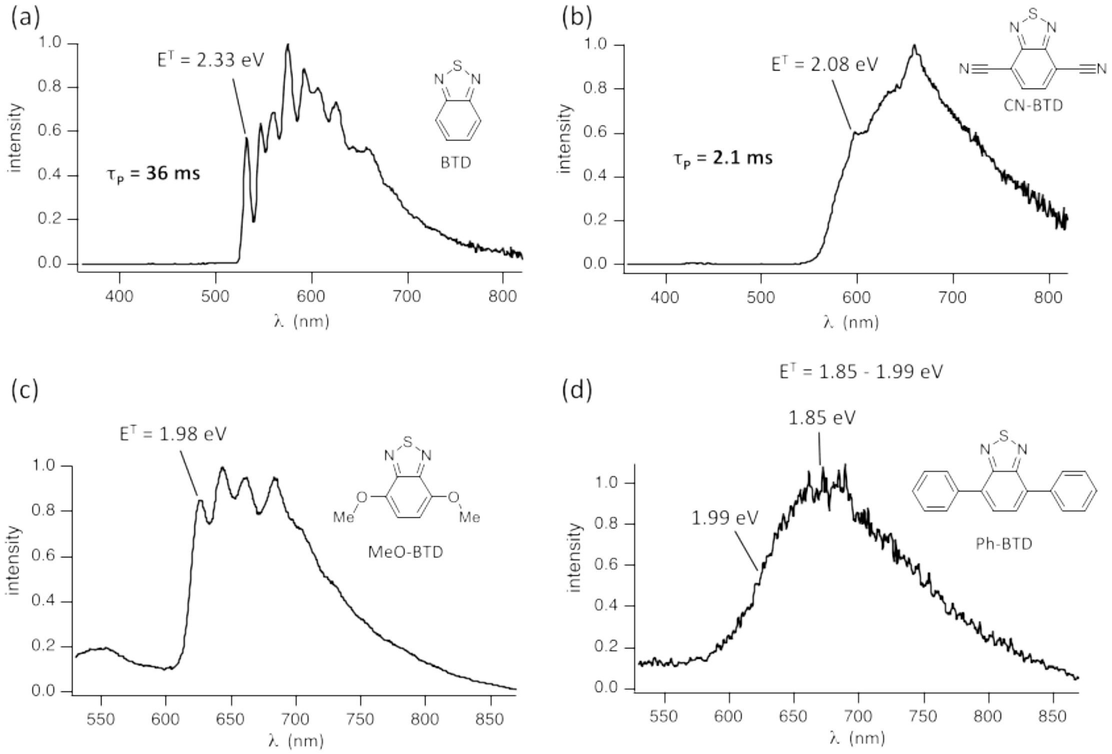


Figure 3.15: Time-resolved phosphorescence spectra of BTD derivatives in methylcyclohexane (a, b) or methylcyclohexane / iodomethane mixture (2:1, v/v) (c, d) at 77 K. λex = 340 nm (a, b) or λex = 355 nm (c, d). Detection window after light pulse: 10 - 30 ms (a), 1 - 6 ms (b), 100 - 600 μs (c), 20 - 1020 μs (d).

3.6.2 Additional Calculation Details

All AFQMC calculations in this work use the same procedures and parameters (e.g. Cholesky decomposed two-electron integrals, time-step size, precision for floating-point operations, orthonormalization interval, etc.) detailed in Ref. 72, with walker populations of around 1200.

We additionally explored the effects of extrapolating our triplet energies, as computed with ph-AFQMC in the cc-pVTZ (TZ) basis, to the complete basis set (CBS) limit using DLPNO-CCSD(T) values in TZ and QZ (see Ref. 47 for details of this protocol). For the subset of molecules in Section 3.1.1 we find that with DLPNO-CCSD(T) the difference between the TZ result and our extrapolated estimate of the CBS limit is never larger than 0.06 eV. Given that our QMC error bars are larger than this, further extrapolation of the AFQMC/U results appears to be unnecessary for screening, and the following results reflect use of the TZ basis. This is consistent with previous

CC[268] and DFT[269] studies showing TZ basis sets to be sufficient for screening T1 excitation energies of organic chromophores. This was not the case for the BTD series, and we therefore make use of CBS numbers in the main text. Species with no CBS values indicate that we were unable to converge QZ DLPNO-CCSD(T) calculations in <1 week computational wall time.

Triplet energies

Below are our calculated adiabatic T1 energies for all systems involved.

Table 3.6: A comparison in eV of AFQMC/U, DLPNO-CCSD(T), KS-DFT, and TD-DFT results for the T1 energies of anthracene derivatives. Parentheses denote statistical error; the value to the right of "/" denotes the value extrapolated to the basis set limit.

	KS-B3LYP	B3LYP	CAM-B3LYP	M062X	DLPNO-CCSD(T)(TZ/CBS)	AFQMC(TZ/CBS)
DCA	1.50	1.757	1.825	1.956	1.76/1.80	1.74(7)/1.80(7)
DPA	1.66	1.940	2.023	2.123	2.02/2.04	1.89(11)/1.93(10)
CF3	1.66	1.875	1.943	2.075	2.02/2.03	2.09(11)/2.14(10)
CN-updn-Me	1.28	1.688	1.766	1.895	1.74/1.76	1.70(12)/1.74(11)
CN-diag-Me	1.42	1.790	1.865	1.996	1.79/1.82	1.80(11)/1.86(10)
DMA	1.54	1.951	2.025	2.148	1.91/1.92	1.81(7)/1.84(7)
OMe	1.60	2.017	2.093	2.207	1.94/1.94	1.87(10)/1.90(9)
Ac	1.41	1.688	1.749	1.537	1.68	1.68(8)
Ac-CN1	1.38	1.649	1.718	1.502	1.67	1.56(6)
Ac-CN2E	1.36	1.611	1.686	1.464	1.65	1.64(7)
Ac-CN2Z	1.37	1.622	1.697	1.480	1.67	1.67(7)

Species	KS-B3LYP	B3LYP	CAM-B3LYP	M062X	DLPNO-CCSD(T)	AFQMC
BTD	3.11	2.594	2.647	2.828	2.87/2.67	2.85(6)/2.66(7)
CN-BTD	1.99	2.299	2.358	2.528	2.28/2.37	2.33(7)/2.43(6)
MeO-BTD	1.89	2.148	2.350	2.509	2.17/2.22	2.18(7)/2.20(6)
Ph-BTD	1.78	2.154	2.289	2.437	2.07/2.13	1.77(6)/1.81(6)
BSeD	2.81	2.339	2.381	2.558	2.60/2.84	2.65(4)/2.89(3)
CN-BSeD	1.83	2.143	2.188	2.357	2.13/2.36	2.25(7)/2.50(7)
MeO-BSeD	1.66	1.913	2.098	2.246	1.98	1.87(5)
Ph-BSeD	1.62	1.961	3.352	2.233	1.92/1.96	2.11(10)/2.17(9)

Table 3.7: A comparison in eV of AFQMC/CAS, DLPNO-CCSD(T), KS-DFT, and TD-DFT results for T1 of a set of substituted BTD and BSeD compounds. Parentheses denote statistical error; the value to the right of "/" denotes the value extrapolated to the basis set limit.

Species	KS-B3LYP	B3LYP	CAM-B3LYP	M062X	DLPNO-CCSD(T)(TZ/CBS)	AFQMC(TZ/CBS)
MAD	0.249	0.171	0.257	0.405	0.216	0.170(73)

Table 3.8: Mean Absolute Deviations for all experimental values (DCA, DPA, BTD, CN-BTD, MeO-BTD, Ph-BTD

Solvent and vibrational corrections

For a small subset of anthracene derivatives, the effect of vibrational zero-point energies and solvent corrections were evaluated. Solvent corrections were included using an implicit solvent model dielectric (CPCM) for toluene at the DFT level, and the corrected ST gaps are given in Table 3.9. As neither ZPE nor solvent correction resulted in an energetic shift greater than the statistical error of AFQMC, we neglect vibrational and solvent effects in the main manuscript, and expect (and find) that while the reported energies are gas phase, they accurately reflect solvated energetics in a nonpolar solvent such as toluene.

The vibrational corrections were estimated from numerical vibrational frequency calculations with an increment of 0.001 Bohr. Zero-point energies and thermal vibrational corrections are scaled by a factor obtained from the NIST computational chemistry database (0.965 for B3LYP/cc-pVTZ)[270] and added to the resulting energies. Corrections due to solvation effects were obtained from geometries and vibrational frequencies computed within the conductor-like polarizable continuum model (CPCM) to estimate the effects of toluene solvent. The DLPNO-CCSD(T) and AFQMC/U values are obtained by first calculating the S0 - T1 energy difference in the gas-phase; energetic corrections to this gap due to ZPEs and thermal occupation of vibrational states, computed in the gas-phase, are scaled and then added. Finally, a solvent correction is obtained via DFT as the difference between the vibrationally corrected ST gaps in toluene, as represented by the CPCM model, and in gas-phase.

Table 3.9: Dielectric corrections to the gas-phase energy to account for solvent (toluene) effects, within the CPCM formalism. All corrections were found at the B3LYP/TZ level, using ORCA. Note that tthe solvation correction is minimal in this low-dielectric solvation model, as expected, aside from the MeO-substituted diazoles, which exhibit the strongest charge-transfer due to the electron donating character of MeO; the corrections were -0.15 and -0.13 eV for MeO-BTD and MeO-BSeD, respectively.

Species	Solvation Correction
DMA	-0.036
DCA	0.063
OMe	0.012
CN-1,5-Me	-0.066
CN-2,6-Me	-0.008
CF3	0.028
DPA	-0.005
Ac	0.000
Ac-CN1	-0.006
Ac-CN2E	-0.013
Ac-CN2Z	0.005
BTD	-0.043
CN-BTD	0.032
MeO-BTD	-0.148
Ph-BTD	-0.003
BSeD	-0.046
CN-BSeD	0.042
MeO-BSeD	-0.128
Ph-BSeD	0.000

Table 3.10: Triplet energies including both vibrational and solvent (toluene) corrections within the CPCM formalism.

Species	KS-B3LYP	DLPNO-CCSD(T)	AFQMC/U
DMA	1.54	1.77	1.69
DCA	1.50	1.75	1.75
OMe	1.60	1.85	1.81
CN-1,5-Me	1.28	1.58	1.56
CN-2,6-Me	1.42	1.70	1.73
CF3	1.66	1.95	2.05
DPA	1.66	1.91	1.82

Dimerization Energies

Due to prior computational evidence for strong chalcogen bonding in anti-square thiadiazole dimers[259], we evaluated the electronic bonding energy for each pair. The bonding energies, evaluated at the KS-DFT ωB97X-V/cc-pVTZ level of theory, are shown in Table 3.11, in addition to DLPNO-CCSD(T)/cc-pVTZ estimations of the triplet energy for each dimer. Each species electronically favors the dimer, with CN-BTD and BTD having the strongest and second strongest bonding energies, respectively. Moreover, the computational triplet energies agree with the experimental phosphorescence results, suggesting that the low temperature phosphorescence measurements may reflect dimerized BTD and CN-BTD species. To further evaluate the likelihood of dimerization, we calculated the dimerization free energies at a range of temperatures (Table 3.12). For a concentrated solution at low temperatures, the dimer is favored in both BTD and CN-BTD at ratios of 1000:1 and 1000,000,000:1, respectively. However, the likelihood of dimerization at low concentrations such as those used in the phosphorescence measurements ($\simeq \mu$M) is expected to be low.

	BTD	CN-BTD	MeO-BTD	Ph-BTD
Electronic Bond energy	0.179	0.279	0.125	0.124
Experimental T1 energy	2.33	2.07	1.98	1.85(1.99)
Dimer T1 energy (DLPNO-CCSD(T)/TZ)	2.42	2.28	2.22	2.10
Monomer gap (AFQMC/CBS)	2.62(6)	2.46(7)	2.09(7)	1.81(6)

Table 3.11: Electronic bond energies (eV) for the anti-square dimer (at the wB97X-V/cc-pVTZ level of theory), along with a comparison of triplet energies (eV) for the dimer(calculated by DLPNO-CCSD(T)) and monomer (AFQMC) vs experiment. All species exhibit chalcogen bonding, with CN-BTD and BTD exhibiting the highest and second highest bonding strengths, respectively. The phosphorescence for Ph-BTD lacked the fine structure necessary to specify the 0-0 transition energy, and so we report two estimates, one based on the phosphorescence maximum, and one (in parentheses) determined from the 1/2 intensity of the phosphorescence maximum. Triplet energies for the dimers, calculated with DLPNO-CCSD(T), agree with experiment for BTD and CN-BTD, within 0.04 and 0.1 eV, respectively.

Temperature (K)	BTD	CN-BTD	MeO-BTD	Ph-BTD
77	-0.040	-0.139	0.030	0.132
100	0.002	-0.096	0.076	0.199
150	0.097	-0.001	0.176	0.355
200	0.192	0.096	0.277	0.520
250	0.289	0.194	0.380	0.693
300	0.383	0.325	0.515	0.900

Table 3.12: Dimerization free energies evaluated using KS-DFT at the B3LYP/cc-pVTZ level. Note that a negative free energy favors the formation of a dimer; all species are monomers at room temperature, but BTD and CN-BTD transition to a favored dimer state at low temperatures. Inspection of the geometries reveal in-plane dimerization for all but Ph-BTD, which is sterically limited to out of plane, likely resulting in the higher observed dimerization free energies.

Species	S1 (TD-DFT)
DPA	3.136
DMA	3.101
OMe	3.059
CF3	3.033
CN-2,6-Me	2.926
DCA	2.922
CN-1,5-Me	2.790
Ac	2.88
Ac-CN1	2.82
Ac-CN2E	2.78
Ac-CN2Z	2.77

Species	S1 (TD-DFT)
BTD	3.86
CN-BTD	3.64
MeO-BTD	3.05
Ph-BTD	3.04
BSeD	3.57
CN-BSeD	3.41
MeO-BSeD	2.78
Ph-BSeD	2.85

Table 3.13: CAM-B3LYP TD-DFT results for S1 energies in eV.

Excited State Tuning

Here we attempt to rationalize shifts in the S0-S1 transition energies (Table 3.13) between functionalized chromophores based on intuitive concepts within chemistry. We find that two factors, that of π-system extension and a combination of electron-withdrawing and electron-rich substituents, can be used in tandem to lower the HOMO-LUMO gap.

π-extension

Extending the π system should result in lowering T1 and S1, thus substantially lowering the 2T1-S1 gap. In the series of anthracenes, it can be seen that the amount of S1 lowering from the parent system can be directly correlated to the extent of π-extension within the functional groups. Interestingly, Me groups contribute significantly to π-extension through hyper-conjugation, as seen in the HOMO-LUMO plots of DMA:

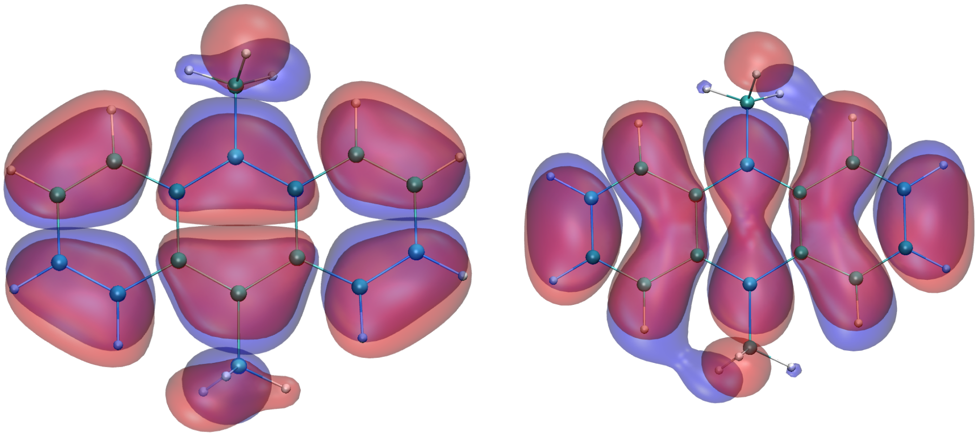


Figure 3.16: HOMO(left) and LUMO (right) orbital plots for DMA. Note the extension of the π-system into the Methyl groups.

Whereas these particle in a box arguments are helpful in predicting S1, T1 is much more dependent on the exact nature of the functional group. We are currently unable to resolve any correlations that could be used as a design principle for T1 in the anthracenes, and in some cases T1 actually follows opposite trends from S1 (i.e. DMA → OMe → CF3). This, however, underlines the importance of accurate computational methods such as AFQMC.

Donor-Acceptor Character

π-system extension can have neutral, donating, or withdrawing character, and excitations can involve varying degrees and directions of charge-transfer, which can in principle be harnessed to lower the ST gap via judicious choice of solvent polarity[271]. This may also be a factor in increasing the energy of certain excitation states by pulling electrons away from highly electronegative groups within certain geometries. In a similar fashion, it is well known that conjugation of an donor-acceptor pair can lead to a large decrease in the band gap.[272] We can therefore use the relative electron withdrawing/donating characters of the parent compound and functional groups to predict the extent of S1 lowering. The BTD series illustrates this well. BTD is a highly electron-poor

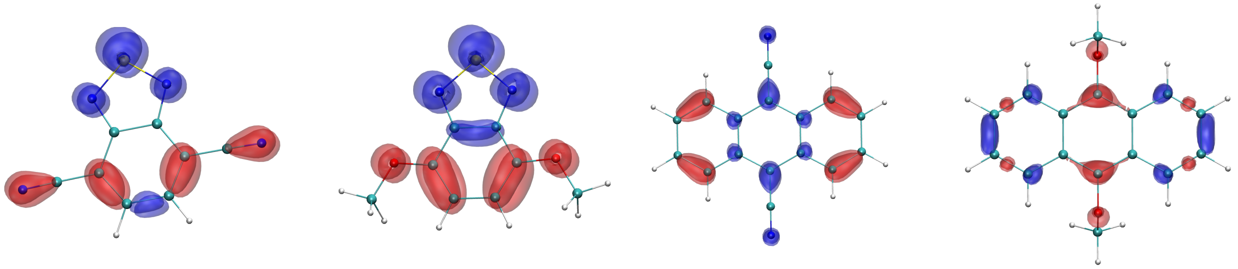


Figure 3.17: Difference-density plots for the S0-S1 transition for CN and MeO substituted BTD and anthracene. A red surface denotes electron loss upon excitation, while the blue is electron density gained. Note the significant charge transfer in both BTD cases from the functional groups to the sulfur ring, regardless of electron withdrawing (CN) or donating (OMe) character; withdrawal of electrons from CN likely raises the excited state energy in comparison to OMe, even with similar HOMO-LUMO π-extension. In comparison, DCA and OMe-anthracene show the opposite trend, while the difference density plot shows electron withdrawing to the CN and electron donation from the OMe.

system used frequently in donor acceptor paradigms, and so electron donating groups such as OMe lower the gap significantly more than electron withdrawing groups such as CN, even though OMe is arguably less effective at extending the π-system than CN (Figure 3.17).

Singlet Excited states using AFQMC

Ideally, one would obtain both S1 and T1 from a consistent level of theory. In the case of calculating S1 for a molecule with a singlet ground state, fixing the spin cannot be used to orthogonalize the trial wave-function against the ground state, as it can for the triplet. In the case of anthracene, we fix the symmetry of the wavefunction to be B2u, that of the first bright excited singlet state. For a general non-symmetric molecule, one can use state-averaging techniques to obtain orthogonal CASSCF trial wavefunctions for an arbitrary number of excited states. We include all π orbitals in the active space, and use the same random number seeds for S0, S1, and T1 calculations to accelerate the convergence of the energy gaps, as in correlated sampling (procedure outlined in Ref. 56). Since correlated sampling attains maximum efficiency when the same geometries are used, we compute the vertical excitation energies, and thus compare to experiments corrected with a geometry reorganization energy calculated with DLPNO-CCSD(T) (from which, given our results for the anthracene derivatives, we expect reasonable accuracy). Indeed, in Table 3.14 we report vertical excitation energies within 0.05 eV accuracy versus experiment.

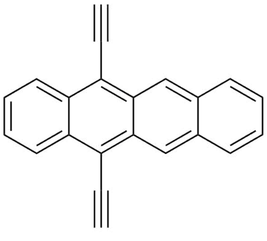
Table 3.14: AFQMC results for both S1 and T1 of Anthracene using Correlated Sampling

State	AFQMC/CAS	Experiment[247]
S1	3.65(4)	3.60
T1	2.27(1)	2.23[a]

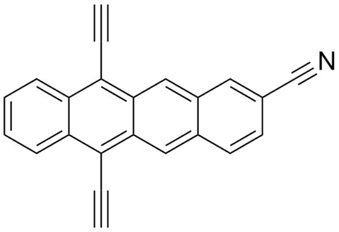
[a] Corrected from experimental adiabatic energy using DLPNO-CCSD(T) reorganization energy

Functionalized Tetracene structures

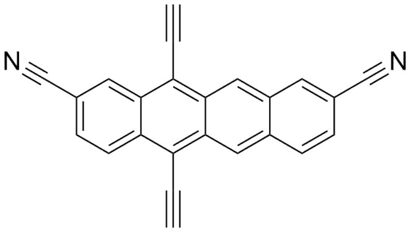
We have used a set of structures based on functionalizing tetracene with cyano groups in order to validate the use of CAM-B3LYP TDDFT to obtain S1 energies against experiment (see Table 3.13). The structures for these compounds can be seen below:



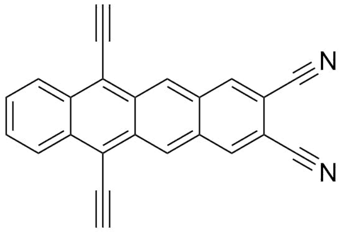
CN0



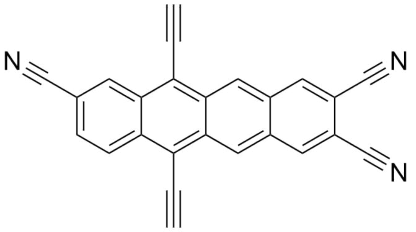
CN1



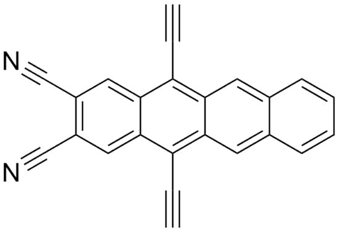
CN2E_cis



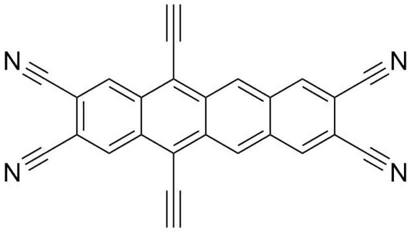
CN2T



CN3



CN2H



CN4

Chapter 4: A Localized-Orbital Energy Evaluation for Auxiliary-Field Quantum Monte Carlo

Reprinted with permission from J. L. Weber *et al.*, *J. Chem. Theory Comput.*, vol. 18, no. 6, 2022. Copyright 2022 American Chemical Society.

4.1 Introduction

Solving the electronic structure of large, strongly correlated molecules at zero temperature represents a quintessential challenge of quantum chemistry. Quantum Monte Carlo methods, in particular projector-based Monte Carlo techniques such as Auxiliary-Field Quantum Monte Carlo (AFQMC), have emerged as promising approaches capable of producing benchmark results for small, strongly correlated systems.[48] This method relies on solving the Schrodinger equation in imaginary time τ, projecting out the exact ground state $|\phi_0\rangle$ under the requirement that the initial state $|\phi_i\rangle$ has a non-zero overlap, $\langle\phi_i|\phi_0\rangle \neq 0$

$$|\phi_0\rangle \propto \lim_{\tau\to\infty} e^{-\tau\hat{H}}|\phi_i\rangle. \tag{4.1}$$

AFQMC does so by reformulating the many body propagator via a Hubbard-Stratonovich transformation to an integral of a set of one body propagators coupled to auxiliary fields. This large multi-dimensional integral is estimated by Monte Carlo sampling a random walk of non-orthogonal Slater determinants, called walkers. Each walker can be propagated independently, aside from occasional communication to maintain efficient sampling populations (see later discussion on population control), and so this algorithm is nearly embarrassingly parallel. This allows one to make effective use of graphical processing units (GPUs), in favorable cases resulting in a ≃ 200× reduction in the prefactor. Aside from a few model systems, however, one encounters a fermionic sign problem (or phase problem in the case of AFQMC). This leads to an exponential decrease in the signal to noise ratio, and thus an exponential increase in the required computational effort to obtain statistically meaningful results as the system grows. Algorithms for which the phase problem is uncontrolled, i.e. free projection AFQMC (fp-AFQMC), are formally exact, and the use of systematically better trial wavefunctions can reduce variance and allow convergence of exact results for small systems.[273] In order to render AFQMC polynomially scaling, the phase problem can be avoided via the phaseless approximation (ph-AFQMC), where we remove any phase accumulated in the weight by imposing a boundary condition based on the trial wavefunction.[274] This removes the phase problem and allows extended simulations of realistic materials, at the expense of a systematic bias with respect to the trial wavefunction.

The success of ph-AFQMC for *ab initio* systems is in large part dependent upon the ability to systematically increase the quality of the results by increasing the quality of the trial wavefunction. Thus a fundamental question in the field centers on the optimal construction of trial functions, and there have been many studies exploring the space of quantum chemistry methods for efficient and accurate trial selection. In some cases, mean-field trial wavefunctions such as Hartree-Fock (HF) or Kohn-Sham Density Functional Theory (KS-DFT) are sufficient.[3,72,113] In many cases, however, multi-determinant trials are necessary, particularly in transition metal containing complexes, with Configuration Interaction (CI) expansions including Complete Active Space Self-Consistent Field (CASSCF),[1–3,47,243,244,275–277] Non-Orthogonal CI (NOCI),[244] and Semi-stochastic Heat-bath CI (SHCI) having been used for this purpose.[273,278] Interestingly, even in some transition metal systems which are not expected to exhibit strong static correlation effects, multi-determinant trials have been necessary for ph-AFQMC to converge to within experimental results.[2] However, use of a multideterminantal trial function leads to increased costs in both generating the trial and using it in AFQMC propagation and energy evaluation. Addressing these issues is a key challenge inherent in the development of a scalable AFQMC approach with benchmark accuracy across a wide range of chemical problems.

For transition metal containing molecules, we have found that it is often necessary to use a multi-determinantal trial such as one provided by a CASSCF wavefunction (AFQMC/CAS) with a limited number of determinants ($N_{det} \simeq 300 - 1000$), in order to converge the result to chemical accuracy or near chemical accuracy.[1,2,47] However, the use of AFQMC/CAS has been limited to small, single transition-metal systems (< 1000 basis functions, 200 electrons), in part due to the enhanced computational cost and scaling of the local energy evaluation. Optimization of this step, therefore, may allow for the extension of AFQMC to larger transition metal systems of importance in biology, catalysis, and materials science.

For single determinant trials, the propagation of a walker within ph-AFQMC algorithms exhibits cubic scaling, whereas the energy evaluation exhibits quartic scaling. The most expensive steps in the propagation are the formation of the propagator, as well as the formation of the a shift in the auxiliary fields necessary for efficient importance sampling,[54] both of which scale as XM^2, or asymptotically $O(M^3)$, where M is the full basis dimension and X is defined as the number of auxiliary fields, $X \simeq 4 - 10M$ for Cholesky decomposition thresholds between 10^{-4} and 10^{-5} Ha.[279] In the most straightforward implementation,[111] the energy evaluation has a cost of $O(N^2M^2)$, where N is the number of electrons in the molecule. There have been numerous attempts to decrease the scaling of the energy evaluation for single determinant AFQMC. Many take advantage of the low-rank nature of the ERIs beyond that of the literature standard, namely the modified Cholesky decomposition. For example, Motta *et al.* formed a low-rank compression of each individual Cholesky matrix for single determinant trials.[280] Additionally, Tensor Hyper-Contraction (THC) has emerged as a cubically scaling method with only quadratically scaling memory usage, albeit with a large prefactor that limits the utility of the approach to systems with $M > 2000$ to 4000.[112,236] In addition to these low rank algorithms, Lee and Reichman applied a stochastic resolution of the identity (sRI) to reduce the scaling of the sum over Cholesky matrices, resulting in a cubic scaling algorithm with minimal overhead, although the accuracy has not been demonstrated for correlated systems.[281] It is also possible to take advantage of the sparsity of the Hamiltonian, where in a specified representation only elements above a certain threshold are explicitly stored. However, linear operations on sparse matrices are fairly inefficient, particularly on graphical processing units (GPUs), and thus doing so is only favorable in cases with extreme sparsity where nearly 90% of the elements are removed, which is in our experience typically not the case for molecular systems.

When extending the trial to include multiple determinants, there are many choices of algorithms which take advantage of the excitation structure of CI expansions. The simplest involves using the Sherman Morrison Woodbury (SMW) algorithm to update the overlap between the walker and trial determinants.[57] This approach significantly reduces the computational cost of multideterminantal propagation, but results in a scaling of $O(N_{det}N^2M^2)$ for the energy evaluation, where N_{det} is the number of determinants in the trial, resulting in a quickly intractable code for large determinant expansions. Recently, Mahajan *et al.* have described an algorithm for the local energy evaluation of AFQMC with multi-determinant trials, which effectively reduces the scaling dependence on the number of determinants to $O(NM^3+N_{det}M)$.[273,278,282] This algorithm enables the use of extremely accurate trial wavefunctions based on large SHCI expansions, which in turn reduces the variance of free-projection AFQMC enough to treat significantly larger systems than previously possible. While this opens the door to the usage of larger determinant expansions, for large systems the energy evaluation is still quartic, and thus remains the effective bottleneck. It is possible to modify THC to use within this framework, but this is again subject to the associated large prefactor and is likely to be useful only for very large systems.

Here we report an alternative (and in some cases complementary) method, directly applicable to both single determinant and multideterminant CI expansion trials, which takes advantage of the low rank of the electron repulsion integrals (ERIs) when working in a localized orbital basis. It does so by compressing the block of the “half-rotated" ERI tensor corresponding to the interaction between each electron pair $[ij]$ using singular value decomposition (SVD), which is systematically controllable by an energetic threshold in addition to being compatible with dense linear algebra routines. This results in a reduced memory scaling from quartic to cubic for this tensor, with a concomitant reduction of scaling for the energy evaluation for a single determinant. We implement and outline the use of these localized ERIs within the SMW algorithm for multideterminant CAS trials, resulting in an energy evaluation algorithm that scales as $\langle M_{SVD} \rangle (N^2 M + N_{det} N^2)$, where $\langle M_{SVD} \rangle$ represents the average rank of the compressed [ij] block, and we omit other values which are constant with system size, such as the maximum number of excitations from the reference for the set of determinants, ϵ. This low rank localized structure is additionally compatible with the modified generalized Wick's theorem approach of Mahajan *et al.*, and we outline one such algorithm in the SI, which results in a theoretical scaling of $\langle M_{SVD} \rangle N^2 M A + N_{det}$, where A is the size of the active space. We move on to show the accuracy of Localized Orbital AFQMC (LO-AFQMC) in population control ph-AFQMC to within the statistical error of ph-AFQMC ($\simeq$<1 mHa) for a variety of representative molecules of nontrivial size and complexity, including polyacenes, metallocenes, and a set of benchmark platonic hydrocarbon cages.

We emphasize that although we focus on the application of this localized orbital compression to a particular set of algorithms within ph-AFQMC, much of this work is applicable to other orbital based Quantum Monte Carlo methods. We expect that the relative simplicity of the approach described here will enable these extensions. Due to the many advantages of GPU parallelization, we make efforts in this work to tailor our algorithm towards GPU architectures by maintaining an emphasis on both the reduction of memory requirements and the ability to perform operations efficiently in parallel using standard CUDA libraries.[283]

The manuscript is organized as follows: In Sec. 2 we provide a brief overview of AFQMC (for a more in depth discussion see Ref 279), along with relevant multi-determinant algorithms for the local energy evaluation. In Sec. 3 we present the localized-orbital based algorithm, including some discussion of optimizing an implementation on GPU architectures. In Sec. 4 we then present the accuracy, memory usage, and timings for a series of test molecules. Lastly, we discuss the impact of such results, and conclude with our current outlook for AFQMC calculations in medium to large systems with strong correlation.

4.2 Theory

4.2.1 Overview of ph-AFQMC

Here we give a brief introduction to the framework of AFQMC; for a more in depth review of the method, we suggest some recent reviews.[73,111] In AFQMC, initial states are propagated in imaginary time τ according to

$$|\phi(\tau+\Delta\tau)\rangle = e^{-\Delta\tau\hat{H}}|\phi(\tau)\rangle = \int d\mathbf{x}P(\mathbf{x})\hat{B}(\mathbf{x})|\phi(\tau)\rangle \approx \sum_{w}\hat{B}(\mathbf{x}_w)|\phi_w(\tau)\rangle, \tag{4.2}$$

where $\hat{H}$ denotes the electronic Hamiltonian, $\hat{H} = \hat{H}_1+\hat{H}_2 = \sum_{pq} h_{pq}c_p^\dagger c_q+\frac{1}{2}\sum_{pqrs} V_{pqrs}c_p^\dagger c_q^\dagger c_s c_r$, $|\phi_w\rangle$ denotes the Slater determinant associated with walker w, and $\hat{B}|\phi_w(\tau)\rangle = |\phi_w(\tau+\Delta\tau)\rangle$ is another Slater determinant with rotated orbitals as given by the Thouless theorem.[284] V_{pqrs} are the two electron integrals referred to as $(pr|qs)$ and $\langle pq|rs\rangle$ in chemists and physicists notation, respectively. For long imaginary times, computing observables using this Monte Carlo representation of the wavefunction will recover ground-state properties as long as the initial wavefunction has a non-zero overlap with the ground state. Using a symmetric Suzuki-Trotter decomposition, we separate the one- and two-body terms in $\hat{H}$ with an error quadratic in the imaginary time τ,

$$e^{-\Delta\tau(\hat{H}_1+\hat{H}_2)} \simeq e^{-\frac{\Delta\tau\hat{H}_1}{2}}e^{-\frac{\Delta\tau\hat{H}_2}{2}}e^{-\frac{\Delta\tau\hat{H}_1}{2}} + \mathcal{O}(\Delta\tau^3). \tag{4.3}$$

In practice, we mitigate the Trotter error by restricting our calculations to a small timestep, $\Delta\tau$ (0.005 Ha^{-1} in this work). If we write the electronic two-body operator as a sum of one-body operators squared, $V_{pqrs} = \sum_\alpha L_{pr,\alpha}L_{qs,\alpha}$, which can be accomplished exactly via diagonalization or approximately via a density fitting or a modified Cholesky decomposition,[146] we can then use the Hubbard-Stratonovich identity to convert the two-body operators into a multi-dimensional integral over a set of fluctuating "auxiliary-fields" x_α,

$$e^{-\frac{\Delta\tau}{2}(\sum_\alpha L_\alpha^2)} = \prod_\alpha \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_\alpha^2}{2}} e^{\sqrt{\Delta\tau} x_\alpha L_\alpha} dx_\alpha + O(\Delta\tau^2). \tag{4.4}$$

It is this multi-dimensional integral on which we perform Monte Carlo sampling

$$|\phi(\tau+\Delta\tau)\rangle = \prod_\alpha \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{x_\alpha^2}{2}} e^{\sqrt{\Delta\tau} x_\alpha L_\alpha} dx_\alpha |\phi(\tau)\rangle = \int d\mathbf{x} P(\mathbf{x}) \hat{B}(\mathbf{x}) |\phi(\tau)\rangle, \tag{4.5}$$

where $\mathbf{x}$ is the vector of auxiliary fields. This Monte Carlo simulation can be reformulated as a branching, open-ended ensemble of random walkers w over the manifold of Slater determinants, each represented by a single Slater determinant $\phi_{\tau,w}$ and corresponding weight $W_{\tau,w}$ and overlap with the trial $\langle\Phi_T|\Phi_{\tau,w}\rangle$. As each walker is propagated forward by $\hat{B}(\mathbf{x}_{\tau,w})$, with the space of auxiliary fields $\mathbf{x}$ being sampled from the Gaussian probability defined in Eq. 5.4, the weights are updated according to the ratio of the new overlap with the trial to the old overlap

$$|\Phi_{\tau+\Delta\tau,w}\rangle = \hat{B}(\mathbf{x}_{\tau,w})|\Phi_{\tau,w}\rangle, \tag{4.6}$$

$$W_{\tau+\Delta\tau,w} e^{i\theta_{\tau+\Delta\tau,w}} = \frac{\langle\Phi_T|\Phi_{\tau+\Delta\tau,w}\rangle}{\langle\Phi_T|\Phi_{\tau,w}\rangle} W_{\tau,w} e^{i\theta_{\tau,w}}. \tag{4.7}$$

All theory thus far is formally exact; however, the fermionic phase problem leads to an exponential decrease in signal-to-noise ratio as the walkers are propagated, resulting in an exponentially growing population of walkers (and thus computational time) necessary to achieve a given statistical error. In order to mitigate this noise, we perform importance sampling by shifting the auxiliary fields to favor sampling in regions with high(er) overlap with the trial wavefunction.[54,285] This is complemented by the phaseless constraint, in which we multiply each walker's weight by a factor corresponding to projecting the accumulated phase with respect to the trial wavefunction back onto the real axis, namely we multiply by max(0, cos($\Delta\theta$)).

Within both free projection and phaseless AFQMC, there are numerous ways to perform the sampling over auxiliary fields. It is advantageous for computational efficiency to restrict the number of walkers to some set value. This is most typically done by periodically annihilating walkers with low weights and duplicating those with high weights with a probability proportional to the distance of the walker's weight from one, and keeping the total number of walkers constant, referred to as Population Control (PC). PC introduces a bias which scales linearly with the inverse number of walkers, which is typically minimal when running with hundreds to a few thousands of walkers, and in the absence of the phase problem leads to stable calculations over hundreds of Ha^{-1}. In this work we implement population control via a "comb" algorithm,[286,287] doing so every 20 time steps, along with measurements of energy. Additionally, the walkers themselves need to be periodically orthonormalized for numerical stability; we do so every two steps. A mean field subtraction is performed prior to propagation to reduce variance.[279]

As an alternative to PC, we have recently introduced a sampling approach based on direct calculation of energy differences between two relevant states (e.g. +2 and +3 states to compute ionization potentials) which is designated 'correlated sampling' (CS).[56] For problems which can be effectively formulated to utilize CS, i.e. similar geometries, it can provide an attractive combination of speed and accuracy, at least for the systems tested to date.[1,2,47] While we do not report values for free projection or correlated sampling in this work, we note that as the local energy evaluation remains identical for both, we expect LO-AFQMC to exhibit similar results.

Evaluation of the Local Energy

The local energy associated with a walker is defined as

$$E = \frac{\langle \phi_T | \hat{H} | \phi \rangle}{\langle \phi_T | \phi \rangle}. \tag{4.8}$$

The two-body contribution to the energy can be written as

$$E_2 = \frac{1}{2} \sum_{pqrs}^{M} V_{pqrs} \frac{\langle \phi_T | a_p^\dagger a_q^\dagger a_s a_r | \phi \rangle}{\langle \phi_T | \phi \rangle}, \tag{4.9}$$

which, when the generalized Wick's theorem is used, gives

$$E_2 = \frac{1}{2} \sum_{pqrs}^{M} V_{pqrs} \sum_{\sigma\tau} (G_{p\sigma,r\sigma} G_{q\tau,s\tau} - G_{p\sigma,s\tau} G_{q\tau,r\sigma}), \tag{4.10}$$

where σ and τ are spin indices. The equal-time Green's function is defined as

$$G_{p\sigma,q\tau} = \frac{\langle \phi_T | a^{\dagger}_{p\sigma} a_{q\tau} | \phi \rangle}{\langle \phi_T | \phi \rangle} = \delta_{\sigma\tau} [\Phi(\Phi_T^{\dagger}\Phi)^{-1}\Phi_T^{\dagger}]^{\mathrm{T}}, \tag{4.11}$$

where $\delta_{\sigma\tau}$ is the Kronecker delta function. Hereafter, indices i, j run over the number of electrons N, and p, q, r, s index the entire set of basis functions M. $|\phi_T\rangle$, the trial wavefunction, constrains the paths of the random walk according to the phaseless constraint. Φ is used to denote the matrix representation of a Slater determinant, with columns representing orbitals, e.g. linear combinations of orthonormal basis functions. These matrices have dimensions $M \times N$ (M = basis size, N = number of electrons). In the case of a CASSCF trial, $|\phi_T\rangle = \sum_d^{N_{det}} c_d |\phi_T^d\rangle$, where every $|\phi_T^d\rangle$ corresponds to a determinant with a maximum of ϵ excitations from the reference determinant.

The expression for the two-body contribution to the local energy is given by

$$E_2 = \sum_{d}^{N_{det}} \frac{c_d \langle \Phi_T^d | \Phi \rangle}{\sum_d^{N_{det}} c_d \langle \Phi_T^d | \Phi \rangle} \times \frac{1}{2} \sum_{pqrs}^{M} V_{pqrs} \sum_{\sigma\tau} (G^d_{p\sigma,r\sigma} G^d_{q\tau,s\tau} - G^d_{p\sigma,s\tau} G^d_{q\tau,r\sigma}), \tag{4.12}$$

where

$$G^d_{p\sigma,q\tau} = \frac{\langle \phi_T^d | a^{\dagger}_{p\sigma} a_{q\tau} | \phi \rangle}{\langle \phi_T^d | \phi \rangle} = \delta_{\sigma\tau} [\Phi(\Phi_T^{d\dagger}\Phi)^{-1}\Phi_T^{d\dagger}]_{qp}. \tag{4.13}$$

In practice, the four index ERI tensor is represented as a three-index factorized tensor $V_{pqrs} = \sum_\alpha^X L^\alpha_{pr} L^\alpha_{qs}$, where X ranges from $4M$ to $10M$ for Cholesky decompositions with thresholds between 10^{-4} and 10^{-6}. A naive implementation of the local energy evaluation would thus scale as $O(N_{det}M^4)$. In typical ph-AFQMC calculations the local energy must be evaluated approximately 4 to 6 million times; the energy evaluation can thus quickly become a computational bottleneck. However, there are many algorithmic tricks one can use that can significantly reduce the scaling of this step, allowing for the extension of AFQMC/CAS to treat large scale *ab initio* systems in an accurate manner. We describe one such possibility for scaling reduction below.

4.2.2 Half-Rotated ERIs

We begin by recognizing that the trial wavefunction is known prior to propagation. We define an intermediate matrix, Q, which partially excludes the dependence of the Green's function on the trial

$$G^{d}_{p\sigma,q\tau} = \sum_{i}^{N} Q^{d}_{qi} \Phi^{d\dagger}_{T,ip}, \tag{4.14}$$

$$Q^{d}_{qi} = [\Phi(\Phi^{d\dagger}_{T}\Phi)^{-1}]_{qi}. \tag{4.15}$$

As both the trial wavefunction and ERIs are known, we can precompute the "half-rotated" cholesky vectors, $\bar{L}^{\alpha}_{ri} = \sum_{p}^{M} L^{\alpha}_{rp}\Phi_{T,pi}$, at a cost of XM^2N_{cas} at the beginning of the calculation, where N_{cas} corresponds to the inactive occupied plus active orbitals (aka each orbital that can be occupied in a trial determinant). N_{cas} can thus be as small as N (for a single determinant) or as large as M for a full CI expansion; for the case of interest here, relatively limited CASSCF active space CI expansions, $N_{cas} \simeq N$. Use of this precomputed tensor and the Q intermediates results in an algorithm which scales as $N_{det}XMN^2$

$$E_2 = \sum_{d}^{N_{det}} \frac{c_d\langle\Phi_d|\Phi\rangle}{\sum_{d}^{N_{det}} c_d\langle\Phi_d|\Phi\rangle} \times \frac{1}{2} \sum_{ij}^{N_{cas}} \sum_{rs}^{M} \sum_{\alpha}^{X} \bar{L}^{\alpha}_{ri}\bar{L}^{\alpha}_{sj} \sum_{\sigma\tau}(Q^{d}_{r\sigma,i\sigma}Q^{d}_{s\tau,j\tau} - Q^{d}_{s\sigma,i\tau}Q^{d}_{r\tau,j\sigma}). \tag{4.16}$$

The cost of this energy evaluation can further be reduced via precomputation of the sum over X, resulting in $O(N_{det}M^2N^2)$ scaling at the cost of storing a $N_{cas}^2M^2$ dimensional tensor in memory, $Y_{ijrs} = \sum_{\alpha}^{X} \bar{L}^{\alpha}_{ri}\bar{L}^{\alpha}_{sj}$, giving

$$E_2 = \sum_{d}^{N_{det}} \frac{c_d\langle\Phi_d|\Phi\rangle}{\sum_{d}^{N_{det}} c_d\langle\Phi_d|\Phi\rangle} \times \frac{1}{2} \sum_{ij}^{N_{cas}} \sum_{rs}^{M} Y_{ijrs} \sum_{\sigma\tau}(Q^{d}_{r\sigma,i\sigma}Q^{d}_{s\tau,j\tau} - Q^{d}_{s\sigma,i\tau}Q^{d}_{r\tau,j\sigma}). \tag{4.17}$$

Note that if i or j are in the active space of a CAS trial, they might not be present in every determinant. In order to effectively perform the sum over determinants in parallel, we find it advantageous to store the set of Q^d matrices as if they include every orbital, substituting with zeros when this is not the case. While the HR-ERI algorithm is generally faster for small CI expansions and smaller systems in general, this approach can quickly present a bottleneck for large systems. This is especially apparent in calculations utilizing GPUs, due to the inability to store the half-rotated Y_{ijrs} on a single GPU. This renders large (and expensive) interconnected GPU clusters a requirement, and results in the efficiency being primarily controlled by memory passing protocols, which can easily lead to major slowdowns for medium to large systems around M greater than around 1000. When running on clusters with fast GPU to GPU transfer speeds, however, the effect can be minimal, and it is often then advantageous to precompute Y_{ijrs} even in the case where it cannot fit on a single GPU.

4.2.3 LO-AFQMC Algorithms

We here present the general structure of localized ERIs, before outlining the application and use of localization to the HR-ERI algorithm.

Compression of the HR-ERI Tensor

We begin by noting that the structure of the half-rotated tensor, Y_{ijrs}, corresponds to N_{cas}^2 blocks of dimension $M \times M$ corresponding to the interaction integrals between the occupied orbital pair $[ij]$, which we denote $\{Y_{rs}^{[ij]}\}$. It is well known that the low-rank structure of the Coulombic integrals is best revealed when dealing with interactions between distinct localized orbitals.[288] This low rank structure can be taken advantage of in numerous ways; we do so via a block-wise singular value decomposition for each localized orbital pair $[ij]$, which is then truncated according to a threshold beyond which singular values are discarded, here denoted T_{SVD}. The result of this procedure is an $N_{cas} \times N_{cas}$ list of rectangular matrices with dimensions $M \times M_{SVD}$ and $M_{SVD} \times M$

with M_{SVD} being the truncated dimension. Thus, the expression for $Y_{rs}^{[ij]}$ may be written

$$Y_{rs}^{[ij]} \simeq \sum_{K^{[ij]}}^{M_{SVD}^{[ij]}} U_{rK}^{[ij]} \Sigma_K^{[ij]} V_{Ks}^{[ij]} = \sum_{K}^{M_{SVD}} \bar{U}_{rK}^{[ij]} \bar{V}_{Ks}^{[ij]}, \tag{4.18}$$

where $U^{[ij]}$ and $V^{[ij]}$ are the set of unitary left and right singular vectors, and $\Sigma^{[ij]}$ are their associated singular values. For notational clarity, we omit the indices [ij] from M_{SVD} and K throughout the remainder of the paper. At $M_{SVD} = M$ (i.e. $\text{T}_{SVD} = 0$), this expression becomes exact. This compression is performed once in the beginning of the simulation, at a cost of $O(N^2M^3)$, which is easily parallelized over GPUs for a speedup of $\frac{1}{N_{GPU}}$, and represents a minimal ($\simeq$ 1 to 2%) addition to the time of the calculations outlined in this study. We choose to allow M_{SVD} to vary between [ij] pairs, instead fixing T_{SVD} to obtain a given accuracy. It is additionally possible to employ approximate versions of the singular value decomposition, which would reduce the scaling of this step significantly, although we did not explore this in this work. In practice, $\langle M_{SVD} \rangle$, the average compressed dimension, does not scale with the system size (and in fact decreases with a given T_{SVD} as a function of the system size for a given class of molecules, as we demonstrate in section 4.3), and thus the memory scaling is effectively reduced from quartic to cubic in system size, namely $N_{cas}^2 M$.

LO-AFQMC in the HR-ERI Algorithm

We next explicitly express the energy for a CASSCF trial using the HR-ERI algorithm within the compressed framework using

$$E_2 = \sum_{d}^{N_{det}} \frac{c_d \langle \Phi_d | \Phi \rangle}{\sum_d^{N_{det}} c_d \langle \Phi_d | \Phi \rangle} \times \frac{1}{2} \sum_{ij}^{N_{cas}} \sum_{K}^{M_{SVD}} \sum_{rs}^{M} \bar{U}_{rK}^{[ij]} \bar{V}_{Ks}^{[ij]} (4Q_{r,i}^d Q_{s,j}^d - 2Q_{s,i}^d Q_{r,j}^d). \tag{4.19}$$

The complexity for a direct summation is now $O(N_{det}N^2M\langle M_{SVD}\rangle)$, and for a single determinant trial this is straightforward and results in $O(N^2M\langle M_{SVD}\rangle)$ scaling. Due to the introduction of a low rank index K, however, we can extend the savings for the half-rotated algorithm further with respect to the number of determinants by taking advantage of the structure of the CAS trial. In the HR-ERI algorithm we use a Sherman-Morrison-Woodbury (SMW) algorithm to update the overlap of each walker with the determinants of the trial. In other words, we can write the N_{det} intermediate Green's functions Q^{d}_{ri} of Eq. 4.15 as a single contribution from the reference, along with rank ϵ corrections specific to each determinant

$$Q^{d}_{ri} = \Phi * (A + U^{d}V^{d\,\mathrm{T}})^{-1} = \Phi A^{-1} - \Phi A^{-1}U^{d}(I + V^{d\,\mathrm{T}}A^{-1}U^{d})^{-1}V^{d\,\mathrm{T}}A^{-1}, \tag{4.20}$$

where $A = \Phi^{0\dagger}_{T}\Phi$, $\Phi^{0\dagger}_{T}$ is the reference determinant, and U^{d} and $V^{d,\mathrm{T}}$ are N by ϵ matrices corresponding to the rows which are changed via excitation from the reference for a given determinant. From here on ϵ will refer to the maximum number of excitations possible in the CAS trial, which is typically $\simeq 6$ and bounded by the size of the active space. The scaling of the formation of the set of all Q matrices has now been transformed from $O(N_{det}N^{2}M)$ to $O(N^{2}M + N_{det}N\epsilon^{2})$ when in the molecular orbital basis (see Sec. 4.5.4). If one formed the full set of N_{det} $N \times M$ matrices Q^{d}_{ri}, as in the full HR-ERI algorithm, the resulting operations in Eq. 4.19, computed in a loop over determinants, would scale as $O(N_{det}N^{2}M\langle M_{SVD}\rangle)$. However, If we rearrange the sums in Eq. 4.19 so as to avoid the explicit computation of Q^{d}_{ri}, we can effectively replace another factor of the full basis dimension, M, with $\langle M_{SVD}\rangle$.

Since the LO-ERI are specific to the pairs $[ij]$, we must operate on a column of the Q matrix, $Q^{d}_{r,i}$, corresponding to the specific orbital i,

$$Q^{d}_{r,i} = \Phi A^{-1}_{i} - \Phi A^{-1}U^{d}(I + V^{d\,\mathrm{T}}A^{-1}U^{d})^{-1}V^{d\,\mathrm{T}}A^{-1}_{i}. \tag{4.21}$$

Note that the U^{d} matrix has a block structure of zero rows followed by an identity (or permutation matrix) of maximum dimensions of $\epsilon \times \epsilon$, and so $A^{-1}U^{d}$ is equal to the last ϵ columns (or permutations thereof) of A^{-1} and requires no computation. Additionally, note that each Q has N total columns, whereas the index i in the energy evaluation (see Eq. 4.22), is summed over N_{cas}; thus the index "i" in Eq. 4.22 corresponds to the index of the column of the d-th determinant in the original CAS MOs. This means that for a specific determinant, the i-th column might not exist - thus, to allow for efficient operations in parallel on GPUs ("batching"), which is most effective for processes with equivalent operation count, we assume that they do in memory, substituting with zeros if needed.

We can now combine Eqns. 4.17 and 4.21. This involves effectively performing the sum over full basis dimensions r and s first. We then form four vectors of dimension $\langle M_{SVD}\rangle$, $\bar{Q}^d_{K,L,i} = \bar{U}^{[ij],\mathrm{T}}_{rK} * Q^d_{ri}$, $\bar{Q}^d_{K,R,i} = \bar{V}^{[ij]}_{Ks} * Q^d_{si}$, $\bar{Q}^d_{K,L,j} = \bar{U}^{[ij],\mathrm{T}}_{rK} * Q^d_{rj}$, and $\bar{Q}^d_{K,R,j} = \bar{V}^{[ij]}_{Ks} * Q^d_{sj}$, for every $[ij]$ pair and every determinant. If we multiply by the determinant coefficients and overlaps in this step, the resulting sums over $[ij]$, determinants, and K, are simply a dot product (e.g. for the Coulomb interaction $\bar{Q}^{d,[ij]}_{K,L,i} * \bar{Q}^{d,[ij]}_{K,R,j}$, scaling as $O(N_{det}N^2\langle M_{SVD}\rangle)$). Explicitly, we write

$$E_2 = \sum_{d}^{N_{det}} \frac{c_d\langle\Phi_d|\Phi\rangle}{\sum_d^{N_{det}} c_d\langle\Phi_d|\Phi\rangle} \times \frac{1}{2}\sum_{ij}^{N_{cas}}\sum_{K}^{M_{SVD}} (4\bar{Q}^{d,[ij]}_{K,L,i}\bar{Q}^{d,[ij]}_{K,R,j} - 2\bar{Q}^{d,[ij]}_{K,L,j}\bar{Q}^{d,[ij]}_{K,R,i}), \tag{4.22}$$

$$\bar{Q}^{d,[ij]}_{K,L,i} = \bar{U}^{[ij],\mathrm{T}}_{rK} * Q^d_{ri} = \bar{U}^{[ij]}_{K,r}\Phi A^{-1}_i - \bar{U}^{[ij]}_{Kr}\Phi A^{-1}U^d(I + V^{d,\mathrm{T}}A^{-1}U^d)^{-1}V^{d,\mathrm{T}}A^{-1}_i \tag{4.23}$$

Consider now the formation of $\bar{Q}^{d,[ij]}_{K,L,i}$, for which the scaling of key intermediates has been outlined in Table 4.1. Note that for every $[ij]$ pair, we only require $\bar{U}^{[ij]}_{Kr}$ intersect with vectors A^{-1}_i and A^{-1}_j, instead of the entire $M \times M$ matrix A^{-1}. For the reference section, i.e. the first term of the right hand side of Eq. 4.23, the computational order is straightforward. Specifically, we store the full ΦA^{-1}, then define vectors corresponding to i and j, and lastly multiply $\bar{U}^{[ij]}_{Kr}$ by these vectors. This set of steps scales as $N^2M\langle M_{SVD}\rangle$, and must occur four times, once for every $\bar{Q}$ in Eq. 4.22.

Intermediate	Operation	Memory scaling	Computational scaling
$\{Q^{\text{int}}_L\}_{[ij],d,K}$	$\bar{U}^{[ij]}_{Kr} \times (\Phi A^{-1}U^d)$	$N^2\langle M_{SVD}\rangle\epsilon$	$N^2M\langle M_{SVD}\rangle\epsilon$
$\{S\}_d$	$(I + V^{T,d}A^{-1}U^d)^{-1} \times V^{T,d}A^{-1}$	$N_{det}N\epsilon$	$N_{det}N\epsilon^2$
$\{\bar{Q}_{L,i}\}_{[ij],d,K}$	$Q^{\text{int}}_L \times S_i$	$N_{det}N^2\langle M_{SVD}\rangle$	$N_{det}N^2\langle M_{SVD}\rangle\epsilon$

Table 4.1: List of key intermediates in the LO-AFQMC evaluation of the 2-body energy using SMW, including the memory and computational scaling. The nomenclature $\{X\}_{a,b,c}$ indicates that the scaling listed is for the formation of the set of intermediates X for all valid indices a, b, c.

For the SMW correction term (the rightmost term in Eq. 4.23), we must take advantage of the block identity form of U^d, namely that the first N_{inact} (occupied orbitals which are not in the active space) columns of ΦA^{-1} are zero regardless of the determinant. Thus we need only compute $Q_L^{\text{int}} = \bar{U}_{Kr}^{[ij]} \Phi A_{act}^{-1}$ once, which scales as $N^2 M \langle M_{SVD} \rangle \epsilon$, where A_{act}^{-1} is the last ϵ columns of A^{-1}. This must be performed twice, once for $\bar{U}_{Kr}^{[ij]}$ and once for $\bar{V}_{Ks}^{[ij]}$.

If one generates the N_{det} intermediates of size $\epsilon \times N$, corresponding to $\{S\}_d = (I + V^{d,\text{T}} A^{-1} U^d)^{-1} V^{d,\text{T}} A^{-1}$ (see details in Sec. 4.5.4), it is straightforward to determine the column of these intermediates corresponding to i for each determinant and then multiply the stored Q_L^{int} by these column vectors and subtract them from the reference value. This series of steps then scales as $N_{det} N^2 \langle M_{SVD} \rangle \epsilon$. The total scaling of the HR-ERI energy evaluation is thus reduced from $(N_{det} N^2 M \langle M_{SVD} \rangle + N_{det} N^2 \langle M_{SVD} \rangle)$ to $(N^2 M \langle M_{SVD} \rangle (\epsilon + 1) + N_{det} N^2 \langle M_{SVD} \rangle \epsilon)$. Keeping only dimensions that scale with the system size, this is thus an asymptotically $O(N^2 M + N_{det} N^2)$ scaling algorithm. While the use of SMW to speed up calculations has been well documented, we additionally provide details with respect to the generation of the Green's function in the MO basis, for 1-body energetic terms and the evaluation of the force bias in Sec. 4.5.4.

4.3 Results

Here we present results illustrating the efficiency and accuracy of the LO-AFQMC approach. The specifics of the localization method used were not observed to significantly affect results, thus only a Foster-Boys localization scheme was used to produce the results shown here.[289] Only the inactive orbitals were localized; localization of the active orbitals after CASSCF iterations reduced the efficiency of the CI expansion, resulting in orders of magnitude more determinants being required in the CI expansion to produce the same total sum of CI weight, and thus a loss of efficiency.

We first investigate dodecane in a cc-pVDZ basis as a medium sized test case. In order to estimate the error due to compression of the HR-ERI tensor, we take the difference in energies between that arising from the full and localized HR-ERI tensors to evaluate the local energy of a single walker (Figure 4.1). The error estimated in this manner converges with respect to the threshold T$_{SVD}$, with T$_{SVD}$ = 0 reproducing the full AFQMC result. For ph-AFQMC, our goal is simply to reduce systematic errors to below the statistical error that one is seeking, which is generally around 1 mHa for molecular systems. This is achieved for a single energy evaluation for dodecane around T$_{SVD}$ = 5×10^{-4}, with an overall compression efficiency ($\frac{M - \langle M_{SVD} \rangle}{M} * 100$) of 96 percent. We note that the reduction in error with T_{SVD} does not occur monotonically.Error reduction does decrease systematically, however, and the magnitude of fluctuations at any given threshold are proportional to the threshold itself. To assess the increased scaling due to this compression, we performed calculations for 16 walkers on a single GPU, and present the speedup versus storing and using the full HR-ERI tensor. Even the tightest threshold tested (T$_{SVD}$ = 10^{-5}) is over 15x faster than our HR-ERI code for a single energy evaluation.

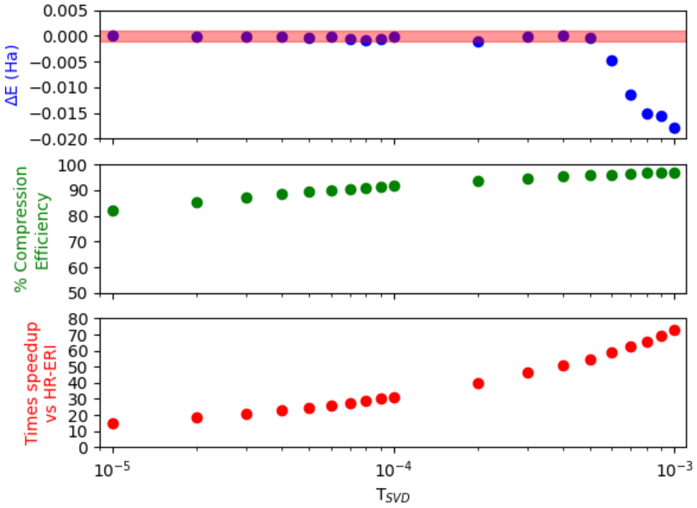


Figure 4.1: Convergence of the LO-AFQMC error with regards to T$_{SVD}$ for dodecane in the cc-pVDZ basis. The error vs regular AFQMC (top, blue) converges to within 1 mHa (red shaded region) around T$_{SVD}$ = 0.0005. Note that values to the left have a tighter threshold, and that the x axis is in a logarithmic scale. The compression efficiency (middle, green) ranges between 82 and 97 %, with an intermediate T_{SVD} = 10^{-4} threshold reducing M by 91.75%. The reduced computational scaling due to the compression is reflected in the speedup vs regular HR-ERI (bottom, red) of the energy evaluation for 16 walkers, which even for the tightest threshold tested (T$_{SVD}$ = 10^{-5}) is over 15x faster than our HR-ERI code, with T_{SVD} = 10^{-4} being $\simeq$ 31 times faster. These calculations were performed on a single NVIDIA GeForce RTX 3090.

To illustrate the necessity of using localized orbitals, we compare LO-AFQMC (T$_{SVD}$ = 10^{-4})

using both localized and non-localized orbitals for dodecane (Fig. 4.2). It can be seen that the degree of compression is significantly reduced when using the canonical basis, highlighting the intuitive fact that localized and distant orbital pairs exhibit low rank blocks within the HR-ERIs, similar to observations in other local correlation methods.[290]

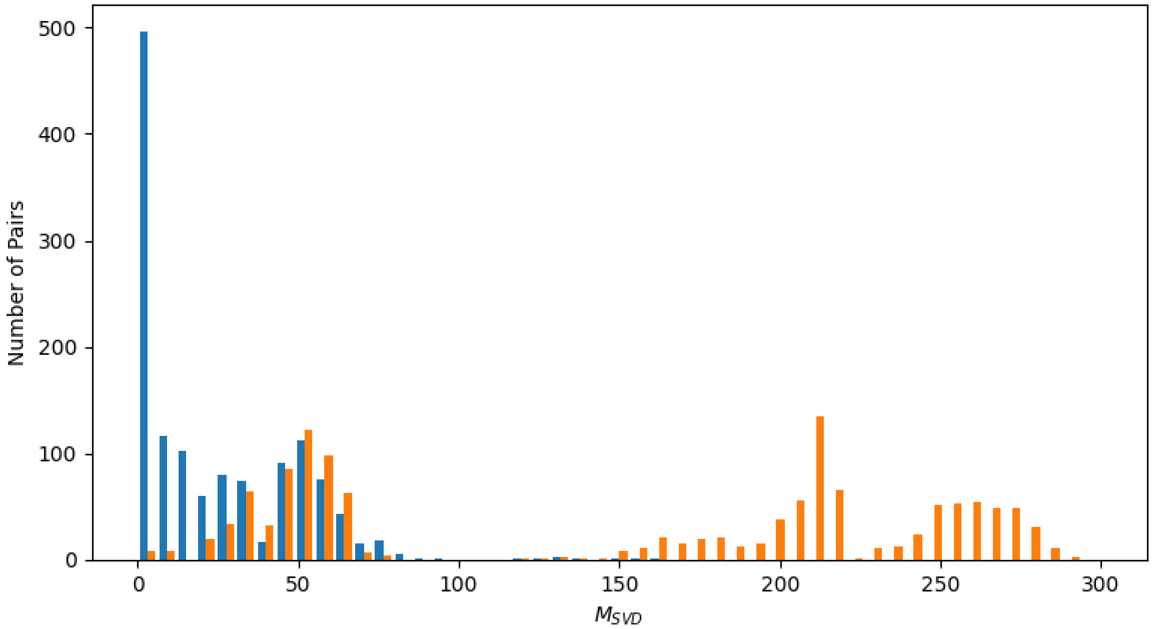


Figure 4.2: A histogram of M_{SVD} for LO-AFQMC in the canonical (orange) and localized (blue) molecular orbital bases for dodecane with $T_{SVD} = 10^{-4}$. Localization results in an increase in compression efficiency from 48.89% to 91.75%. Note that the full basis size M = 298.

4.3.1 Metallocenes

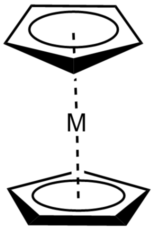


Figure 4.3: The chemical structure of a generic metallocene, $M(Cp)_2$

While our results for dodecane are promising, it is necessary to test whether or not this localization scheme works for more challenging molecules where AFQMC may be a method of choice.

Table 4.2: Errors in total energy for ferrocene (cc-pVTZ-dkh) vs an identical AFQMC calculation with full ERIs for a series of SVD truncation thresholds T_{SVD}. $T_{SVD} = 10^{-4}$ provides a good balance between memory efficiency and accuracy.

T_{SVD}	ΔE(Ha)	% Compression
0.00005	0.0003(13)	70.84
0.0001	0.0010(10)	77.50
0.00025	0.0013(15)	85.04
0.0005	0.0034(11)	89.52
0.00075	0.0049(11)	91.64
0.001	0.0086(11)	92.91

Recently, we presented ph-AFQMC data on the gas phase ionization energy of a series of metallocenes (Fig. 4.3), for which we found it necessary to use multideterminant CASSCF trials. Here, we use calculations on ferrocene as benchmark results for LO-AFQMC, using geometries from Ref. 2. Population control calculations for Fe(III)(Cp)$_2$ were run with a series of T_{SVD} values and compare our results to the full AFQMC total electronic energies in a cc-pVTZ-dkh basis (Table 4.2). We again find $T_{SVD} = 10^{-4}$ to be a good choice to reduce the error in the total energy to the level of statistical error, while still maintaining a compression efficiency of 77.5%. We hereafter then use $T_{SVD} = 10^{-4}$ as the default threshold for the remaining results, although one might need to increase the threshold if one desires an increased accuracy.

4.3.2 Polyacenes

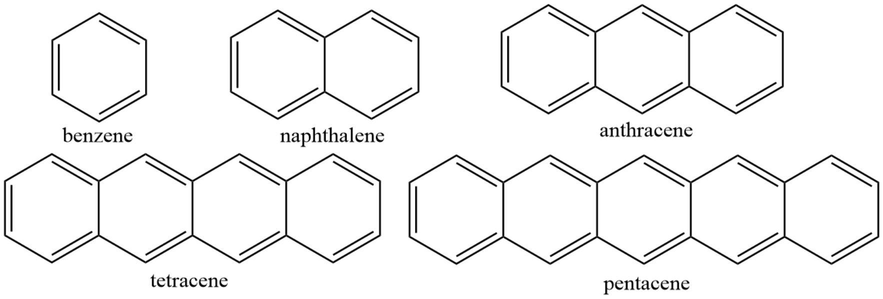


Figure 4.4: Chemical structure of the polyacenes 1-5

While the performance for metallocenes is reasonable, the localized orbital approximation is likely to have a greater effect on the accuracy of ph-AFQMC (as well as a better compression efficiency) as more electron pair interactions are approximated. Thus it is important to assess the scaling of the error of the approximation with respect to the system size. To this end, the singlet-triplet (ST) gaps of the linear polyacenes provide a set of aromatic, quasi one-dimensional systems for which ph-AFQMC has already been shown to be robust.[72] As these systems exhibit significant delocalization throughout their π systems, we expect the calculation of this quantity to provide a stringent test of the accuracy and efficiency of LO-AFQMC. Minimal CASSCF trials (6e 6o) were obtained for each system in the cc-pVTZ basis, for both singlet and triplet spin states using the 3 highest and 3 lowest energy occupied and unoccupied RHF orbitals, respectively, following Ref. 275 for benzene. Geometries were taken from Ref. 72.

In order to evaluate the error due to the LO approximation, both regular AFQMC and LO-AFQMC with $T_{SVD} = 10^{-4}$ were run for the entire polyacene set. 1656 walkers were run for 300 Ha^{-1}, maintaining 99.5% of the CI weight for each trial wavefunction, and sampling the same auxiliary fields to attempt to remove discrepancies due to rare sampling events. Table 4.3 shows the total energies for each system, as well as the LO error, defined as the difference between the total energy derived from AFQMC and that from LO-AFQMC. In all cases, the LO error remains below 1 kcal/mol, and additionally below that of the statistical error of AFQMC, rendering LO-AFQMC statistically equivalent to AFQMC for these cases. As expected, as the system size increases, the efficacy of localization increases, with pentacene exhibiting a 93.4% compression. Note that the error does not necessarily cancel between similar species, and additionally does not scale with system size.

The AFQMC and LO-AFQMC singlet triplet gaps for the entire series are presented in Table 4.4 and compared to previous ph-AFQMC calculations as well as experiment. In most cases the use of an CASSCF trial, even the minimal 6e 6o ones used here, results in a smaller error compared to experiment, with an MAE of 1.3 ± 0.8 kcal/mol vs the previously reported ph-AFQMC (using an unrestricted Kohn-Sham (UKS) trial) MAE of 2.8 kcal/mol. LO-AFQMC retains this accuracy with an MAD of 1.8 ± 0.8 kcal/mol versus experiment, and the LO error ranging from 0.4 to 1.1 kcal/mol, and always within the inherent statistical error of AFQMC.

Table 4.3: Total energies (Ha) for the singlet (S_0) and triplet (T_1) states of the series of polyacenes, as calculated by AFQMC and an identical LO-AFQMC calculation with $T_{SVD} = 10^{-4}$.

System	State	% Comp.	AFQMC (Ha)	LO-AFQMC (Ha)	LO Error (kcal/mol)
benzene	S_0	67	-231.9932(5)	-231.9928(4)	0.2±0.4
	T_1	69	-231.8436(6)	-231.8439(5)	-0.2±0.5
naphthalene	S_0	81	-385.4733(6)	-385.4739(6)	-0.4±0.5
	T_1	81	-385.3682(7)	-385.3674(6)	0.5±0.6
anthracene	S_0	86	-538.9477(8)	-538.9488(8)	-0.7±0.7
	T_1	87	-538.8735(10)	-538.8741(7)	-0.4±0.8
tetracene	S_0	90	-692.4214(8)	-692.4207(9)	0.4±0.8
	T_1	91	-692.3684(13)	-692.3684(10)	0.1±1
pentacene	S_0	94	-845.8904(14)	-845.8918(11)	-0.9±1.1
	T_1	93	-845.8555(8)	-845.8552(13)	0.2±0.9

Table 4.4: ST gaps in kcal/mol for the polyacenes (cc-pVTZ). The experimental values and ph-AFQMC with a UKS trial were obtained from Ref 72; the experimental values are corrected to compare to electronic energies from AFQMC using vibrational free energy corrections from B3LYP as described in Ref 72.

System	CAS/AFQMC	CAS/LO-AFQMC	LO Error	UKS/AFQMC[72]	Expt.[72]
benzene	93.9±0.5	93.4±0.4	-0.4±0.6	N/A	
naphthalene	65.9±0.6	66.8±0.5	0.9±0.8	68(1.2)	64.4
anthracene	46.6±0.8	46.9±0.7	0.4±1.1	46.2(1.2)	45.4
tetracene	33.2±1	32.9±0.8	-0.4±1.3	34(1.6)	31.2
pentacene	21.9±1	23±1.1	1.1±1.4	25.2(1.6)	21.3
MAD vs Expt.	1.3±0.8	1.8±0.8		2.8±1.4	

As we sample walkers along the imaginary time trajectory, there is a small variation in the LO error. To quantify this we take a sample of around 100,000 walkers for the case of benzene, and evaluate the local energy error due to compression (Fig. 4.5). The error follows a normal distribution at each timestep, which then varies slightly over imaginary time and results in an additional small ($\simeq 0.1$kcal/mol) increase in the statistical error.

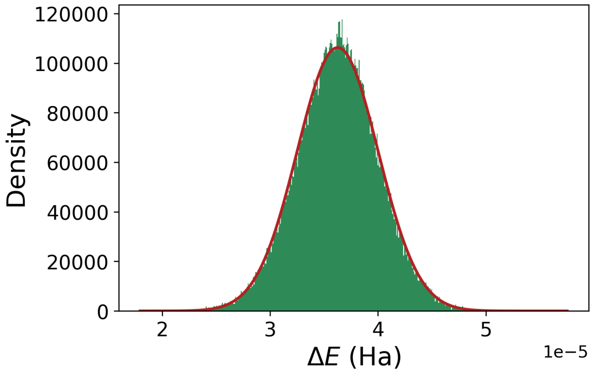


Figure 4.5: Distribution of LO error over 100,244 walkers after a single block (0.1 Ha^{-1}) of propagation for benzene, fitted to a normal distribution with a mean of 3.6×10^{-5} and standard deviation of 3.75×10^{-6}.

To investigate whether the insensitivity of the LO error to the size of the molecule was consistent across imaginary time (i.e. independent of the particular set of walkers we ran an energy evaluation using the non-compressed HR-ERI tensor at a single time step (somewhat arbitrarily chosen to be at 280 Ha^{-1}), using an identical set of walkers as the LO-AFQMC calculation. The resulting compression errors for the singlet total energies (Fig. 4.6) do in fact scale near linearly with system size (M), as was originally expected for systems with similar electronic structure. For all systems, the localization error is less than 1 kcal/mol, although for the case of pentacene the error due to compression as estimated at a single time step is actually larger than the statistical error after reblocking over imaginary time. This suggests that the error due to localization, while averaging out to a consistently small value, can vary significantly depending on the stochastic paths taken by the set of walkers. Due to this, we believe it may be prudent for applications which depend on the accuracy of the total energy to tighten T$_{SVD}$ as the system size grows, although the scaling of T_{SVD} was not necessary for any of the applications presented in this work. As the efficacy of compression demonstrably increases as the system size increases, we expect the overall savings for percent reduction in memory to remain sizeable, should one tighten T$_{SVD}$ for larger systems.

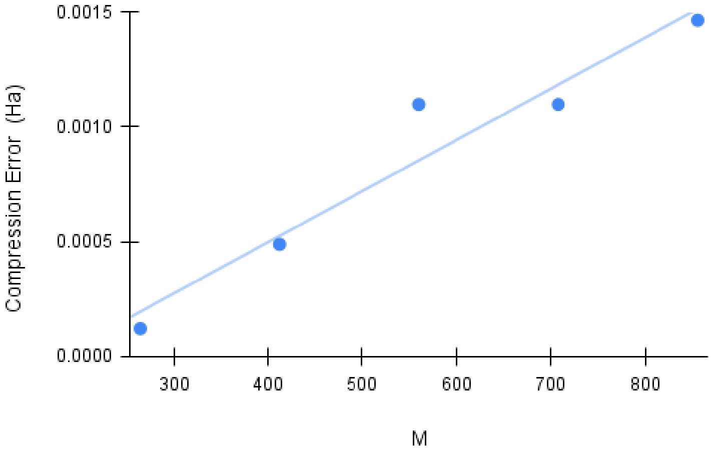


Figure 4.6: Scaling of the error of LO-AFQMC vs full ph-AFQMC, estimated at a time slice of 280 Ha^{-1}, with respect to the basis size for fixed value of $T_{SVD} = 10^{-4}$. Error bars are not plotted since the error is reported at a single timestep.

4.3.3 Platonic Hydrocarbons

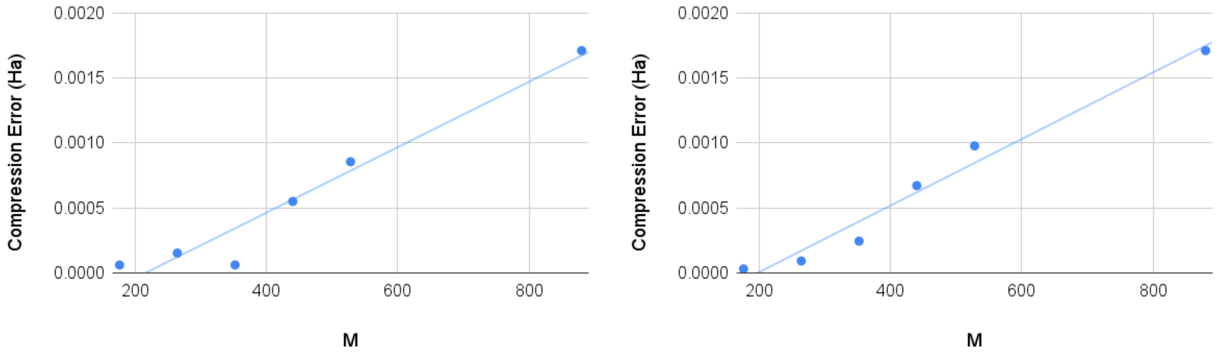
While the polyacenes are a good test to observe the scaling of localization error in a quasi one-dimensional setting, it is useful to test three-dimensional systems. Here, we choose a set of hydrocarbon cage molecules in the shape of the platonic solids, for which benchmark CCSD(T) heats of formation and geometries are available.[291] Total heats of formation for LO-AFQMC using both cc-pVDZ and cc-pVTZ basis sets are presented in Table 4.5, along with LO-AFQMC errors versus full, standard ph-AFQMC calculations and the compression efficiency. Although AFQMC at the cc-pVTZ level reaches the bounds of chemical accuracy versus the cited CCSD(T)/CBS values for the smaller systems, the errors versus the reference CCSD(T) energies increase with the system size. This is not, however, attributed to the LO approximation; the error versus a full ph-AFQMC calculation was in all cases less than the statistical error of ph-AFQMC (LO Error in Table 4.5). Two tests (Fig. 4.7) using LO-AFQMC and full ph-AFQMC to measure the energy of specific sets of walkers at imaginary times 2 Ha^{-1} and 250 Ha^{-1}, representing the beginning and end of an AFQMC simulation respectively, suggest that the linear increase in LO error seen in the polyacenes is indeed reproduced in the platonic hydrocarbons, with some expected variation in error between timesteps. It thus appears that the averaging over imaginary time can again lead to

	CCSD(T)	LO-AFQMC	LO Error	% Comp.
	CBS	cc-pVTZ	cc-pVTZ	cc-pVTZ
C_4H_4	793.9	794.22±0.34	0±0.48	50.4
C_6H_6	1254.32	1255.06±0.49	-0.02±0.69	67.1
C_8H_8	1706.99	1708.58±0.46	-0.15±0.64	72.1
$C_{10}H_{10}$	2196.46	2201.26±0.5	0.76±0.71	82.4
$C_{12}H_{12}$	2718.29	2724.36±0.56	-0.23±0.8	85.9
$C_{20}H_{20}$	4621.46	4633.52±0.8	-0.42±1.41	84.2

Table 4.5: Total Atomization Energies in kcal/mol for the platonic hydrocarbon cages; CCSD(T)/CBS values were taken from Ref. 291. LO Error refers to the difference between the LO-AFQMC result and an identical ph-AFQMC calculation run without compression.

some favorable cancellation of LO error, which is further supported by a small test calculation in the case of C_8H_8 in section 4.5.5. Again, as before, the efficacy of the compression increases with system size, with $\mathrm{C}_{20}\mathrm{H}_{20}$ exhibiting 84% compression.

Figure 4.7: Scaling of the error of LO-AFQMC vs full ph-AFQMC, estimated at a time step of 2 Ha^{-1} (a) and 250 Ha^{-1} (b), with respect to the basis size (cc-pVTZ). Error bars are not plotted since the error is evaluated at a single timestep. Note that the error for each time step scales linearly, although there is notable variation as the walker distributions change.



(a)

(b)

4.3.4 GPU implementation and timing

LO-AFQMC effectively mediates the memory bottlenecks apparent in the HR-ERI algorithm, in addition to reducing the computational complexity and scaling with respect to system size. However, we have not discussed to what degree the actual wall time can be reduced on a GPU system, as for these system sizes wall time is, more often than not, limited by the ability to perform operations in parallel. Since we are dealing with dense linear algebra, it is possible to perform most operations in parallel for all determinants (aka “batch” operations over determinants). As we allow M_{SVD} to vary between each pair $[ij]$, great care must be taken to enable efficient simultaneous batching of operations, especially if one intends to use standard cuBLAS libraries which are designed for calculations with consistent dimension.[283] By storing the compressed ERIs contiguously in GPU memory, we were able to batch nearly all operations over the total number of electron pairs and determinants simultaneously, resulting in a significant gain in efficiency. Some details are given in section 4.5.3. Additionally, as each walker is undergoing an identical set of operations, we were able to simultaneously evaluate the local energy for all walkers stored on a GPU at once. Such batching is ineffective in the HR-ERI algorithm due to the large memory requirements (and thus inability to store intermediates for each walker), but with the reduced memory requirements, batching LO-AFQMC over walkers significantly improves efficiency, as was previously reported for alternate algorithms in Ref 112. To illustrate this, we ran LO-AFQMC for $Fe(Cp)_2$ in the cc-pVTZ-DK basis, with 200 determinants, 1656 walkers, and propagated to 200 Ha^{-1}, while modifying the number of walkers for which operations are performed in parallel on a single GPU. The resulting timings in GPU-hours can be seen in Fig. 4.8; at 16 walkers performed in parallel, the total GPU-hours is reduced by over 3 times versus in series, to 65.9 GPU-hours.

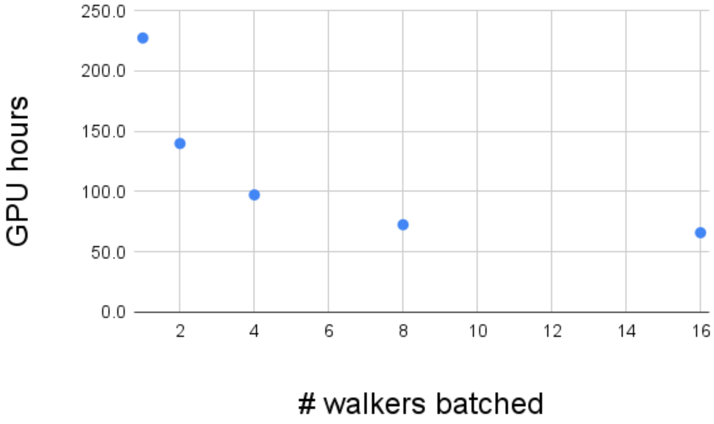


Figure 4.8: Effect of performing operations for increasing numbers of walkers in parallel on a single GPU, for Fe(Cp)$_2$ in the cc-pVTZ-DK basis (508 basis functions, 200 determinants). The GPU-hours represent the time necessary to propagate 1656 walkers over 200 Ha^{-1}, evaluating the energy 2000 times in total per walker, on the Summit supercomputer at Oak Ridge National Laboratory.

To showcase the speedups possible due to the LO-AFQMC algorithm, as well as the stability of our previous results with respect to use of the LO-AFQMC algorithm, we additionally re-ran the calculations in both cc-pVTZ-DK and cc-pVQZ-DK for the adiabatic ionization energy of Ni(Cp)$_2$ from Ref. 2 using $T_{SVD} = 10^{-4}$. For a significantly reduced cost, we were able to reproduce our previous result in the complete basis set (CBS) limit to within statistical error (-141.86±1.76 kcal/mol versus our previous -143.39±1.64 kcal/mol), which is additionally within error bars of the experimental result (-143.8±1.5 kcal/mol). To put this in perspective, our previous ph-AFQMC calculation of the Ni(III)(Cp)$_2$ species in the QZ basis (954 basis functions) required 1615 GPU-hours, whereas LO-AFQMC took 102 GPU hours to propagate the same number of walkers the same distance in imaginary time, reproducing the energy to within statistical error in 6.3% of the time. Of those 102 GPU hours, approximately 18.9% were spent in evaluating the energy (both 1- and 2-body terms), slightly more than the approximately 18.7% that were spent forming the propagation matrix ($B = \sum_\alpha x_\alpha L_\alpha$, scaling as XM^2). Although the formation of B is ≃20 times faster than the energy evaluation, it must be performed every step, whereas we only measure the energy once every 20 steps. The compression rates of the neutral/cationic Ni(Cp)$_2$ species were

84/86% and 86/89% for the cc-pVTZ-DK and cc-pVQZ-DK bases, respectively, slightly higher than that of $Fe(Cp)_2$.

4.4 Conclusions

We have presented a localized orbital compression of the ERIs as a means to reduce the memory footprint and computational scaling of the energy evaluation in AFQMC calculations. Our algorithm has been demonstrated to be efficient and robust for a variety of difficult test systems in manner which is independent of cancellation of error between two electronic states. Although here we restricted our focus to the HR-ERI algorithm of phaseless AFQMC, the approach and approximation errors reported are general.

As the value of interest is not the controlled error of SVD (namely the L^2 norm of the matrix error), but rather the total energy estimated using the compressed integrals, we have taken care to estimate the error due to our localization approximation in multiple ways. We compare to full ph-AFQMC for a single energy evaluation, a weighted average over walkers at a single time step, and full population control AFQMC calculations averaged over imaginary time. In all cases we observe controllable if not entirely monotonic errors for each metric. In addition, we compare our results to experiment and CCSD(T) calculations, highlighting that the accuracy of ph-AFQMC with respect to these external benchmarks remains unaffected due to the LO approximation.

While the scaling of the algorithms outlined is dependent on the number of electron pairs N^2 (i.e. $O(N^2(M+N_{det}))$), the formal asymptotic rank of the ERIs is expected to be quadratic $O(M^2)$, and so we foresee the possibility to improve upon this localization procedure significantly for larger systems. In particular, one might envision screening pairs of electrons, and only including those that contribute meaningfully to the total energy. This would asymptotically reduce the number of pairs to N from N^2 and thus reduce the memory scaling of this tensor to quadratic, namely $O(NM)$, as well as producing a corresponding reduction in computational effort.

While LO-AFQMC significantly reduces the computational effort required for ph-AFQMC, it is still prohibitively expensive compared to, say, density functional theory (DFT). We thus do not expect LO-AFQMC to directly compete with fast quantum chemistry methods. However, the performance of DFT for transition metal containing molecules has been sub-standard, in no small part due to the lack of trainable reference data for relevant cases. AFQMC can obtain reference values on which to train DFT functionals, and when extended in viable system size via methods like LO-AFQMC, the task of producing a large benchmark dataset with a variety of represented chemistries begins to become feasible. This philosophy is akin to some recent work in the coupled cluster community, where pair-natural orbital (PNO) based localization approaches can reduce the scaling of CCSD(T) to near-linear, with respectable accuracy,[51,102,103,292,293] and some benchmark datasets on transition metals have been recently produced using DLPNO-CCSD(T).[41,176] However, for statically correlated systems, or even otherwise, the accuracy of full CCSD(T) can come into question,[49] and alternate benchmark-level methods such as AFQMC are necessary to provide complementary or even more accurate results.[1,2,47] It is worth noting that a PNO-based ph-AFQMC approach using approximate PNOs from MP2 was initially attempted, but was found to be inaccurate upon projection in imaginary time, presumably due to the approximate nature of the PNO transformation. One advantage of directly compressing the HR-ERI tensor as described here is that it is not dependent on any lower level methods, and so can effectively reproduce full AFQMC values. Even so, the idea to directly solve for a low rank structure for the ERIs is not new to AFQMC, and is shared by tensor hypercontraction and the low-rank compression of the cholesky vectors in Ref. 280. The LO-AFQMC algorithm presented here stands out in utility by virtue of it's relative simplicity, high performance, and low prefactor, which depending on the degree of electronic localization within a given molecule can lead to better performance for systems of nearly all sizes.

Based on the timing results above for the metallocenes test cases, we can estimate the cost of generating AFQMC benchmark data. As an example of a potentially useful benchmark data set, consider the set of experimental redox potentials for octahedral transition metal complexes in solution assembled in Ref. 107. The number of heavy atoms in these systems ranges from 7 (e.g. $M(H_2O)_6$) to approximately 45; however, most of the molecules have 25 or fewer heavy atoms, and the data set could be restricted to those cases. The cost of running 50 such calculations at the TZ/QZ level would be on the order of 70,000 GPU hours on the Summit supercomputer. A gas phase comparison of DFT and AFQMC results would enable the development of an improved DFT functional for transition metal containing systems; using such a functional, the continuum model needed to compute redox potentials in solution could then be fit to the experimental data assembled in Ref. 107.

The ability to reproduce full ph-AFQMC for considerably larger molecules than have heretofore been amenable to AFQMC calculations represents a major step towards obtaining benchmark-quality simulations of large transition metal complexes (and other challenging systems) relevant to biology and materials science, such as the water splitting cluster of Photosystem II (PSII). Hundreds to thousands of calculations will be required to obtain relative energies of different metal (Mn) spin and charge states, and different oxygen protonation states, that are needed to precisely specify the intermediates in the Kok cycle, the catalytic process in which the energy of four photons is used to convert water into oxygen and hydrogen gas.[294] A minimal PSII model would contain approximately 50 heavy atoms (based on the smallest possible truncations of the relevant amino acids). The cost of a single calculation at the TZ/QZ level would be on the order of 10,000 GPU hours. While a cost of this magnitude would not be sufficient to evaluate every possible model for the S states of the Kok cycle, 50-100 selected (perhaps on the basis of DFT energetics) individual states could readily be investigated, enabling benchmark relative energies to be determined and compared with various DFT approaches and with experiment. Application to a system of this size would likely benefit from some of the improvements proposed above, such as screening distant pairs.

The above applications presume that suitable trial functions can be generated for the systems to be studied, and that a moderate number of determinants (in the 300-1000 range that we have used to date in our AFQMC modeling of metal-containing systems) can produce accurate results. For the redox potential benchmarks, this assumption is likely to be valid, although there could be surprises for particularly challenging individual cases. For the PSII model, which contains four

Mn atoms and one Ca, the requirements for a trial function are far from clear. If an ultralarge number of determinants are needed, it may be necessary to switch to alternate algorithms for better scaling with N_{det} (although, as pointed out above, the LO formulation described in this paper, or some alternate, will still be necessary to make such calculations tractable).

4.5 Appendices

4.5.1 MS Algorithm

Mahajan and Sharma (MS) have recently developed an algorithm for evaluating the two-body local energy for multideterminant trials which takes advantage of the rank ϵ excitation structure of CI expansions in a general way.[273,282] Here we reproduce key algorithmic details of this approach in our notation for clarity. The trial wavefunction is represented first as an expansion of excitation operators operating on the reference

$$|\Psi\rangle = \sum_{d}^{N_{det}} c_d \prod_{\mu_d}^{\epsilon_d} a^{\dagger}_{t_{\mu_d}} a_{p_{\mu_d}} |\psi_0\rangle, \tag{4.24}$$

where the operators $a^{\dagger}_{t_{\mu_d}} a_{p_{\mu_d}}$ excite a configuration from the occupied orbital p_μ into t_μ, and ϵ_d denotes the number of excitations for the determinant indexed by the label d.

One can choose to rearrange the local energy as

$$E = \frac{\langle \Phi_T|\hat{H}|\Phi\rangle}{\langle \Phi_T|\Phi\rangle} = \frac{\langle \Psi|\hat{H}|\Phi\rangle}{\langle \psi_0|\Phi\rangle} / \frac{\langle \Psi|\Phi\rangle}{\langle \psi_0|\Phi\rangle}, \tag{4.25}$$

$$E = \frac{\sum_d^{N_{det}} c_d \langle \psi_0| \prod_{\mu_d} a^{\dagger}_{p_{\mu_d}} a_{t_{\mu_d}} \hat{H}|\Phi\rangle}{\langle \psi_0|\Phi\rangle} / \frac{\sum_d^{N_{det}} c_d \langle \psi_0| \prod_{\mu_d} a^{\dagger}_{p_{\mu_d}} a_{t_{\mu_d}} |\Phi\rangle}{\langle \psi_0|\Phi\rangle}, \tag{4.26}$$

where the overlap ratios in the denominator can be computed with $O(N^2M+N_{det})$ cost if one works in the MO basis, and $O(NM^2+N_{det})$ cost if not.[273] Skipping the one body contribution, which can be calculated at a similar cost, we then rewrite the resulting two body energy contribution in terms of the reference CI determinant,

$$E_2 = \sum_{d}^{N_{det}} c_d \cdot \frac{1}{2} \sum_{pqrs}^{M} \sum_{\alpha}^{X} L^{\alpha}_{rp} L^{\alpha}_{sq} \frac{\langle \psi_0 | (\prod_{\mu_d} a^{\dagger}_{p_{\mu_d}} a_{t_{\mu_d}}) a^{\dagger}_{p} a^{\dagger}_{q} a_r a_s | \phi \rangle}{\langle \psi_0 | \phi \rangle}. \tag{4.27}$$

The generalized Wick's theorem enables the expansion of this term as a function of the reference Green's function. Dropping the determinant index d, we find

$$\frac{\langle \psi_0 | \prod_{\mu_d} a^{\dagger}_{p_{\mu_d}} a_{t_{\mu_d}} a^{\dagger}_{p} a^{\dagger}_{q} a_r a_s | \phi \rangle}{\langle \psi_0 | \phi \rangle} = \det \begin{pmatrix} G^{\{p,q\}}_{\{r,s\}} & \mathcal{G}^{\{p,q\}}_{\{t_\mu\}} \\ G^{\{p_\mu\}}_{\{r,s\}} & G^{\{p_\mu\}}_{\{t_\mu\}} \end{pmatrix} = \det \begin{pmatrix} G_{pr} & G_{ps} & \mathcal{G}_{p\{t_\mu\}} \\ G_{qr} & G_{qs} & \mathcal{G}_{q\{t_\mu\}} \\ G_{\{p_\mu\}r} & G_{\{p_\mu\}s} & G_{\{p_\mu\}\{t_\mu\}} \end{pmatrix}, \tag{4.28}$$

where $\mathbf{G}$ is the Green's function associated with the reference determinant, calculated at cost N^2M in the MO basis, and the sets $\{p_\mu\}$ and $\{t_\mu\}$ are the ϵ occupied and virtual excitation orbitals for the particular determinant. $\mathcal{G}_{ab}$ is defined as $\mathcal{G}_{ab} = G_{ab} - \delta_{ab}$. This expression can then be Laplace expanded across the first two columns, yielding the following equation for the determinant in Eq. 4.28,

$$\begin{aligned} det \begin{pmatrix} G^{\{p,q\}}_{\{r,s\}} & \mathcal{G}^{\{p,q\}}_{\{t_\mu\}} \\ G^{\{p_\mu\}}_{\{r,s\}} & G^{\{p_\mu\}}_{\{t_\mu\}} \end{pmatrix} &= [G_{pr}G_{sq} - G_{qr}G_{ps}] \det(G_{\{p_\mu\}\{t_\mu\}}) \\ &+ \sum_{\nu}^{\epsilon} (-1)^{\nu} [G_{ps}G_{p_\nu r} - G_{pr}G_{p_\nu s}] \det \begin{pmatrix} \mathcal{G}_{q\{t_\mu\}} \\ G_{\{p_{\mu \neq \nu}\}\{t_\mu\}} \end{pmatrix} \\ &+ \sum_{\nu,\nu',\lambda,\lambda'}^{\epsilon} (-1)^{\nu+\nu'+\lambda+\lambda'} G_{p_\nu r} G_{p_{\nu'} s} \mathcal{G}_{pt_\lambda} \mathcal{G}_{qt_{\lambda'}} \det(G_{\{t_{\mu \neq \{\nu,\nu'\}}\}\{p_{\mu \neq \{\lambda,\lambda'\}}\}}). \end{aligned} \tag{4.29}$$

Plugging Eq. 4.29 into Eq. 4.27 results in an equation that allows for a separation of sums over

Hamiltonian and determinant indices. For each determinant we have

$$
\begin{aligned}
E_2^d = E_2^0 \times \det(G_{\{p_\mu\}\{t_\mu\}}) + \sum_{\nu}^{\epsilon}(-1)^{\nu}\det\begin{pmatrix}[D_1]_{p_\nu\{t_\mu\}} \\ G_{\{p_{\mu\neq\nu}\}\{t_\mu\}}\end{pmatrix} \\
+\sum_{\alpha}^{X}\sum_{\nu,\nu',\lambda,\lambda'}^{\epsilon}(-1)^{\nu+\nu'+\lambda+\lambda'}[[D_2]^{\alpha}_{p_\nu t_\lambda}[D_2]^{\alpha}_{p_{\nu'}t_{\lambda'}} - [D_2]^{\alpha}_{p_\nu t_{\lambda'}}[D_2]^{\alpha}_{p_{\nu'}t_\lambda}] \times \det(G_{\{t_{\mu\neq\{\nu,\nu'\}}\}\{p_{\mu\neq\{\lambda,\lambda'\}}\}}),
\end{aligned}
\tag{4.30}
$$

where E_2^0, $[D_1]_{pt}$, and $[D_2]^{\alpha}_{pt}$ are expressed as

$$
\begin{aligned}
E_2^0 &= \sum_{\alpha}^{X}\sum_{pqrs} L^{\alpha}_{rp}L^{\alpha}_{sq}[G_{pr}G_{qs} - G_{qr}G_{ps}], \\
[D_1]_{p_\mu t_\mu} &= \sum_{\alpha}^{X}\sum_{pqrs} L^{\alpha}_{rp}L^{\alpha}_{sq}[G_{ps}G_{p_\mu s} - G_{pr}G_{p_\mu r}]\mathcal{G}_{qt_\mu}, \\
[D_2]^{\alpha}_{p_\mu t_\mu} &= \sum_{pr} L^{\alpha}_{rp}G_{p_\mu r}\mathcal{G}_{pt_\mu}.
\end{aligned}
\tag{4.31}
$$

These terms can be calculated at a cost of XNM^2, $XNAM$, and $XNAM$, respectively, where A is the size of the active space. Note that for $[D_2]^{\alpha}_{p_\mu t_\mu}$, we have a partial summation of Hamiltonian indices, leaving the last term of Eq. 4.30 to be calculated at a cost of $N_{det}X$. The resulting algorithm thus scales as $O(N_{det}S + S^4)$, where S is a general proxy for the system size. It is possible to precompute the sum over α as well, leading to an algorithm scaling as $O(N_{det} + S^5)$, but this becomes intractable for even medium system sizes.

4.5.2 LO-AFQMC in the MS Algorithm

The compression of the HR-ERI tensor allows for two significant sources of improvement in scaling versus Eqns. 4.29-4.31. Firstly, the sum over auxiliary fields is precomputed and folded into the integrals. Savings due to this are additionally possible if one uses the full HR-ERIs, but at a prohibitive memory cost. Upon localization, that memory cost becomes insignificant, and we can additionally replace a factor of M with $\langle M_{SVD}\rangle$. The most straightforward applications are to E_2^0

and $[D_1]_{p_\mu t_\mu}$, which take the form of sums similar to Eq. 4.22, but for the reference determinant only,

$$E_2^0 = \sum_{ij}^{N} \sum_{K}^{M_{SVD}} \sum_{rs}^{M} \bar{U}_{rK}^{[ij]} \bar{V}_{Ks}^{[ij]} [Q_{ir} Q_{js} - Q_{jr} Q_{is}], \tag{4.32}$$

$$[D_1]_{p_\mu t_\mu} = \sum_{ij}^{N_{cas}} \sum_{K}^{M_{SVD}} \sum_{rs}^{M} \bar{U}_{rK}^{[ij]} \bar{V}_{Ks}^{[ij]} [Q_{is} Q_{p_\mu r} - Q_{ir} Q_{p_\mu s}] \mathcal{G}_{j t_\mu}. \tag{4.33}$$

The largest scaling step is the contraction over full basis indices r and s, which scales as $O(N^2 M \langle M_{SVD} \rangle)$, reducing the scaling from quartic to cubic. For the last term of Eq. 4.30, it is most advantageous to perform the summation of all four indices at once,

$$[D_3]_{p_\nu p_{\nu'} t_\lambda t_{\lambda'}} = \sum_{ij}^{N_{cas}} \sum_{K}^{M_{SVD}} \sum_{rs}^{M} \bar{U}_{rK}^{[ij]} \bar{V}_{Ks}^{[ij]} Q_{p_\nu r} Q_{p_{\nu'} s} \mathcal{G}_{i t_\lambda} \mathcal{G}_{j t_{\lambda'}}. \tag{4.34}$$

Note that this is equivalent to $[D_3]_{p_\nu p_{\nu'} t_\lambda t_{\lambda'}} = \sum_\alpha^X [D_2]^\alpha_{p_\nu t_\lambda} [D_2]^\alpha_{p_{\nu'} t_{\lambda'}} - [D_2]^\alpha_{p_\nu t_{\lambda'}} [D_2]^\alpha_{p_{\nu'} t_\lambda}$ in Eq. 4.30, but using the localized HR-ERIs and changing the order of summation so that it is not necessary to explicitly form the set of $[D_2]^\alpha$. As all Green's functions in this Eq. are reference Green's functions, we can make use of the form due to being in the MO basis, and replace all general indices p and q with occupied indices i and j. The formation of this four index tensor of size $A_{occ}^2 A_{virt}^2$ now scales as $N^2 M A_{occ} \langle M_{SVD} \rangle$, where A_{occ} and A_{virt} are the number of occupied and virtual orbitals in the active space, respectively. E_2 can now be calculated as

$$\begin{aligned} E_2^d = {} & E_2^0 \det(G_{\{p_\mu\}\{t_\mu\}}) + \sum_{\nu}^{\epsilon} (-1)^\nu \times \det \begin{pmatrix} [D_1]_{p_\nu \{t_\mu\}} \\ G_{\{p_{\mu \neq \nu}\}\{t_\mu\}} \end{pmatrix} \\ & + \sum_{\nu, \nu', \lambda, \lambda'}^{\epsilon} (-1)^{\nu + \nu' + \lambda + \lambda'} [D_3]_{p_\nu p_{\nu'} t_\lambda t_{\lambda'}} \times \det(G_{\{t_{\mu \neq \{\nu, \nu'\}}\}\{p_{\mu \neq \{\lambda, \lambda'\}}\}}). \end{aligned} \tag{4.35}$$

This expression can then be used to calculate the energy at a cost of $O(N_{det})$, with an overall scaling for the energy evaluation of $O(N^2 M A_{occ} + N_{det})$.

4.5.3 Additional Details of GPU implementation

It is an unfortunate consequence of GPU architectures that operations involving lists of small matrices with inconsistent dimension are not easily parallelized. We therefore restrict ϵ to be equivalent to the maximum number of excitations, regardless of the number of excitations in that specific determinant, allowing these operations to be easily batched over N_{det} using, for example, `cublas<T>gemmBatched()` functions. While it is possible to use a similar restriction in M_{SVD}, this would result in significant loss of memory and computational efficiency. We have found it possible to perform the largest scaling operations over M_{SVD} for all $[ij]$ pairs concurrently by storing matrix lists contiguously in memory. An example of this difficulty can be seen in an intermediate step in the formation of the reference section of $\bar{Q}^{d,[ij]}_{R,i}$ (Eq. 31):

$$X^{[ij]}_{L,i} = \sum_{r}^{M} [\bar{U}^{[ij]}_{rK}] * [\Phi_{ra} A^{-i}_{ai}]. \tag{4.36}$$

Ideally, we would like to batch the sum over r for all [ij] pairs, all determinants, and all walkers at the same time. This is particularly difficult, as all [ij] pairs have a different dimension of K. Additionally, the vector on the right hand side depends on i, and so batching is only possible for all [ij] pairs with a common i. Batching is enabled by storing $\bar{U}^{[ij]}_{rK}$ in memory as a contiguous list of matrices of size M by $M^{[ij]}_{SVD}$, with the pairs [ij] in “i major” format:

$$\underbrace{\bar{U}^{[0,0]}, \bar{U}^{[0,1]}, \dots \bar{U}^{[0,N_{cas}]}}_{i=0}, \underbrace{\bar{U}^{[1,1]}, \bar{U}^{[1,2]}, \dots \bar{U}^{[1,N_{cas}]}}_{i=1}, \dots, \underbrace{\bar{U}^{[N_{cas},N_{cas}]}}_{i=N_{cas}}$$

The formation of a particular $X^{[ij]}_{L,i}$ then takes the form of a simple matrix-vector multiplication. Further parallelization is possible when the GPU is not saturated by taking advantage of CUDA streams.

We note that for the ‘j’ section, $\bar{Q}^{d,[ij]}_{L,j}$, we lose the contiguous nature of each i (aka all [ij] pairs with the same j are not contiguous). This is treated by a custom memory copying CUDA kernel, which takes all $\bar{U}^{[ij]}_{rK}$ with a given j and forms an intermediate tensor which is contiguous.

As GPU memory copying is generally fast, this leads to a significant increase in performance, with the formation of $X_{L,j}^{[ij]}$ being 21.7 times faster than performing the matrix vector calls in series for benzene in the cc-pVTZ-DK basis. There is a similar step in the active section (last term on the right hand side) of equation 31; the rest of the operations can be performed in parallel for all pairs [ij].

4.5.4 SMW algorithm for the formation of the Green's Function

Here we give details on the implementation and scaling of the formation of the Green's function for use in evaluating both the force bias and the 1-body energy. In contrast to the 2-body energy, we can perform the sum over determinants in the formation of the Green's function,

$$E_1 = \sum_{pr}^{M} \left[\sum_{d}^{N_{det}} c_d G_{pr}^d \right] K_{pr} = \sum_{pr}^{M} \bar{G}_{pr} K_{pr} \tag{4.37}$$

where K_{pr} are the 1-body integrals (or in the case of the force bias, K_{pr} would be a Cholesky vector L_{pr}^{α}). The summed Green's function $\bar{G}_{pr} = \sum_{d}^{N_{det}} c_d G_{pr}^d$ is in practice expanded via SMW along the lines of the Q matrix written out in eq. 4.20,

$$\bar{G} = \sum_{d}^{N_{det}} Q^d c_d * \Phi_T^{d,\dagger} = \sum_{d}^{N_{det}} c_d \left[\Phi A^{-1} \Phi_T^{d,\dagger} - \Phi A^{-1} U^d (I + V^{T,d} A^{-1} U^d)^{-1} V^{T,d} A^{-1} \Phi_T^{d,\dagger} \right]. \tag{4.38}$$

Since we can now include the sum over determinants in this expression (as was not the case when forming the intermediate Q matrices for the 2-body energy), we can rewrite this as

$$\bar{G} = \sum_{d}^{N_{det}} c_d \left[\Phi A^{-1} \Phi_T^{d,\dagger} \right] - \sum_{d}^{N_{det}} c_d \left[\Phi A^{-1} U^d (I + V^{T,d} A^{-1} U^d)^{-1} V^{T,d} A^{-1} \Phi_T^{d,\dagger} \right], \tag{4.39}$$

$$= \Phi A^{-1} \sum_{d}^{N_{det}} c_d \left[\Phi_T^{d,\dagger} \right] - \Phi A^{-1} \sum_{d}^{N_{det}} c_d \left[U^d (I + V^{T,d} A^{-1} U^d)^{-1} V^{T,d} A^{-1} \Phi_T^{d,\dagger} \right] \tag{4.40}$$

If propagating in the MO basis, $\Phi_T^{d,\dagger}$ has a form of identity plus some permutation, rendering all operations involving it trivial memory transfers, and thus the reference term (first term of right hand side of 4.40) scales as $O(MN^2)$. In Table 4.6 we outline key steps and scaling involved in evaluating the rightmost term, the correction term.

Intermediate	Operation	Memory scaling	Computational scaling
$C_1 = V^{T,d} A^{-1}$	$V^{T,d} \times A^{-1}$	$N_{det} N \epsilon$	$\min(N_{det} N^2 \epsilon, N^2 N^a_{occ} N^a_{unocc})$
$C_2 = (I + V^{T,d} A^{-1} U^d)^{-1}$	Inverse	$N_{det} \epsilon^2$	$N_{det} \epsilon^3$
$C_3 = (I + V^{T,d} A^{-1} U^d)^{-1} V^{T,d} A^{-1}$	$C_2 \times C_1$	$N_{det} N \epsilon$	$N_{det} N \epsilon^2$

Table 4.6: List of key intermediates in the formation of the Green's function using SMW, including the formal memory and computational scaling of forming them. N^a_{occ} and N^a_{unocc} refer to the number of occupied orbitals in the active space and number of unoccupied orbitals in the active space, respectively.

Note that the ϵ by N matrices U^d and $V^{T,d}$ have structures that can lead to increased savings. If N_{inact} is the number of inactive occupied orbitals, the first N_{inact} rows of U^d are zero, whereas the rest are either identity or a permutation matrix corresponding to the excitation structure of the determinant. Thus, the operation $U^d \times C_3$ would involve taking the last $N - N_{inact}$ rows of C_3 and performing some memory transfers, similar to how operations involving the trial determinants reduce to memory transfers when in the MO basis. $V^{T,d}$, while being a dense matrix, consists of a series of rows each corresponding to a single excitation from orbital a to b, $c_b^\dagger c_a$. While there may be many determinants in the CI expansion which have a given excitation, there are only $N^a_{occ} N^a_{unocc}$ possible single excitations, and one must only compute the rows of C_1 for each excitation once, maximally scaling as $N^2 N^a_{occ} N^a_{unocc}$ in the case where all possible excitations are present in the CI expansion. Including all steps outlined here, the formal scaling of the formation of the Green's function in the MO basis is $O(N^2(M + N^a_{occ} N^a_{unocc}) + N_{det} N)$. For CAS trials, which become intractable over about 18 orbitals, $N^a_{occ} N^a_{unocc}$ is typically much smaller than the basis size M. Once the Green's function is formed, the 1-body energy can be evaluated at cost M^2, and the force bias with cost XM^2. We note that it is likely possible to extend the savings due to compression to the evaluation of the force bias, but as this would neither reduce the formal scaling of the overall AFQMC nor reduce the memory overhead (as one still requires the full cholesky vectors to compute the propagation matrix $\sum_{\alpha}^{X} x_{\alpha} L^{\alpha}$ at cost XM^2), we do not explore this further.

4.5.5 LO error cancellation through imaginary time

Plotting LO errors over multiple time steps is made difficult by population control, which can be sensitive to small deviations in the energy (and thus weight) and mask the error due to the LO approximation with the intrinsic variance of AFQMC. We thus propagated 20 independent walkers for 100 Ha^{-1} (Fig. 4.9), evaluating the mean and standard deviation at every step for C_8H_8, the platonic hydrocarbon which exhibited the greatest difference in error between the two times steps in Fig. 4.7. The variance of the LO error over imaginary time then leads to the slight increase in statistical error mentioned in the main text. Note that the error does not increase dramatically as a function of imaginary time.

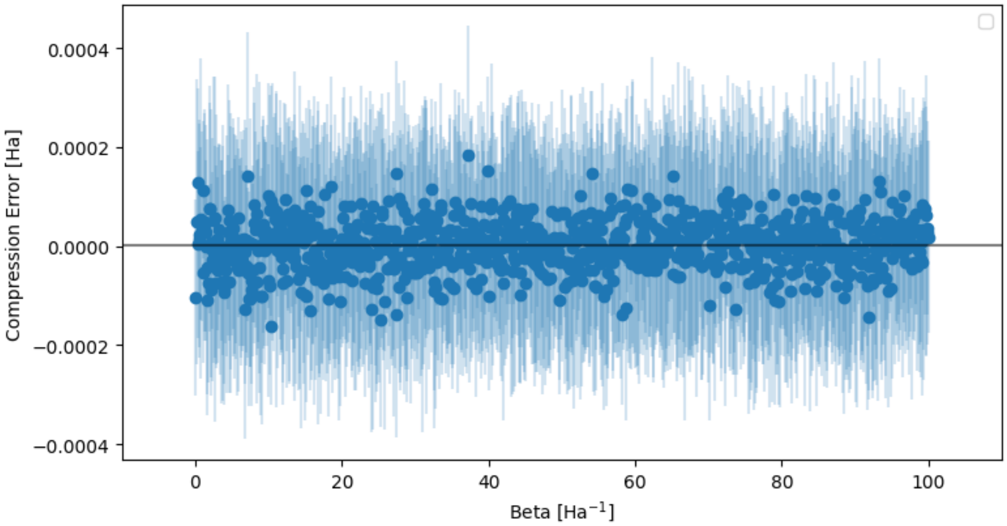


Figure 4.9: Plot of the mean and statistical error of the LO error of a set of 20 independent walkers for the system C_8H_8. The error exhibits a variance which contributes minimally to the statistical error, and does not grow significantly in imaginary time. The mean error in this case was 0.002 mHa, with a standard deviation of 0.2 mHa, representing a favorable cancellation of error over imaginary time.

Chapter 5: The Design of New Practical Constraints in Auxiliary-Field Quantum Monte Carlo

The following chapter was published as: J. L. Weber *et al.*, *arXiv preprint arXiv:2306.09207*, 2023.

5.1 Introduction

Projector Quantum Monte Carlo (PQMC) algorithms formally provide a polynomial-scaling route to chemical accuracy for a variety of molecular and solid state systems. However, as the exact projection of an initial guess onto the ground state necessarily encounters a sign or phase problem for general fermionic systems, the key to the extended success of PQMC has been the formulation of constraints which effectively remove the sign problem while retaining a high level of accuracy. Phaseless Auxiliary-Field Quantum Monte Carlo (ph-AFQMC) has provided a prominent example of such a constrained method in the basis of non-orthogonal Slater determinants.[5] It does so by exactly restricting the propagation of walkers to a single gauge in the complex plane via a "force bias," defined in relation to a trial wave function. This cancels the accumulation of phase from the propagator to first order in the imaginary timestep, $\Delta\tau$. In practice, one must additionally remove the origin from the calculation to avoid the sign problem. There are many possible ways to do so, as originally described by Zhang *et. al.* in the paper that first introduced the phaseless approximation.[274] Zhang *et. al.* found that the most accurate and stable algorithms for model systems were obtained when using a cosine projection to scale the weight of each walker according to the stepwise growth of phase, namely

$$W_w = |W_w| \times \max(\cos(\Delta\theta), 0), \tag{5.1}$$

where W_w is the weight of the walker indexed by w, and $\Delta\theta$ refers to the stepwise change of phase with regard to the trial wave function, $\theta = \arg(\langle \Phi_T | \Phi_w \rangle)$. In the case of an exact trial wave function this procedure provides exact results, however for an inexact trial wave function, this results in a bias, commonly referred to as the “phaseless bias."

The behavior of the phaseless bias is difficult to quantify, and thus the vast majority of studies to date focus on benchmarking ph-AFQMC with different trials in order to devise strategies to obtain increasingly accurate benchmark results on difficult systems at minimal cost. Such approaches have been largely successful, with calculations using multi-reference trials performing well on many transition metal complexes for which single reference methods such as coupled cluster can fail.[**neugebauer2023toward**, 1,2,48,72]

The most popular strategy to reduce the cost of AFQMC, and thus render it scalable to larger, more interesting systems, is to provide more affordable ways to converge the trial wave function, e.g. by using selected CI wave functions[273,278] or by updating a single determinant trial until self consistency is obtained.[296,297] Recently, coupled cluster (CC) trials were used with the assistance of a quantum computer[298]. Additionally, one can devise systematically convergable approximations[4,236] or sampling strategies[56] that reduce the cost of running AFQMC with a fixed trial, which can, in some cases, take advantage of an enhanced cancellation of error for energy differences. While these approaches are useful in practice, there still exist many cases where the performance of ph-AFQMC is worse than expected, given the nature of the problem and the quality of the trial. For example, recent work by Lee *et. al.* pointed out the failure of ph-AFQMC with an unrestricted Hartree Fock (UHF) trial to adequately describe open shell atoms.[299] Indeed, we have found that even given a small CASSCF trial wave function describing the valence S and P orbitals (Xe4o), ph-AFQMC does not attain chemical accuracy for any of the open shell main group atoms (Fig. 5.1). While numerically exact energies may be obtained with a converged trial or via non-constrained (free projection) AFQMC, these results highlight the limitations of exclusively using previous benchmarks to determine when a trial is sufficient for ph-AFQMC. Additionally, many single reference transition metal problems still appear to require a large number of determinants in the trial wave function to adequately converge the total energy for ph-AFQMC, which can result in significant increases in cost for each ph-AFQMC calculation. The prospect of modifying the phaseless constraint in the hopes of obtaining less biased results given a fixed trial thus remains very enticing.

Two decades after the invention of ph-AFQMC we now have sufficient computational power to rigorously test alternatives to the cosine projection on benchmark sets, and efforts to do so have already emerged in the literature. For example, Sukurma et. al.[300] recently tested a reformulation of the cosine projection in which they delay projection until evaluating observables, instead removing the origin by killing walkers if their total phase grows beyond $\frac{\pi}{2}$. They found qualitatively similar yet statistically distinct results to ph-AFQMC on a small benchmark set of small main group molecules.

Another approach to systematically improve the accuracy of ph-AFQMC is to interface it in some way with the exact non-constrained free projection AFQMC (fp-AFQMC), known as constraint release.[273,301] For ph-AFQMC, it is necessary to reformulate the importance sampling procedure upon releasing the constraint, as left in the constrained form the force bias effectively removes walkers which violate the phaseless constraint by reducing the weight to zero even in the absence of explicitly removing them upon crossing the origin. This procedure is typically unstable, and one is thus restricted to fully removing the constraint at once, which necessarily reduces the applicability of constraint release techniques in AFQMC.

Recently, Xiao *et. al.* introduced a framework which bypasses this instability by initiating a non-constrained Markov chain Monte Carlo calculation with a population of walkers from a constrained ph-AFQMC random walk which was subsequently back-propagated.[302] With sufficiently long imaginary times, as estimated by the variance in energy decreasing beyond a certain threshold, they were able to obtain results starting from a ph-AFQMC calculation which are exact within error bars. Moreover, by freezing a portion of the back-propagated trajectory within the Markov chain simulation, they were able to lessen the severity of the observed sign problem at the cost of retaining some bias.

In constrast to this study, we choose to forego the goal of eliminating the phase entirely, and report on a new class of "phaseful" constraints. We do not make use of importance sampling with the force bias, and we allow arbitrary phase accumulation up to $|\theta| = \alpha\pi$. This effectively breaks the symmetry of the random walk in the complex plane while simultaneously removing the origin, allowing for a retention of signal with a bounded noise amplitude. For the open shell atoms, we find that this constraint is stable up to $\alpha = 1.0$, contrary to the common expectation of the emergence of an exponential phase problem for $\alpha > 0.5$ (see Fig. 5.1). For $\alpha > 1.0$ our approach reduces to free projection. This formalism introduced here enables a novel means to release the constraint and thus allows for a systematic partial sampling of the unbiased electronic structure problem.

This paper is organized as follows: We first provide a brief overview of the details of the phaseless AFQMC algorithm and introduce our new approach in this context. We then present results for the open shell atoms when scanning α, showing stability up to $\alpha = 1.0$ and a consistent improvement in accuracy versus ph-AFQMC. We then explore the accuracy of the $\alpha = 1.0$ constraint, which we call "linecut" (lc-) AFQMC, for a series of model molecular systems for which exact total energies are available, including H_4 dissociation, N_2 dissociation, C_2, and benzene. We test for size consistency using an increasing number of N_2 molecules, and observe similar behavior to ph-AFQMC when using a large timestep. We additionally show consistent stability as the system size increases, reproducing a phaseless (exact) calculation for $Ni(Cp)_2$ in the cc-pVTZ basis within statistics. Finally, we present an algorithm for the partial release of the linecut constraint (lcR-AFQMC), asymptotically reproducing the unbiased free projection. Partially released calculations are shown to allow for the convergence of AFQMC calculations for difficult systems given a fixed trial. We conclude with an outlook on the practical use of our approach within the context of AFQMC calculations.

5.2 Theory

Here we give a brief overview of the formalism of phaseless AFQMC, followed by a description of the proposed constraint. For a more in depth review of ph-AFQMC, we suggest some recent reviews.[5,73,111] In PQMC, initial states $|\Phi_i\rangle$ which are non-orthogonal with respect to the exact ground state will recover the exact ground state $|\Phi_0\rangle$ upon projection in imaginary time τ,

$$\lim_{\tau \to \infty} e^{-\tau \hat{H}} |\Phi_i\rangle = |\Phi_0\rangle, \tag{5.2}$$

where $\hat{H}$ denotes the electronic Hamiltonian, $\hat{H} = \hat{H}_1 + \hat{H}_2 = \sum_{pq} h_{pq} c_p^\dagger c_q + \frac{1}{2} \sum_{pqrs} \langle pq|rs\rangle c_p^\dagger c_q^\dagger c_s c_r$. Time evolution is formulated in terms of a series of finite steps in imaginary time, $e^{-\tau \hat{H}} = (e^{-\Delta\tau \hat{H}})^n$, each of which are then separated into one- and two- body terms via a symmetric Suzuki-Trotter decomposition,

$$e^{-\Delta\tau(\hat{H}_1 + \hat{H}_2)} \simeq e^{-\frac{\Delta\tau \hat{H}_1}{2}} e^{-\frac{\Delta\tau \hat{H}_2}{2}} e^{-\frac{\Delta\tau \hat{H}_1}{2}} + O(\Delta\tau^3). \tag{5.3}$$

If we write the electronic two-body operator as a sum of one-body operators squared, $\langle pq|rs\rangle = \sum_\alpha L_{pr,\alpha} L_{qs,\alpha}$ via exact diagonalization or approximate methods such as density fitting or modified Cholesky decomposition,[146] we can then use the Hubbard-Stratonovich identity to convert the two-body operators into a multi-dimensional integral over a set of fluctuating “auxiliary-fields” x_α,

$$e^{-\frac{\Delta\tau}{2}(\sum_\alpha L_\alpha^2)} = \prod_\alpha \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_\alpha^2}{2}} e^{\sqrt{-\Delta\tau}\, x_\alpha L_\alpha} dx_\alpha + O(\Delta\tau^2). \tag{5.4}$$

It is this multi-dimensional integral on which we perform Monte Carlo sampling

$$|\Phi(\tau + \Delta\tau)\rangle = \prod_\alpha \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{x_\alpha^2}{2}} e^{\sqrt{-\Delta\tau}\, x_\alpha L_\alpha} dx_\alpha |\Phi(\tau)\rangle = \int d\mathbf{x} P(\mathbf{x}) \hat{B}(\mathbf{x}) |\Phi(\tau)\rangle, \tag{5.5}$$

where $\mathbf{x}$ is the vector of auxiliary fields. In AFQMC, this simulation can be formulated as a branching, open-ended ensemble of random walkers w over the manifold of Slater determinants, each represented by a single Slater determinant $|\Phi_{\tau,w}\rangle$ and corresponding weight $W_{\tau,w}$. Each walker is propagated forward at a given time τ by $\hat{B}(\mathbf{x}_{\tau,w})$, with the space of auxiliary fields $\mathbf{x}$ being sampled from the Gaussian probability defined in Eq. 5.4. As $\hat{B}(\mathbf{x}_{\tau,w})$ is represented by purely one body operators, propagation respects the Thouless theorem and produces another Slater determinant, thereby maintaining the anti-symmetry of the wave function.[284]

To reduce the necessary sampling, the weights are updated according to the ratio of the new overlap with the trial to the previous overlap

$$|\Phi_{\tau+\Delta\tau,w}\rangle = \hat{B}(\mathbf{x}_{\tau,w})|\Phi_{\tau,w}\rangle, \tag{5.6}$$

$$W_{\tau+\Delta\tau,w}e^{i\theta_{\tau+\Delta\tau,w}} = \frac{\langle\Phi_T|\Phi_{\tau+\Delta\tau,w}\rangle}{\langle\Phi_T|\Phi_{\tau,w}\rangle}W_{\tau,w}e^{i\theta_{\tau,w}}. \tag{5.7}$$

The AFQMC representation of the wave function in this importance sampling framework is thus given by

$$|\Phi\rangle = \sum_w W_w \frac{|\Phi_w\rangle}{\langle\Phi_T|\Phi_w\rangle}. \tag{5.8}$$

So far, the algorithm presented is formally exact and is referred to as free projection (fp-) AFQMC; however, the un-constrained fermionic phase problem leads to an exponential decrease in the signal-to-noise ratio with respect to imaginary time duration as the walkers are propagated. In the phaseless constraint, we exactly shift the auxiliary fields by a complex "force bias",[54,285]

$$\prod_\alpha \int_{-\infty}^{\infty} \frac{1}{2\pi}e^{-\frac{x_\alpha^2}{2}} e^{\sqrt{-\Delta\tau}x_\alpha L_\alpha} dx_\alpha |\Phi(\tau)\rangle = \prod_\alpha \int_{-\infty}^{\infty} \frac{1}{2\pi}e^{-\frac{(x_\alpha-\bar{x}_\alpha)^2}{2}} e^{\sqrt{-\Delta\tau}(x_\alpha-\bar{x}_\alpha)L_\alpha} dx_\alpha |\Phi(\tau)\rangle, \tag{5.9}$$

where $\bar{x}_\alpha$ is chosen to cancel the accumulation of phase with respect to the trial to first order,

$$\bar{x}_\alpha = \frac{\langle\Phi_T|\hat{L}_\alpha|\Phi_w\rangle}{\langle\Phi_T|\Phi_{\tau,w}\rangle}. \tag{5.10}$$

This modifies the weight W_w by a factor $e^{x_\alpha\bar{x}_\alpha-\frac{\bar{x}_\alpha^2}{2}}$. In practice we must additionally remove the origin with respect to the trial, which is accomplished via the cosine projection, which projects the weight back onto the real axis after each step and removes walkers which become negative, namely

$$W_w = W_w \times \max(0, \cos(\Delta\theta)). \tag{5.11}$$

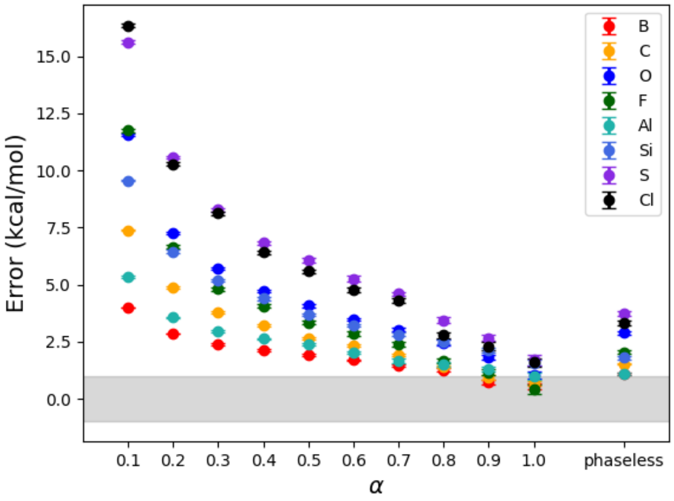


Figure 5.1: Systematic convergence with α for the open shell atoms of AFQMC with the constraint as specified in Eq. 5.12, versus ph-AFQMC. All calculations use CASSCF trials including the valence S and P orbitals in the active space. The grey bar indicates chemical accuracy, or 1 kcal/mol.

In our new constraint, we do not make use of the force bias, instead beginning from the free projection projector defined in Eq. 5.5. We update the weights as in Eq. 5.7, keeping track of the overall phase accumulated. As soon as the magnitude of the phase of a walker exceeds $\alpha\pi$, the walker weight is updated to zero and thus annihilated,

$$W_w = \begin{cases} W_w, & \text{if } |\theta_w| < \alpha\pi \\ 0, & \text{if } |\theta_w| > \alpha\pi \end{cases} . \tag{5.12}$$

This constraint with $\alpha = 1.0$ is what we henceforth refer to as linecut (lc-) AFQMC.

5.3 Results

5.3.1 Open Shell Atoms

Recently, Lee *et. al.* reported substantial biases for the set of open shell atoms using ph-AFQMC with a UHF trial,[299] making them ideal small candidates for the testing of new constraints. We thus use these atoms with small RCAS trials to test the performance and stability of the new approach with increasing α. We find the observed bias decreases monotonically as a function of α for all atoms (Fig. 5.1). As expected, the variance increases as a function of α, with $\alpha \leq 0.5$ exhibiting variance on par with the phaseless constraint, albeit producing much less accurate results due to the removal of the force bias, which thus allows for an accumulation of phase due to propagation. The most interesting cases are those with α above 0.5, where walkers begin to accrue a partial sign. In these cases, while the variance does increase, it does not exponentially increase, as one would expect for a full sign problem (and, indeed, as we see with $\alpha > 1$). Instead, the variance is limited to $\simeq 2 - 4$ times that found with the phaseless constraint, while the accuracy continues to improve. In the most extreme case, $\alpha = 1$, we see consistently improved results versus ph-AFQMC, bordering on and in some cases obtaining chemical accuracy. For the remainder of the study, we focus on this case, which we refer to as "linecut" (lc-) AFQMC.

5.3.2 The C_2 Dimer

Recently we reported the convergence of ph-AFQMC for the C_2 dimer total energy with respect to the number of determinants used in a small selected CI trial.[shee2022potentially] To test for convergence with respect to the quality of the trial, we perform the same calculations using lc-AFQMC. Initially we observe worse performance from lc-AFQMC when using up to 10 of the highest weighted determinants in a SCI wave function, but lc-AFQMC quickly overtakes that of ph-AFQMC, reaching chemical accuracy at $\simeq$100 determinants, as opposed to the 220 determinants required for ph-AFQMC (Fig. 5.2) to reach a comparable level of accuracy. Both methods converge to the exact ground state energy within statistical error bars as the number of determinants increase. It is interesting that with the lower quality (as measured by a small number of determinants) trials, ph-AFQMC outperforms lc-AFQMC, highlighting that although lc-AFQMC converges to chemical accuracy at a faster rate, the behavior of the observed bias is distinct and in some cases can produce less accurate results than standard ph-AFQMC.

The stability of lc-AFQMC for values of α up to $\alpha = 1$ is somewhat surprising. To shed liught on this behavior, we investigate the population of walkers throughout the simulation for the

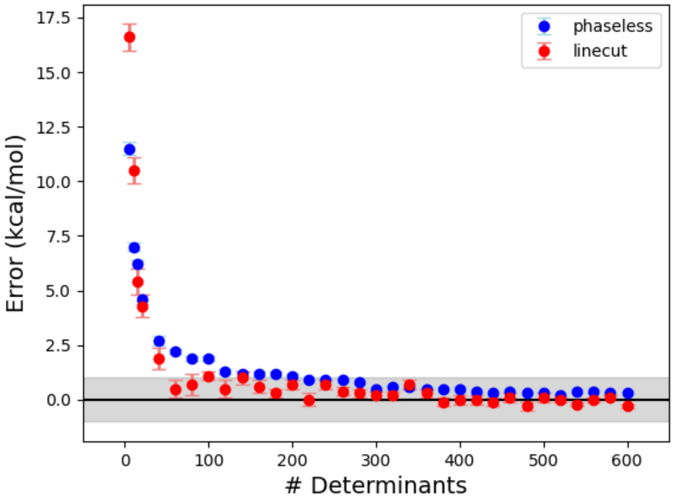


Figure 5.2: Systematic convergence of lc- and ph- AFQMC for C_2 with an increasing number of determinants retained in a heat bath CI trial (active space of 8e16o). In each case the determinants of a specified number with the highest weight in the wave function are kept.

case of the carbon dimer. Upon equilibration, we find that the removal of walkers as they cross the negative real axis (i.e. going from $\theta = \pm(\pi - \epsilon)$ to $\theta = \pm(\pi + \epsilon)$, Fig. 5.3), along with the use of a finite time step, results in a smooth population density reduction which goes to zero on the negative real axis (see Fig. 5.4-5.5). Allowing phase accumulation for $|\theta| > \frac{\pi}{2}$ does result in phase cancellation, which, along with the lack of importance sampling, is the cause of the observed increase in variance. However, breaking the symmetry of the random walk in the complex plane enables the signal to be measured above the noise. The largest signal emanates from walkers with overlap closest to one with respect to the trial. As there is no phase cancellation at this point, a more extreme dependence on the quality of the trial can result. Crucially, the lc-AFQMC removes the origin from the calculation and thus removes the sign problem even while allowing phase accumulation. These calculations show no increase in noise as a function of τ once the phase distribution has equilibrated, and have been run stably for upwards of 1200 Ha^{-1}.

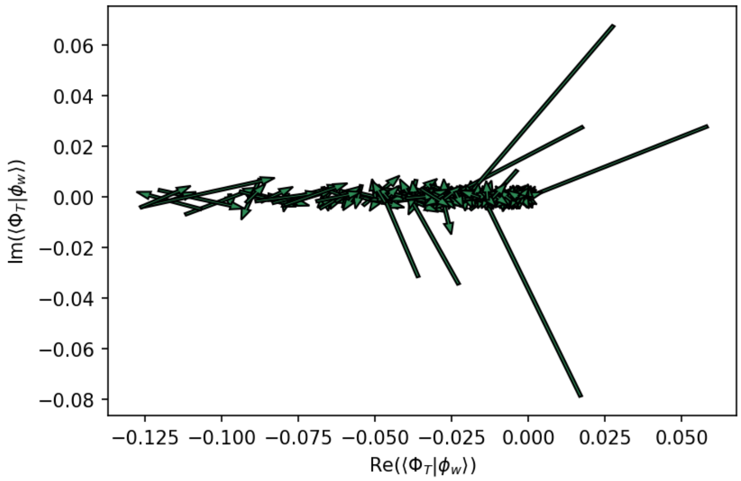


Figure 5.3: A sample of walker steps from a lc-AFQMC trajectory which cross the negative real axis, resulting in the elimination of the walker by setting the weight to zero. The tails of the arrows represent the overlap of the walker with the trial before the propagation step, while the head represents the overlap after. lc-AFQMC is stable due to the annihilation of walkers whose overlaps with respect to the trial wave function cross the negative real axis, including the origin, from either side.

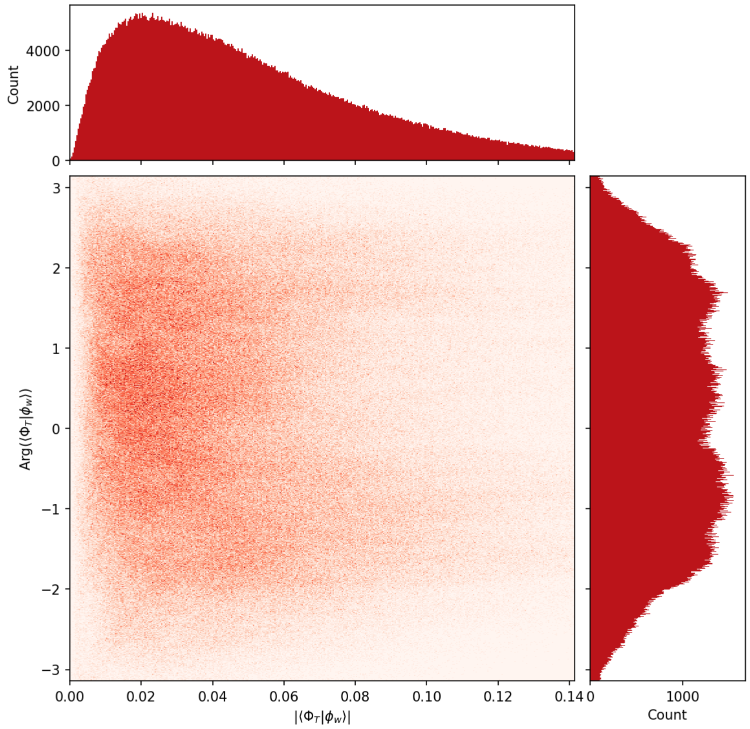


Figure 5.4: Distribution of overlaps in the complex plane with respect to the trial wave function for 256 walkers from 100 to 120 Ha^{-1} in a lc-AFQMC calculation. Note the effect of the linecut on the space surrounding the negative real axis.

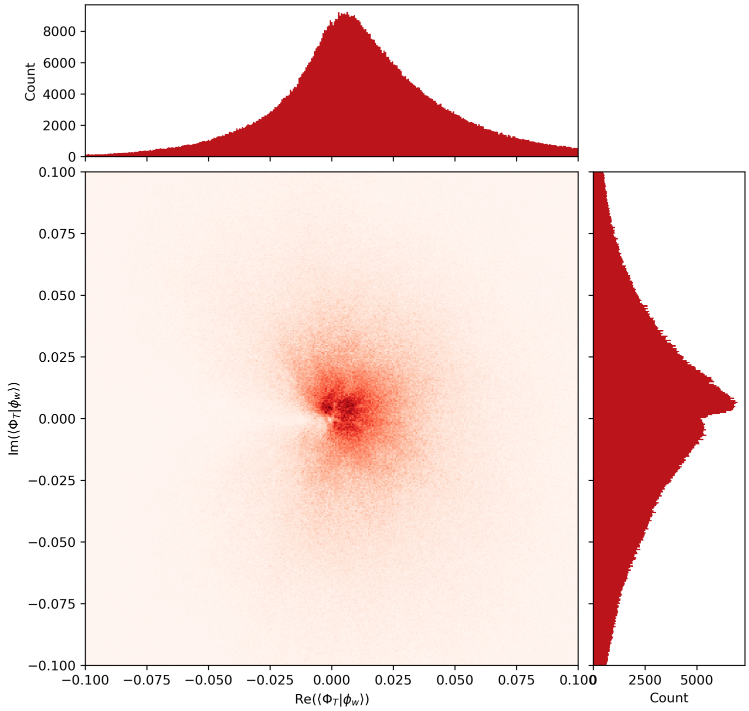


Figure 5.5: Distribution of the norm and phase of overlaps with respect to the trial wave function for 256 walkers from 100 to 120 Ha^{-1} in a lc-AFQMC calculation.

5.3.3 Bond Breaking Model Systems

Two widely studied model systems for strong correlation are the dissociation of the H_4 and N_2 molecules, for which ph-AFQMC results have been recently reported.[299] Following this study, we investigate the performance of lc-AFQMC versus ph-AFQMC for these systems using UHF trials. For H_4 in the STO3G basis, which is thought to accentuate the effects of strong correlation, lc-AFQMC obtains chemical accuracy for all bond distances, including obtaining the correct dissociation limit, albeit with some non-monotonic noise fluctuations as a function of R (Fig. 5.6). This behavior is not maintained, however, when increasing the basis set size to cc-pVQZ, where lc-AFQMC performs poorly, especially in the regions of strong correlation surrounding the equal distance points (Fig. 5.7). Interestingly, the equal distance point at $R = 1.23$ is one of the only cases where lc-AFQMC performs better than ph-AFQMC, and the errors of the two constraints oppose each other. This may suggest that the two constraints complement each other, with lc-AFQMC performing well for systems dominated by static correlation (as in the equal distance point) while

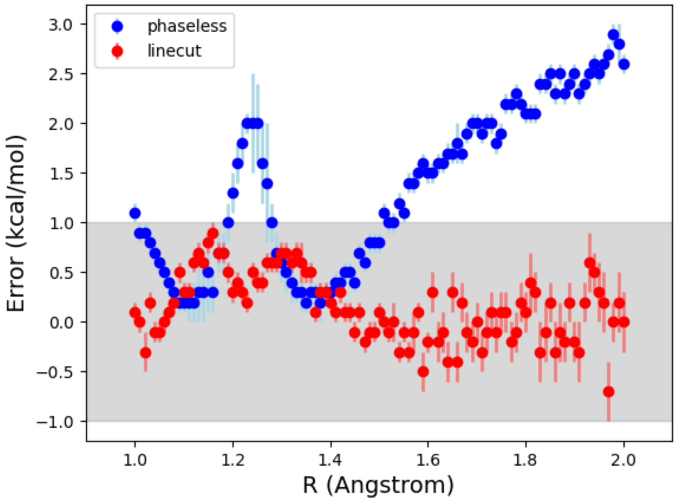


Figure 5.6: Error (kcal/mol) of both phaseless and linecut AFQMC for the dissociation of the H_4 molecule in the STO3G basis using a UHF trial. The grey bar indicates chemical accuracy, or < 1 kcal/mol error.

ph-AFQMC performs well for systems dominated by dynamic correlation, although it is impossible to generalize such observations with this limited dataset. Neither ph-AFQMC nor lc-AFQMC converges to the correct dissociation limit. To test if the linecut bias is remedied in the larger basis by the use of more accurate trial functions, we investigate H_4 in the cc-pVQZ basis at $R = 1.13$, where we had observed the maximum error of lc-AFQMC, using a 4e4o RCAS trial. Using only 2 determinants, the lc-AFQMC bias is reduced to 3.0 ± 0.9 kcal/mol versus 4.0 ± 0.3 kcal/mol for ph-AFQMC. For N_2, lc-AFQMC provides a less dramatic improvement over ph-AFQMC in the minimal basis, but is more accurate for nearly all bond lengths (Fig. 5.8).

To showcase the stability of the linecut constraint in imaginary time, we provide a plot of the N_2 trajectories for both ph- and lc- AFQMC at $R = 1.5$ using the UHF trial; both are stable up to 1200 Ha^{-1} and exhibit similar equilibration times. However lc-AFQMC converges to within statistical error of the exact FCI value (0.3 ± 0.4 kcal/mol off), whereas ph-AFQMC is biased by 1.4 ± 0.1 kcal/mol (Fig. 5.9).

Benchmark total energies for benzene in the cc-pVDZ basis are often additionally used to evaluate high level electronic structure methods for performance in systems dominated by dynamic

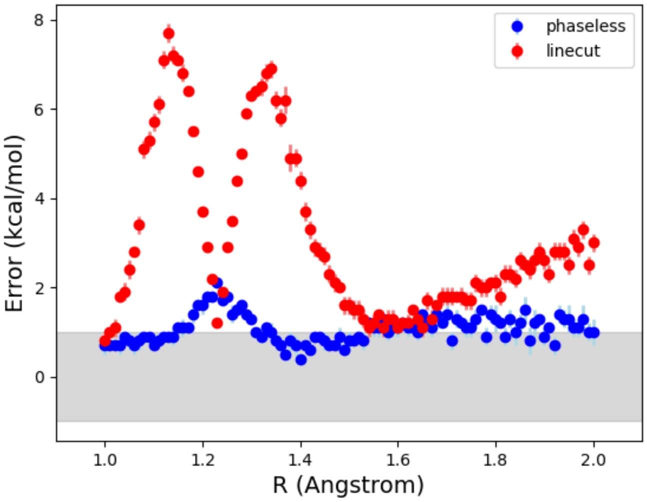


Figure 5.7: Error (kcal/mol) of both phaseless and linecut AFQMC for the dissociation of the H_4 molecule in the cc-pVQZ basis using a UHF trial. The grey bar indicates chemical accuracy, or < 1 kcal/mol error.

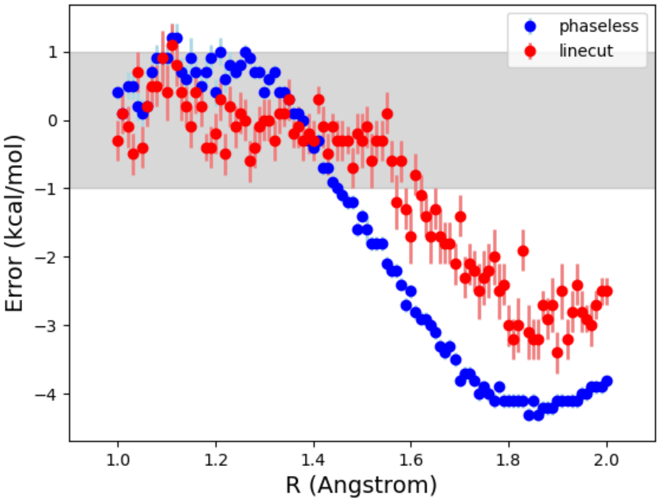


Figure 5.8: Error (kcal/mol) of both phaseless and linecut AFQMC for the dissociation of the N_2 molecule in the STO3G basis using a UHF trial. The grey bar indicates chemical accuracy, or < 1 kcal/mol error.

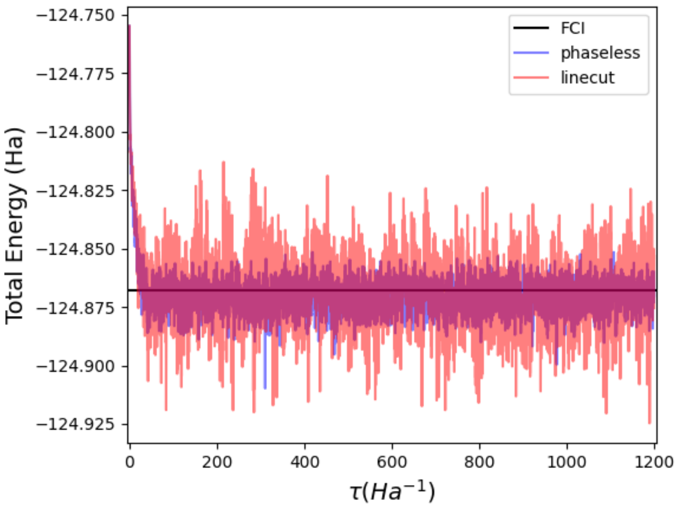


Figure 5.9: Trajectories of ph-AFQMC and lc-AFQMC using the exact trial for N_2 at 1.5 . ph-AFQMC is in blue, with lc-AFQMC in a semi-transparent red overlaid on top.

correlation.[303] Here, we ran lc-AFQMC with a single determinant trial (RHF), and obtain a total energy of $-862.1(1.0)$ Ha, within statistical error of the accepted exact energy (-863), significantly improving upon the RHF/ph-AFQMC value of $-866.1(3)$ reported in Ref. 275.

To test whether the linecut constraint stability decreases as a function of system size (and thus a more severe sign problem), we revisit a calculation from a previous study for which ph-AFQMC obtains chemical accuracy versus experiment, namely the case of Nickelocene in the cc-pVTZ-DK basis using an RCAS trial.[2] We again observe a $\simeq$4x increase in variance versus phaseless that nevertheless remains stable over the course of the simulation. Here, lc-AFQMC ($-2812.432(4)$ Ha) agrees with ph-AFQMC in total energy to within statistical error ($-2812.429(1)$ Ha).

5.3.4 Linecut Release

Due to the lack of importance sampling in the current implementation of lc-AFQMC, it is straightforward to implement a systematic release of the constraint to the free projection limit by relaxing the boundary conditions. In this protocol, walkers are instead tagged to be annihilated in X steps upon violation of the constraint, allowing a partial sampling of the region surrounding the origin. At $X = 0$, this is fully constrained lc-AFQMC, whereas at $X = \infty$ the walkers are freely

X	Error (kcal/mol)	st err
0	8.4	0.6
20	4.0	0.6
40	2.6	0.8
60	0.7	1.5

Table 5.1: Systematic release of the linecut constraint for C_2 using a trial with the ten highest weight CI determinants in the wave function. All calculations use 1920 walkers and propagate for 300 Ha^{-1}.

projected and a full sign problem is encountered. For intermediate X, one can partially remove the bias while damping the resulting noise from the sign problem. We have demonstrated this with C_2 using a 10-determinant trial (Table 5.1). This procedure is similar in spirit to release-node diffusion Monte Carlo.[304]

Whereas partially releasing the linecut constraint over long trajectories does remove the bias while still retaining a reasonable signal to noise ratio for C_2, we expect that releasing the constraint will be more difficult in more strongly correlated systems where the sign problem is more severe. As an example of such a case, we test the FeO dimer, which requires $\mathcal{O}(10^5)$ or more determinants in a full active space selected CI trial for ph-AFQMC to converge to chemical accuracy. Testing fully constrained lc-AFQMC, the convergence to the exact result is significantly slower than ph-AFQMC, similar to the trials with the lowest number of determinants for C_2 (Table 5.2). However, closer analysis of the trajectories reveals another potential advantage of the linecut constraint; as we neglect to perform any importance sampling, lc-AFQMC is rigorously equivalent to fp-AFQMC until the first walker is killed. By plotting the average number of walkers removed by the constraint each step (Fig. 5.10), we can observe the nearly exact (and up until this point, sign problem free) trajectory become biased. As many free projection calculations make use of energies obtained at around 5 Ha^{-1}, the point where the linecut constraint begins removing walkers, any significant deviation past this provides a clear *a priori* estimate of the bias due to the linecut constraint, and thus the quality of the trial.

For particularly problematic cases such as FeO, we propose taking further advantage of this relationship between linecut release AFQMC (lcR-AFQMC) and free-projection AFQMC (fp-

CI %	ph error	ph st error	lc error	lc st error
90	5.5	0.2	88.8	1.8
95	3.2	0.1	69.8	1.6
96	3.0	0.1	62.4	2.2
97	2.4	0.1	45.7	1.3
98	2.1	0.1	22.2	0.8
99	1.3	0.1	7.7	2.4

Table 5.2: Errors (kcal/mol) of phaseless and linecut AFQMC at long imaginary times for FeO (def2-SVP) using different numbers of determinants from a selected CI trial with a full active space. lc-AFQMC performs significantly worse than ph-AFQMC. Note that due to cost we only ran lc-AFQMC at 99% CI weight (26,071 determinants) for 60 Ha^{-1}, hence the higher reported statistical error.

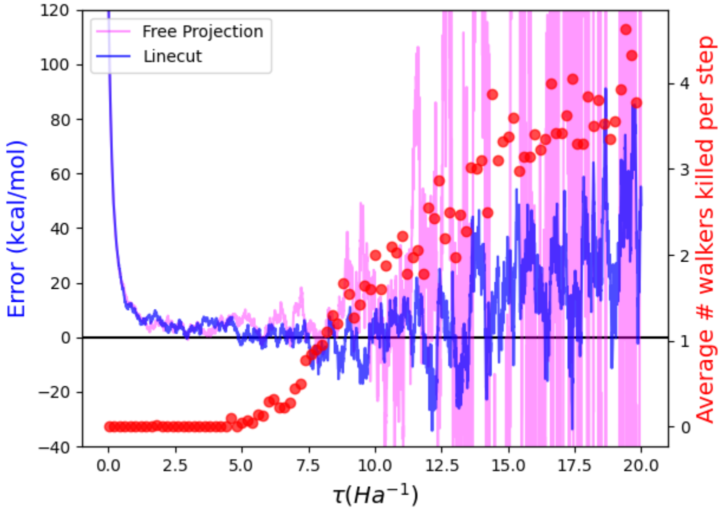


Figure 5.10: Error of AFQMC trajectories versus exact energy for FeO (def2-SVP, 864 determinants) using linecut and free projection AFQMC. The two are statistically equivalent (although run with different random walks) until the linecut constraint activates at around 5 Ha^{-1}, at which point fp-AFQMC encounters the sign problem and loses signal, and lc-AFQMC incurs a large bias. Both trajectories use 5000 walkers.

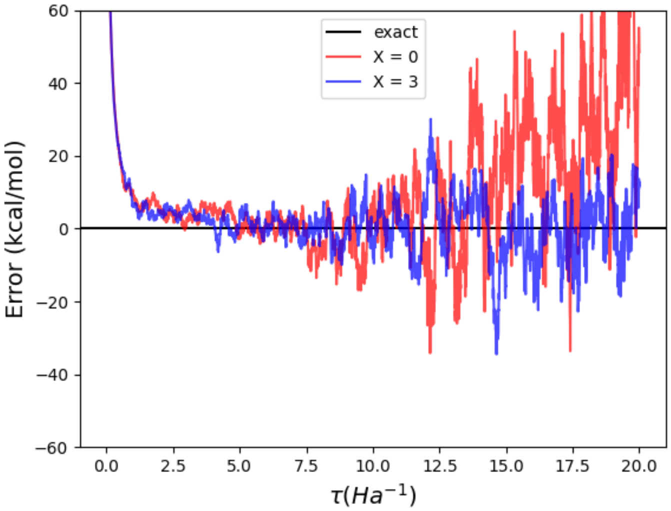


Figure 5.11: Error of AFQMC trajectories versus exact energy for FeO (def2-SVP, 864 determinants) for $X = 0$ and $X = 3$, where we observe the minimum of statistical error. By partially releasing the constraint we delay the onset of both the linecut bias and sign problem noise is delayed for long enough in imaginary time to obtain time-averaged results within chemical accuracy of the exact answer.

AFQMC) by obtaining energies at short imaginary times. By slowly releasing the linecut constraint, one can effectively delay the onset of the linecut bias while minimizing the noise growth of the sign problem, allowing for the estimation of energies over a small portion of the short time trajectory, rather than the typical method of running $O(10^5)$ trajectories to estimate the energy at one timestep as in free projection. Choosing the equilibration time is additionally straightforward, as one can choose the point directly prior to the activation of the linecut constraint. We find for such simulations of FeO, we can systematically reduce the bias with increasing X, and additionally reduce the statistical error, as the overall change in the trajectory that is averaged over is additionally reduced with the bias (Fig. 5.11). At a certain X, the statistical error begins increasing again due to the sign problem, and lcR-AFQMC no longer provides a significant statistical advantage (Table 5.3).

X	Error (kcal/mol)	standard error
0	5.7	3.5
1	3.9	2.9
2	0.7	1.3
3	0.4	0.9
4	1.5	1.5
5	0.3	1.6
FP	-57.6	51.4
phaseless	2.7	0.3

Table 5.3: Systematic release of the linecut constraint for FeO using 864 determinants in the trial. All calculations use 5000 walkers and are averaged over 4 to 20 Ha^{-1}. Note the initial decrease in statistical error with X, due to delaying the onset of the linecut bias, followed by an increase at $X = 4$ due to the sign problem.

5.4 Conclusions

We have formulated a new constraint for use within the AFQMC framework, which allows phase accumulation up to $|\theta| = \pi$ and retains stability for long timescales. Tests on a variety of small systems demonstrate comparable accuracy with qualitatively distinct behavior versus ph-AFQMC, with some complimentary advantages for lc-AFQMC. When sizable errors do arise in lc-AFQMC, they may be remedied by the use of systematically more sophisticated trials, akin to standard practice in ph-AFQMC and are easily identified by increased variance.

The stability of the lc-AFQMC approach suggests that the majority of the phase problem of AFQMC manifests from walkers crossing the origin in the complex plane, rather than generic accumulation of phase with respect to the trial wave function. This suggests the possibility of the design of alternative constraints that avoid origin crossing, and thus may serve as a first step in the creation of useful new constraints.

With respect to the release of the linecut constraint, we emphasize that such calculations are asymptotically unstable, as they allow a finite sampling of the exponential sign problem. Nevertheless this technique appears to be numerically useful for systems in which the trial is already qualitatively correct, as in the case of C_2. Additionally, monitoring the trajectory as the constraint is activated allows for a systematic internal estimate of the constraint bias, which we view as a significant advantage over alternate algorithms. Similar procedures for constraint release may not be as effective within ph-AFQMC due to the use of importance sampling (i.e. the force bias), which effectively removes walkers that violate the constraint indirectly. In this framework, one is restricted to either fully constrained or fully unconstrained calculations, which lessens the applicability, although it has been effectively used in studying lattice models.[301] A recent work[302] describes a method to remove this instability within ph-AFQMC through the use back propagation followed by a non-constrained Markov chain Monte Carlo simulation. It is of course in principle possible to develop an importance sampling method akin to the standard force bias which respects the linecut constraint and thus reduces the variance within lc-AFQMC; in order for this to be useful, however, whatever savings obtained due to reduced sampling would have to outweigh the increased cost and complexity of partially releasing the constraint, as well as the added cost for computing the importance sampling function. If the variance of lc-AFQMC without importance sampling remains 2-4x as large as that of ph-AFQMC, this would result in approximately 4-16x the sampling necessary to obtain comparable error bars. However, it is likely that with increasing system size the increase in variance will eventually become a high enough detriment to necessitate importance sampling.

While the results presented here are promising, we stress that benchmarking on larger datasets is paramount to determine the ultimate utility of lc-AFQMC for generating benchmark energies at lower cost than ph-AFQMC. Indeed, our results even on a small benchmark set do not suggest uniform improvements over ph-AFQMC. We thus suggest that the use of lc-AFQMC in this way is restricted to trials for which systematic extrapolation to the unbiased limit is possible, or to trials which do not show significant deviation upon the onset of the constraint at short imaginary times. To this end, lcR-AFQMC may prove a useful tool by assisting in converging results for difficult systems or in helping diagnose insufficient trials.

Although our conclusions with regards to relative accuracy versus ph-AFQMC are somewhat preliminary, the stability of lc-AFQMC represents a new class of constrained AFQMC which has significant utility independent from direct usage. In particular, the observed distinct and in some cases directly opposing trends of lc-AFQMC versus ph-AFQMC (as in the dissociation curve of H_4), suggest a fertile research pathway towards understanding the behavior of bias in constrained AFQMC. By changing the constraint, lc-AFQMC provides another parameter beyond simply increasing the complexity of the trial wave function with which to study the relationship between the trial and constraint bias, and will likely lead to more insight into the performance and occasional failures of ph-AFQMC. More in-depth studies of both the phaseless and linecut biases are necessary to determine if one can correlate the relative accuracy with specific features of the trial function. In addition to more fundamental research such as this, it is possible to merely use lc-AFQMC as a measure of trial quality for general ph-AFQMC simulations by monitoring the change in energy upon initialization of the constraint for short trajectories. This can additionally provide clear internal metrics with which to study the relationship between trial quality and biases in constrained AFQMC. More extensive benchmark studies which may help answer these and related questions will be pursued in future work.

5.5 Appendix

5.5.1 Details of AFQMC calculations

All ph-AFQMC and lc-AFQMC calculations were run on our custom AFQMC code with 0.005 Ha^{-1} timesteps unless otherwise noted. Trials were generated using PySCF, and the integrals were decomposed using the modified Cholesky Decomposition algorithm with a threshold of 1e-6. Selected CI trials were generated using Dice interfaced with PySCF. All calculations made use of the "comb" population control algorithm upon every energy evaluation, which occurs every 20 steps.

We tested for the convergence of lc-AFQMC with respect to timestep and walker population on H_4 in the cc-pVQZ basis at $R = 1.13$, where we saw the largest error out of all linecut calculations. Neither doubling the walker population nor halving the timestep led to a change in the result outside of statistics, suggesting that the simulation is converged with respect to these parameters.

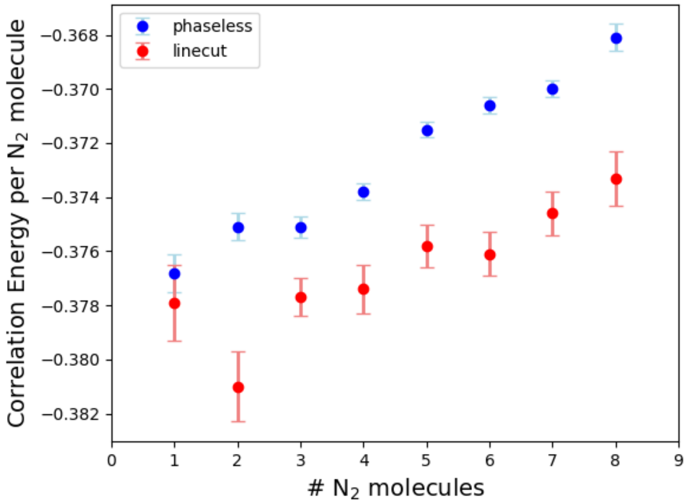


Figure 5.12: Correlation energy per N_2 atom for both phaseless and linecut AFQMC (cc-pVDZ) using a 0.05 Ha^{-1} timestep.

5.5.2 Size Consistency and Stability

It has been recently pointed out that ph-AFQMC for a fixed and finite timestep is not size-consistent. This was attributed to the cosine projection, and posited that the removal of this from the constraint could result in a size consistent method even for finite time step size.[299] To this end, we tested a series of N_2 molecules for size consistency using a fairly large timestep of 0.05 Ha^{-1} (Fig. 5.12). While the better performance of lc-AFQMC is evident vs ph-AFQMC, it still exhibits growth of correlation energy as the number of N_2 molecules is increased, suggesting that the removal of the cosine projection is not the only requirement for a size consistent constraint prior to the limit $\Delta\tau \rightarrow 0$ being taken. It is necessary to again note that this does not preclude general size consistency for constrained AFQMC, which is rigorously size consistent upon converging the timestep.

References

[1] B. Rudshteyn *et al.*, “Predicting ligand-dissociation energies of 3d coordination complexes with auxiliary-field quantum monte carlo,” *J. Chem. Theory Comput.*, vol. 16, no. 5, pp. 3041–3054, 2020.

[2] B. Rudshteyn *et al.*, “Calculation of metallocene ionization potentials via auxiliary field quantum monte carlo: Toward benchmark quantum chemistry for transition metals,” *J. Chem. Theory Comput.*, vol. 18, no. 5, pp. 2845–2862, 2022.

[3] J. L. Weber *et al.*, “In silico prediction of annihilators for triplet–triplet annihilation upconversion via auxiliary-field quantum monte carlo,” *Chemical Science*, vol. 12, no. 3, pp. 1068–1079, 2021.

[4] J. L. Weber *et al.*, “A localized-orbital energy evaluation for auxiliary-field quantum monte carlo,” *J. Chem. Theory Comput.*, vol. 18, no. 6, pp. 3447–3459, 2022.

[5] J. Shee, J. L. Weber, D. R. Reichman, R. A. Friesner, and S. Zhang, “On the potentially transformative role of auxiliary-field quantum monte carlo in quantum chemistry: A highly accurate method for transition metals and beyond,” *The Journal of Chemical Physics*, vol. 158, no. 14, 2023.

[6] G. Li *et al.*, “Accurate quantum chemical calculation of ionization potentials: Validation of the dft-loc approach via a large data set obtained from experiments and benchmark quantum chemical calculations,” *Journal of Chemical Theory and Computation*, vol. 16, no. 4, pp. 2109–2123, 2020.

[7] K. J. Fallon *et al.*, “Molecular engineering of chromophores to enable triplet–triplet annihilation upconversion,” *Journal of the American Chemical Society*, vol. 142, no. 47, pp. 19 917–19 925, 2020.

[8] N. E. S. Tay *et al.*, “Targeted activation in localized protein environments via deep red photoredox catalysis,” *Nature Chemistry*, pp. 1–9, 2022.

[9] C. K. Prier, D. A. Rankic, and D. W. C. MacMillan, “Visible light photoredox catalysis with transition metal complexes: Applications in organic synthesis,” *Chem. Rev.*, vol. 113, no. 7, pp. 5322–5363, 2013.

[10] A. Trautwein, *Bioinorganic Chemistry: Transition Metals in Biology and Their Coordination Chemistry*. Wiley-VCH, 1997.

[11] D. Khomskii, *Transition Metal Compounds*. Cambridge University Press, 2014, ISBN: 9781316060803.

[12] S. Raugei, L. C. Seefeldt, and B. M. Hoffman, "Critical computational analysis illuminates the reductive-elimination mechanism that activates nitrogenase for n_2 reduction," *Proc. Nat. Acad. Sci.*, vol. 115, no. 45, E10521–E10530, 2018.

[13] M. Askerka, G. W. Brudvig, and V. S. Batista, "The O_2-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.*, vol. 50, pp. 41–48, Dec. 2016.

[14] P. E. M. Siegbahn, "Nucleophilic water attack is not a possible mechanism for o–o bond formation in photosystem ii," *Proc. Nat. Acad. Sci.*, vol. 114, no. 19, pp. 4966–4968, 2017.

[15] L. Riccardi, V. Genna, and M. De Vivo, "Metal–ligand interactions in drug design," *Nat. Rev. Chem.*, vol. 2, no. 7, pp. 100–112, 2018.

[16] P. A. Lee, N. Nagaosa, and X.-G. Wen, "Doping a mott insulator: Physics of high-temperature superconductivity," *Rev. Mod. Phys.*, vol. 78, no. 1, p. 17, 2006.

[17] E. Dagotto, "Correlated electrons in high-temperature superconductors," *Rev. Mod. Phys.*, vol. 66, no. 3, p. 763, 1994.

[18] R. A. Friesner, "Modeling metalloenzymes with density functional and mixed quantum mechanical/molecular mechanical (qm/mm) calculations: Progress and challenges," *Encyclopedia of Inorganic and Bioinorganic Chemistry*, 2011.

[19] S. V. Jerome, T. F. Hughes, and R. A. Friesner, "Successful application of the dbloc method to the hydroxylation of camphor by cytochrome p450," *Protein Sci.*, vol. 25, 277–285, 2016.

[20] B. Rudshteyn *et al.*, "Water-nucleophilic attack mechanism for the $cu^{II}(pyalk)_2$ water-oxidation catalyst," *ACS Catal.*, vol. 8, no. 9, pp. 7952–7960, 2018.

[21] J. Agarwal, E. Fujita, H. F. Schaefer III, and J. T. Muckerman, "Mechanisms for co production from co_2 using reduced rhenium tricarbonyl catalysts," *J. Am. Chem. Soc*, vol. 134, no. 11, pp. 5180–5186, 2012.

[22] J. L. Han, J You, H Yonemura, S Yamada, S. Wang, and X. G. Li, "Metallophthalocyanines as triplet sensitizers for highly efficient photon upconversion based on sensitized triplet–triplet annihilation," *Photochem. Photobiol. Sci.*, vol. 15, no. 8, pp. 1039–1045, 2016.

[23] N. Mardirossian and M. Head-Gordon, "ωB97x-v: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, de-

signed by a survival-of-the-fittest strategy," *Phys. Chem. Chem. Phys.*, vol. 16, no. 21, pp. 9904–9924, 2014.

[24] N. Mardirossian and M. Head-Gordon, "ωB97m-v: A combinatorially optimized, range-separated hybrid, meta-gga density functional with vv10 nonlocal correlation," *J. Chem. Phys.*, vol. 144, no. 21, p. 214 110, 2016.

[25] A. V. Marenich, C. J. Cramer, and D. G. Truhlar, "Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions," *J. Phys. Chem. B*, vol. 113, no. 18, pp. 6378–6396, 2009.

[26] D. H. Bross, J. G. Hill, H.-J. Werner, and K. A. Peterson, "Explicitly correlated composite thermochemistry of transition metal species," *J. Chem. Phys.*, vol. 139, no. 9, p. 094 302, 2013.

[27] S. Manivasagam, M. L. Laury, and A. K. Wilson, "Pseudopotential-based correlation consistent composite approach (rp-ccca) for first-and second-row transition metal thermochemistry," *J. Phys. Chem. A*, vol. 119, no. 26, pp. 6867–6874, 2015.

[28] W. Jiang, M. L. Laury, M. Powell, and A. K. Wilson, "Comparative study of single and double hybrid density functionals for the prediction of 3d transition metal thermochemistry," *J. Comput. Theory Comput.*, vol. 8, no. 11, pp. 4102–4111, 2012.

[29] W. Zhang, D. G. Truhlar, and M. Tang, "Tests of exchange-correlation functional approximations against reliable experimental data for average bond energies of 3d transition metal compounds," *J. Chem. Theory Comput.*, vol. 9, no. 9, pp. 3965–3977, 2013.

[30] K. A. Moltved and K. P. Kepp, "Chemical bond energies of 3d transition metals studied by density functional theory," *J. Chem. Theory Comput.*, vol. 14, no. 7, pp. 3479–3492, 2018.

[31] W. Jiang, N. J. DeYonker, J. J. Determan, and A. K. Wilson, "Toward accurate theoretical thermochemistry of first row transition metal complexes," *J. Phys. Chem. A*, vol. 116, no. 2, pp. 870–885, 2011.

[32] R. K. Carlson, G. Li Manni, A. L. Sonnenberger, D. G. Truhlar, and L. Gagliardi, "Multiconfiguration pair-density functional theory: Barrier heights and main group and transition metal energetics," *J. Chem. Theor. Comp.*, vol. 11, no. 1, pp. 82–90, 2014.

[33] J. L. Bao, S. O. Odoh, L. Gagliardi, and D. G. Truhlar, "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.*, vol. 13, no. 2, pp. 616–626, 2017.

[34] J. L. Bao, X. Zhang, X. Xu, and D. G. Truhlar, “Predicting bond dissociation energy and bond length for bimetallic diatomic molecules: A challenge for electronic structure theory,” *Phys. Chem. Chem. Phys.*, vol. 19, no. 8, pp. 5839–5854, 2017.

[35] K. Sharkas, L. Gagliardi, and D. G. Truhlar, “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*, vol. 121, no. 48, pp. 9392–9400, 2017.

[36] L. N. Tran, S. Iskakov, and D. Zgid, “Spin-unrestricted self-energy embedding theory,” *J. Phys. Chem. Lett.*, vol. 9, no. 15, pp. 4444–4450, 2018.

[37] R. J. Bartlett and M. Musiał, “Coupled-cluster theory in quantum chemistry,” *Rev. Mod. Phys.*, vol. 79, no. 1, p. 291, 2007.

[38] M. M. Quintal, A. Karton, M. A. Iron, A. D. Boese, and J. M. Martin, “Benchmark study of dft functionals for late-transition-metal reactions,” *J. Phys. Chem. A*, vol. 110, no. 2, pp. 709–716, 2006.

[39] M. Steinmetz and S. Grimme, “Benchmark study of the performance of density functional theory for bond activations with (ni, pd)-based transition-metal catalysts,” *ChemistryOpen*, vol. 2, no. 3, pp. 115–124, 2013.

[40] R. Kang, W. Lai, J. Yao, S. Shaik, and H. Chen, “How accurate can a local coupled cluster approach be in computing the activation energies of late-transition-metal-catalyzed reactions with au, pt, and ir?” *J. Chem. Theory Comput.*, vol. 8, no. 9, pp. 3119–3127, 2012.

[41] S. Dohm, A. Hansen, M. Steinmetz, S. Grimme, and M. P. Checinski, “Comprehensive thermochemical benchmark set of realistic closed-shell metal organic reactions,” *J. Chem. Theory Comput.*, vol. 14, no. 5, pp. 2596–2608, 2018.

[42] B. Chan, P. M. W. Gill, and M. Kimura, “Assessment of dft methods for transition metals with the tmc151 compilation of data sets and comparison with accuracies for main-group chemistry,” *J. Chem. Theor. Comp*, 2019.

[43] X. Xu, W. Zhang, M. Tang, and D. G. Truhlar, “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.*, vol. 11, no. 5, pp. 2036–2052, 2015.

[44] L. Cheng, J. Gauss, B. Ruscic, P. B. Armentrout, and J. F. Stanton, “Bond dissociation energies for diatomic molecules containing 3d transition metals: Benchmark scalar-relativistic coupled-cluster calculations for 20 molecules,” *J. Chem. Theory Comput.*, vol. 13, no. 3, pp. 1044–1056, 2017.

[45] Z. Fang, M. Vasiliu, K. A. Peterson, and D. A. Dixon, “Prediction of bond dissociation energies/heats of formation for diatomic transition metal compounds: Ccsd(t) works,” *J. Chem. Theory Comput.*, vol. 13, no. 3, pp. 1057–1066, 2017.

[46] Y. A. Aoto, A. P. de Lima Batista, A. Köhn, and A. G. S. de Oliveira-Filho, “How to arrive at accurate benchmark values for transition metal compounds: Computation or experiment?” *J. Chem. Theory Comput.*, vol. 13, no. 11, pp. 5291–5316, 2017.

[47] J. Shee, B. Rudshteyn, E. J. Arthur, S. Zhang, D. R. Reichman, and R. A. Friesner, “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.*, vol. 15, no. 4, pp. 2346–2358, 2019.

[48] K. T. Williams *et al.*, “Direct comparison of many-body methods for realistic electronic hamiltonians,” *Physical Review X*, vol. 10, no. 1, p. 011 041, 2020.

[49] D. Hait, N. M. Tubman, D. S. Levine, K. B. Whaley, and M. Head-Gordon, “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.*, vol. 15, pp. 5370–5385, 2019.

[50] Y. Guo *et al.*, “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.*, vol. 148, no. 1, p. 011 101, 2018.

[51] C. Riplinger, P. Pinski, U. Becker, E. F. Valeev, and F. Neese, “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.*, vol. 144, no. 2, p. 024 109, 2016.

[52] T. Husch, L. Freitag, and M. Reiher, “Calculation of ligand dissociation energies in large transition-metal complexes,” *J. Chem. Theory Comput.*, vol. 14, no. 5, pp. 2456–2468, 2018.

[53] M. R. Momeni and A. Brown, “Why do td-dft excitation energies of bodipy/aza-bodipy families largely deviate from experiment? answers from electron correlated and multireference methods,” *J. Chem. Theory Comput.*, vol. 11, no. 6, pp. 2619–2632, 2015.

[54] S. Zhang and H. Krakauer, “Quantum monte carlo method using phase-free random walks with slater determinants,” *Phys. Rev. Lett.*, vol. 90, no. 13, p. 136 401, 2003.

[55] W. A. Al-Saidi, S. Zhang, and H. Krakauer, “Auxiliary-field quantum monte carlo calculations of molecular systems with a gaussian basis,” *J. Chem. Phys.*, vol. 124, no. 22, p. 224 101, 2006.

[56] J. Shee, S. Zhang, D. R. Reichman, and R. A. Friesner, “Chemical transformations approaching chemical accuracy via correlated sampling in auxiliary-field quantum monte carlo,” *J. Chem. Theory Comput.*, vol. 13, no. 6, pp. 2667–2680, 2017.

[57] J. Shee, E. J. Arthur, S. Zhang, D. R. Reichman, and R. A. Friesner, “Phaseless auxiliary-field quantum monte carlo on graphical processing units,” *J. Chem. Theory Comput.*, vol. 14, no. 8, pp. 4109–4121, 2018.

[58] Y.-R. Luo, *Comprehensive Handbook of Chemical Bond Energies*. CRC press, 2007.

[59] M. T. Rodgers and P. B. Armentrout, “Cationic noncovalent interactions: Energetics and periodic trends,” *Chem. Rev.*, vol. 116, no. 9, pp. 5642–5687, 2016.

[60] D. L. Hildenbrand, “Low-lying electronic states and revised thermochemistry of ticl, ticl$_2$, and ticl$_3$,” *J. Phys. Chem. A*, vol. 113, no. 8, pp. 1472–1474, 2009.

[61] A. D. Becke, “Density-functional thermochemistry. iii. the role of exact exchange,” *J. Chem. Phys.*, vol. 98, no. 7, pp. 5648–5652, 1993.

[62] S. H. Vosko, L. Wilk, and M. Nusair, “Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis,” *Can. J. Phys.*, vol. 58, no. 8, pp. 1200–1211, 1980.

[63] C. Lee, W. Yang, and R. G. Parr, “Development of the colle-salvetti correlation-energy formula into a functional of the electron density,” *Phys. Rev. B*, vol. 37, no. 2, p. 785, 1988.

[64] T. H. Dunning, “Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen,” *J. Chem. Phys.*, vol. 90, 1989.

[65] D. E. Woon and T. H. Dunning, “Gaussian basis sets for use in correlated molecular calculations. iii. the atoms aluminum through argon,” *J. Chem. Phys.*, vol. 98, 1993.

[66] W. A de Jong, R. J. Harrison, and D. A. Dixon, “Parallel douglas-kroll energy and gradients in nwchem: Estimating scalar relativistic effects using douglas-kroll contracted basis sets,” *J. Chem. Phys.*, vol. 114, 2001.

[67] N. B. Balabanov and K. A. Peterson, “Systematically convergent basis sets for transition metals. i. all-electron correlation consistent basis sets for the 3d elements sc–zn,” *J. Chem. Phys.*, vol. 123, no. 6, p. 064 107, 2005.

[68] F. Neese, “The orca program system,” *WIRES Comput. Mol. Sci.*, vol. 2, no. 1, pp. 73–78, 2012.

[69] D. A. Pantazis, X.-Y. Chen, C. R. Landis, and F. Neese, “All-electron scalar relativistic basis sets for third-row transition metal atoms,” *J. Chem. Theory Comput.*, vol. 4, no. 6, pp. 908–919, 2008.

[70] Q. Sun *et al.*, “Pyscf: The python–based simulations of chemistry framework,” *WIRES Comput. Mol. Sci.*, vol. 8, no. 1, e1340, eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1340`.

[71] W. Liu and D. Peng, “Exact two-component hamiltonians revisited,” *J. Chem. Phys.*, vol. 131, p. 031 104, 2009.

[72] J. Shee, E. J. Arthur, S. Zhang, D. R. Reichman, and R. A. Friesner, “Singlet–triplet energy gaps of organic biradicals and polyacenes with auxiliary-field quantum monte carlo,” *J. Chem. Theory Comput.*, vol. 15, no. 9, pp. 4924–4932, 2019.

[73] M. Motta and S. Zhang, “Ab initio computations of molecular systems by the auxiliary-field quantum monte carlo method,” *WIRES Comput. Mol. Sci.*, vol. 8, no. 5, e1364, 2018.

[74] N. J. DeYonker, K. A. Peterson, G. Steyl, A. K. Wilson, and T. R. Cundari, “Quantitative computational thermochemistry of transition metal species,” *J.Phys. Chem. A*, vol. 111, no. 44, pp. 11 269–11 277, 2007.

[75] G. Knizia, “Intrinsic atomic orbitals: An unbiased bridge between quantum theory and chemical concepts,” *J. Chem. Theory Comput.*, vol. 9, no. 11, pp. 4834–4843, 2013.

[76] E. R. Sayfutyarova, Q. Sun, G. K.-L. Chan, and G. Knizia, “Automated construction of molecular active spaces from atomic valence orbitals,” *J. Chem. Theory Comput.*, vol. 13, no. 9, pp. 4063–4078, 2017.

[77] Y. Zhao and D. G. Truhlar, “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,” *Theoretical Chemistry Accounts*, vol. 120, no. 1, pp. 215–241, May 2008.

[78] C. Adamo and V. Barone, “Toward reliable density functional methods without adjustable parameters: The pbe0 model,” *J. Chem. Phys.*, vol. 110, no. 13, pp. 6158–6170, 1999.

[79] S. Kozuch and J. M. Martin, “Dsd-pbep86: In search of the best double-hybrid dft with spin-component scaled mp2 and dispersion corrections,” *Phys. Chem. Chem. Phys.*, vol. 13, no. 45, pp. 20 104–20 107, 2011.

[80] S. Kozuch and J. M. Martin, “Spin-component-scaled double hybrids: An extensive search for the best fifth-rung functionals blending dft and perturbation theory,” *J. Comput. Chem.*, vol. 34, no. 27, pp. 2327–2344, 2013.

[81] L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi, and S. Grimme, “A look at the density functional theory zoo with the advanced gmtkn55 database for general main group thermochemistry, kinetics and noncovalent interactions,” *Phys. Chem. Chem. Phys.*, vol. 19, no. 48, pp. 32 184–32 215, 2017.

[82] G. Santra, N. Sylvetsky, and J. M. L. Martin, “Minimally empirical double hybrid functionals trained against the gmtkn55 database: Revdsd-pbep86-d4, revdod-pbe-d4, and dod-scan-d4,” *J. Phys. Chem. A.*, 2019.

[83] D. Coskun, S. V. Jerome, and R. A. Friesner, “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.*, vol. 12, no. 3, pp. 1121–1128, 2016.

[84] S. Grimme and A. Hansen, “A practicable real-space measure and visualization of static electron-correlation effects,” *Angew. Chem. Int. Ed.*, vol. 54, no. 42, pp. 12 308–12 313, 2015.

[85] C. A. Bauer, A. Hansen, and S. Grimme, “The fractional occupation number weighted density as a versatile analysis tool for molecules with a complicated electronic structure,” *Chem. Eur. J.*, vol. 23, no. 25, pp. 6150–6164, 2017.

[86] J. Lee and M. Head-Gordon, “Distinguishing artificial and essential symmetry breaking in a single determinant: Approach and application to the c_60, c_36, and c_20 fullerenes,” *Phys. Chem. Chem. Phys.*, vol. 21, no. 9, pp. 4763–4778, 2019.

[87] C. W. Bauschlicher, P. Siegbahn, and L. G. Pettersson, “The atomic states of nickel,” *Theor. Chim, Acta*, vol. 74, no. 6, pp. 479–491, 1988.

[88] K. Andersson and B. O. Roos, “Excitation energies in the nickel atom studied with the complete active space scf method and second-order perturbation theory,” *Chem. Phys. Lett.*, vol. 191, no. 6, pp. 507–514, 1992.

[89] L. Goerigk and S. Grimme, “Efficient and accurate double-hybrid-meta-gga density functionals–evaluation with the extended gmtkn30 database for general main group thermochemistry, kinetics, and noncovalent interactions,” *J. Chem. Theory Comput.*, vol. 7, no. 2, pp. 291–309, 2010.

[90] B. Chan and L. Radom, “Accurate quadruple–zeta basis-set approximation for double-hybrid density functional theory with an order of magnitude reduction in computational cost,” *Theor. Chem. Acc.*, vol. 133, pp. 1426–1429, 2014.

[91] D. G. Liakos, M. Sparta, M. K. Kesharwani, J. M. L. Martin, and F. Neese, “Exploring the accuracy limits of local pair natural orbital coupled-cluster theory,” *J. Chem. Theory Comput.*, vol. 11, no. 4, pp. 1525–1539, 2015.

[92] R. A. Friesner, “Ab initio quantum chemistry: Methodology and applications,” *Proc. Nat. Acad. Sci.*, vol. 102, no. 19, pp. 6648–6653, 2005.

[93] L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, “Assessment of gaussian-3 and density functional theories for a larger experimental test set,” *J. Chem. Phys.*, vol. 112, no. 17, pp. 7374–7383, 2000.

[94] L. A. Curtiss, K. Raghavachari, P. C. Redfern, A. G. Baboul, and J. A. Pople, “Gaussian-3 theory using coupled cluster energies,” *Chem. Phys. Lett.*, vol. 314, no. 1-2, pp. 101–107, 1999.

[95] J. P. Perdew, K. Burke, and M. Ernzerhof, “Generalized gradient approximation made simple,” *Phys. Rev. Lett.*, vol. 77, no. 18, p. 3865, 1996.

[96] A. D. Becke, “Density-functional thermochemistry. v. systematic optimization of exchange-correlation functionals,” *J. Chem. Phys.*, vol. 107, no. 20, pp. 8554–8560, 1997.

[97] H. L. Schmider and A. D. Becke, “Optimized density functionals from the extended g2 test set,” *J. Chem. Phys.*, vol. 108, no. 23, pp. 9624–9631, 1998.

[98] M. Ernzerhof and J. P. Perdew, “Generalized gradient approximation to the angle-and system-averaged exchange hole,” *J. Chem. Phys.*, vol. 109, no. 9, pp. 3313–3320, 1998.

[99] F. A. Hamprecht, A. J. Cohen, D. J. Tozer, and N. C. Handy, “Development and assessment of new exchange-correlation functionals,” *J. Chem. Phys.*, vol. 109, no. 15, pp. 6264–6271, 1998.

[100] Q. Ma, M. Schwilk, C. Koppl, and H.-J. Werner, “Scalable electron correlation methods. 4. parallel explicitly correlated local coupled cluster with pair natural orbitals (pno-lccsd-f12),” *J. Chem. Comput. Chem.*, vol. 13, no. 10, pp. 4871–4896, 2017.

[101] Q. Ma and H.-J. Werner, “Accurate intermolecular interaction energies using explicitly correlated local coupled cluster methods [pno-lccsd (t)-f12],” *J. Chem. Theory Comput.*, vol. 15, no. 2, pp. 1044–1052, 2019.

[102] C. Riplinger, B. Sandhoefer, A. Hansen, and F. Neese, “Natural triple excitations in local coupled cluster calculations with pair natural orbitals,” *J. Chem. Phys.*, vol. 139, no. 13, p. 134 101, 2013.

[103] M. Saitow, U. Becker, C. Riplinger, E. F. Valeev, and F. Neese, “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.*, vol. 146, no. 16, p. 164 105, 2017.

[104] B. M. Flöser, Y. Guo, C. Riplinger, F. Tuczek, and F. Neese, "Detailed pair natural orbital–based coupled cluster studies of spin crossover energetics," *J. Chem. Theory Comput.*, vol. 16, no. 4, pp. 2224–2235, 2020.

[105] H. Lin and D. G. Truhlar, "Qm/mm: What have we learned, where are we, and where do we go from here?" *Theor. Chem. Acc.*, vol. 117, no. 2, p. 185, 2007.

[106] G. Bistoni, I. Polyak, M. Sparta, W. Thiel, and F. Neese, "Toward accurate qm/mm reaction barriers with large qm regions using domain based pair natural orbital coupled cluster theory," *J. Chem. Theory Comput.*, vol. 14, no. 7, pp. 3524–3531, 2018.

[107] T. F. Hughes and R. A. Friesner, "Development of accurate dft methods for computing redox potentials of transition metal complexes: Results for model complexes and application to cytochrome p450," *J. Chem. Theory Comput.*, vol. 8, no. 2, pp. 442–459, 2012.

[108] L. E. Roy, E. Jakubikova, M. G. Guthrie, and E. R. Batista, "Calculation of one-electron redox potentials revisited. is it possible to calculate accurate potentials with density functional methods?" *J. Phys. Chem. A*, vol. 113, no. 24, pp. 6745–6750, 2009.

[109] S. J. Konezny, M. D. Doherty, O. R. Luca, R. H. Crabtree, G. L. Soloveichik, and V. S. Batista, "Reduction of systematic uncertainty in dft redox potentials of transition-metal complexes," *J. Phys. Chem. C*, vol. 116, no. 10, pp. 6349–6356, 2012.

[110] J. C. Green, "Gas phase photoelectron spectra of d-and f-block organometallic compounds," in *Bonding Problems*. Springer, 1981, pp. 37–112.

[111] H. Shi and S. Zhang, "Some recent developments in auxiliary-field quantum monte carlo for real materials," *J. Chem. Phys.*, vol. 154, no. 2, p. 024 107, 2021.

[112] F. D. Malone, S. Zhang, and M. A. Morales, "Accelerating auxiliary-field quantum monte carlo simulations of solids with graphical processing units," *J. Chem. Theory Comput.*, vol. 16, no. 7, pp. 4286–4297, 2020.

[113] J. Lee, F. D. Malone, and M. A. Morales, "Utilizing essential symmetry breaking in auxiliary-field quantum monte carlo: Application to the spin gaps of the c_{36} fullerene and an iron porphyrin model complex," *J. Chem. Theory Comput.*, vol. 16, no. 5, pp. 3019–3027, 2020.

[114] M. D. Morse, "Predissociation measurements of bond dissociation energies," *Acc. Chem. Res.*, vol. 52, no. 1, pp. 119–126, 2019. eprint: `https://doi.org/10.1021/acs.accounts.8b00526`.

[115] V. V. Pavlishchuk and A. W. Addison, "Conversion constants for redox potentials measured versus different reference electrodes in acetonitrile solutions at 25 °c," *Inorganica Chimica Acta*, vol. 298, no. 1, pp. 97–102, 2000.

[116] D. E. Richardson, "Applications of gas-phase electron-transfer equilibria in organometallic redox thermochemistry," in *Organometallic Ion Chemistry*. Springer, 1996, pp. 259–282.

[117] M. F. Ryan, J. R. Eyler, and D. E. Richardson, "Adiabatic ionization energies, bond disruption enthalpies, and solvation free energies for gas-phase metallocenes and metallocenium ions," *J. Am. Chem. Soc.*, vol. 114, no. 22, pp. 8611–8619, 1992.

[118] M. F. Ryan, D. E. Richardson, D. L. Lichtenberger, and N. E. Gruhn, "Gas-phase ionization energetics, electron-transfer kinetics, and ion solvation thermochemistry of decamethylmetallocenes, chromocene, and cobaltocene," *Organometallics*, vol. 13, no. 4, pp. 1190–1199, 1994.

[119] C. Cauletti *et al.*, "Photoelectron spectra of metallocenes," *J. Elec. Spec. Rel. Phen.*, vol. 19, no. 3, pp. 327–353, 1980.

[120] R. H. Crabtree, *The Organometallic Chemistry of the Transition Metals*. John Wiley & Sons, 2009, ISBN: 0470257628.

[121] K. Ishimura, M. Hada, and H. Nakatsuji, "Ionized and excited states of ferrocene: Symmetry adapted cluster–configuration–interaction study," *J. Chem. Phys.*, vol. 117, no. 14, pp. 6533–6537, 2002.

[122] L. M. J. Huntington and M. Nooijen, "Application of multireference equation of motion coupled-cluster theory to transition metal complexes and an orbital selection scheme for the efficient calculation of excitation energies," *J. Chem. Phys.*, vol. 142, no. 19, p. 194 111, 2015.

[123] L. M. J. Huntington, O. Demel, and M. Nooijen, "Benchmark applications of variations of multireference equation of motion coupled-cluster theory," *J. Chem. Theory Comput.*, vol. 12, no. 1, pp. 114–132, 2016.

[124] Q. M. Phung, S. Vancoillie, and K. Pierloot, "A multiconfigurational perturbation theory and density functional theory study on the heterolytic dissociation enthalpy of first-row metallocenes," *J. Chem. Theory Comput.*, vol. 8, no. 3, pp. 883–892, 2012.

[125] M. Namazian, C. Y. Lin, and M. L. Coote, "Benchmark calculations of absolute reduction potential of ferricinium/ferrocene couple in nonaqueous solutions," *J. Chem. Theory Comput.*, vol. 6, no. 9, pp. 2721–2725, 2010.

[126] J. Shee, M. Loipersberger, D. Hait, J. Lee, and M. Head-Gordon, "Revealing the nature of electron correlation in transition metal complexes with symmetry-breaking and chemical intuition," *J. Chem. Phys.*, vol. 154, no. 19, p. 194 109, 2021.

[127] J. Shee, M. Loipersberger, A. Rettig, J. Lee, and M. Head-Gordon, "Regularized second-order møller–plesset theory: A more accurate alternative to conventional mp2 for nonco-

valent interactions and transition metal thermochemistry for the same computational cost," *J. Phys. Chem. Lett.*, vol. 12, pp. 12 084–12 097, 2021.

[128] P. Sharpe, J. R. Eyler, and D. E. Richardson, "Free energies of electron attachment to tris (acetylacetonate) and tris (hexafluoroacetylacetonate) transition-metal complexes in the gas phase: Experimental results and ligand field analysis," *Inorg. Chem.*, vol. 29, no. 15, pp. 2779–2787, 1990.

[129] P. Sharpe and D. E. Richardson, "Metal-ligand bond energies and solvation energies for gas-phase transition-metal tris (acetylacetonate) complexes and their negative ions," *J. Am. Chem. Soc.*, vol. 113, no. 22, pp. 8339–8346, 1991.

[130] K. Nawa *et al.*, "Search for the ground-state electronic configurations of correlated organometallic metallocenes from constraint density functional theory," *Phys. Rev. B*, vol. 94, no. 3, p. 035 136, 2016.

[131] S Evans, M. L. H. Green, B Jewitt, A. F. Orchard, and C. F. Pygall, "Electronic Structure of Metal Complexes Containing π-Cyclopentadienyl and Related Ligands. Part 1.—He (I) Photoelectron Spectra of Some Closed-Shell Metallocenes," *J. Chem. Soc., Faraday Trans. 2*, vol. 68, pp. 1847–1865, 1972.

[132] S Evans, M. L. H. Green, B Jewitt, G. H. King, and A. F. Orchard, "Electronic structures of metal complexes containing the Π-cyclopentadienyl and related ligands. part 2.—he i photoelectron spectra of the open-shell metallocenes," *J. Chem. Soc., Faraday Trans. 2*, vol. 70, pp. 356–376, 1974.

[133] K. R. Gordon and K. D. Warren, "Spectroscopic and magnetic studies of the 3d metallocenes," *Inorg. Chem.*, vol. 17, no. 4, pp. 987–994, 1978.

[134] M. F. Rettig and R. S. Drago, "Electron delocalization in paramagnetic metallocenes. i. nuclear magnetic resonance contact shifts," *J. Am. Chem. Soc.*, vol. 91, no. 6, pp. 1361–1370, 1969.

[135] Z.-F. Xu, Y. Xie, W.-L. Feng, and H. F. Schaefer, "Systematic investigation of electronic and molecular structures for the first transition metal series metallocenes m $(c_5h_5)_2$ (m= v, cr, mn, fe, co, and ni)," *J. Phys. Chem. A*, vol. 107, no. 15, pp. 2716–2729, 2003.

[136] S. Trasatti *et al.*, "The absolute electrode potential: An explanatory note (recommendations 1986)," *Pure Appl. Chem*, vol. 58, no. 7, pp. 955–966, 1986.

[137] S. Bai, R. Mansour, L. Stojanović, J. M. Toldo, and M. Barbatti, "On the origin of the shift between vertical excitation and band maximum in molecular photoabsorption," *J. Molec. Model.*, vol. 26, no. 5, pp. 1–9, 2020.

[138] C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*. John Wiley & Sons, 2013.

[139] T. Nakajima and K. Hirao, "The higher-order douglas–kroll transformation," *J. Chem. Phys.*, vol. 113, no. 18, pp. 7786–7789, 2000.

[140] M. Mayer, S. Krüger, and N. Rösch, "Two-component variant of the douglas–kroll relativistic linear combination of gaussian-type orbitals density-functional method: Spin–orbit effects in atoms and diatomics," *J. Chem. Phys.*, vol. 115, no. 10, pp. 4411–4423, 2001.

[141] A. Wolf, M. Reiher, and B. A. Hess, "The generalized douglas–kroll transformation," *J. Chem. Phys.*, vol. 117, no. 20, pp. 9215–9226, 2002.

[142] J. E. Peralta and G. E. Scuseria, "Relativistic all-electron two-component self- consistent density functional calculations including one-electron scalar and spin–orbit effects," *J. Chem. Phys.*, vol. 120, no. 13, pp. 5875–5881, 2004.

[143] E. Caldeweyher, C. Bannwarth, and S. Grimme, "Extension of the d3 dispersion coefficient model," *J. Chem. Phys.*, vol. 147, no. 3, p. 034 112, 2017.

[144] K. A. Peterson and C. Puzzarini, "Systematically convergent basis sets for transition metals. ii. pseudopotential-based correlation consistent basis sets for the group 11 (cu, ag, au) and 12 (zn, cd, hg) elements," *Theor. Chem. Acc.*, vol. 114, no. 4, pp. 283–296, 2005.

[145] D. Figgen, G. Rauhut, M. Dolg, and H. Stoll, "Energy-consistent pseudopotentials for group 11 and 12 atoms: Adjustment to multi-configuration dirac–hartree–fock data," *Chem. Phys.*, vol. 311, no. 1-2, pp. 227–244, 2005.

[146] W. Purwanto, H. Krakauer, Y. Virgus, and S. Zhang, "Assessing weak hydrogen binding on ca^{+} centers: An accurate many-body study with large basis sets," *J. Chem. Phys.*, vol. 135, no. 16, p. 164 105, 2011.

[147] A. Altun, M. Saitow, F. Neese, and G. Bistoni, "Local energy decomposition of open-shell molecular systems in the domain-based local pair natural orbital coupled cluster framework," *J. Chem. Theory Comput.*, vol. 15, no. 3, pp. 1616–1632, 2019.

[148] Y. Minenkov, G. Bistoni, C. Riplinger, A. A. Auer, F. Neese, and L. Cavallo, "Pair natural orbital and canonical coupled cluster reaction enthalpies involving light to heavy alkali and alkaline earth metals: The importance of sub-valence correlation," *Phys. Chem. Chem. Phys.*, vol. 19, no. 14, pp. 9374–9391, 2017.

[149] I. Efremenko and J. M. L. Martin, "Coupled cluster benchmark of new density functionals and of domain pair natural orbital methods: Mechanisms of hydroarylation and oxidative coupling catalyzed by ru(ii) chloride carbonyls," in *AIP Conf. Proc.*, AIP Publishing LLC, vol. 2186, 2019, p. 030 005.

[150] A. Altun, F. Neese, and G. Bistoni, “Extrapolation to the limit of a complete pair natural orbital space in local coupled-cluster calculations,” *J. Chem. Theory Comput.*, vol. 16, no. 10, pp. 6142–6149, 2020.

[151] T. Helgaker, W. Klopper, H. Koch, and J. Noga, “Basis-set convergence of correlated calculations on water,” *J. Chem. Phys.*, vol. 106, no. 23, pp. 9639–9646, 1997.

[152] M. Feldt, C. Martín-Fernández, and J. N. Harvey, “Energetics of non-heme iron reactivity: Can ab initio calculations provide the right answer?” *Phys. Chem. Chem. Phys.*, vol. 22, no. 41, pp. 23 908–23 919, 2020.

[153] K. G. Dyall, “Interfacing relativistic and nonrelativistic methods. i. normalized elimination of the small component in the modified dirac equation,” *J. Chem. Phys.*, vol. 106, no. 23, pp. 9618–9626, 1997.

[154] W. Kutzelnigg and W. Liu, “Quasirelativistic theory equivalent to fully relativistic theory,” *J. Chem. Phys.*, vol. 123, no. 24, p. 241 102, 2005. eprint: `https://doi.org/10.1063/1.2137315`.

[155] W. Liu and D. Peng, “Infinite-order quasirelativistic density functional method based on the exact matrix quasirelativistic theory,” *J. Chem. Phys.*, vol. 125, no. 4, p. 044 102, 2006.

[156] M. Iliaš and T. Saue, “An infinite-order two-component relativistic hamiltonian by a simple one-step transformation,” *J. Chem. Phys.*, vol. 126, no. 6, p. 064 102, 2007.

[157] L. Cheng and J. Gauss, “Analytical evaluation of first-order electrical properties based on the spin-free dirac-coulomb hamiltonian,” *J. Chem. Phys.*, vol. 134, no. 24, p. 244 112, 2011.

[158] F. Ma, S. Zhang, and H. Krakauer, “Excited state calculations in solids by auxiliary-field quantum monte carlo,” *New J. Phys.*, vol. 15, no. 9, p. 093 017, 2013.

[159] W. Purwanto, W. Al-Saidi, H. Krakauer, and S. Zhang, “Eliminating spin contamination in auxiliary-field quantum monte carlo: Realistic potential energy curve of f_2,” *J. Chem. Phys.*, vol. 128, no. 11, p. 114 309, 2008.

[160] B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov, and P.-O. Widmark, “New relativistic ano basis sets for transition metal atoms,” *J. Phys. Chem. A*, vol. 109, no. 29, pp. 6575–6579, 2005.

[161] A. D. Becke, “Density-functional exchange-energy approximation with correct asymptotic behavior,” *Phys. Rev. A*, vol. 38, no. 6, p. 3098, 1988.

[162] J. P. Perdew, “Density-functional approximation for the correlation energy of the inhomogeneous electron gas,” *Phys. Rev. B*, vol. 33, no. 12, p. 8822, 1986.

[163] O. Salomon, M. Reiher, and B. A. Hess, “Assertion and validation of the performance of the b3lyp* functional for the first transition metal row and the g2 test set,” *J. Chem. Phys.*, vol. 117, no. 10, pp. 4729–4737, 2002.

[164] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, “Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids,” *Phys. Rev. Lett.*, vol. 91, no. 14, p. 146 401, 2003.

[165] V. N. Staroverov, G. E. Scuseria, J. Tao, and J. P. Perdew, “Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes,” *J. Chem. Phys.*, vol. 119, no. 23, pp. 12 129–12 137, 2003.

[166] N. Mardirossian and M. Head-Gordon, “Mapping the genome of meta-generalized gradient approximation density functionals: The search for b97m-v,” *J. Chem. Phys*, vol. 142, no. 7, p. 074 111, 2015.

[167] N. Mardirossian and M. Head-Gordon, “Thirty years of density functional theory in computational chemistry: An overview and extensive assessment of 200 density functionals,” *Mol. Phys.*, vol. 115, no. 19, pp. 2315–2372, 2017.

[168] T. F. Hughes and R. A. Friesner, “Correcting systematic errors in dft spin-splitting energetics for transition metal complexes,” *J. Chem. Theory Comput.*, vol. 7, no. 1, pp. 19–32, 2011.

[169] R. A. Friesner, E. H. Knoll, and Y. Cao, “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.*, vol. 125, no. 12, p. 124 107, 2006.

[170] R. A. Friesner and S. V. Jerome, “Localized orbital corrections for density functional calculations on transition metal containing systems,” *Coord. Chem. Rev.*, vol. 344, pp. 205–213, 2017.

[171] M. Feldt, Q. M. Phung, K. Pierloot, R. A. Mata, and J. N. Harvey, “Limits of coupled-cluster calculations for non-heme iron complexes,” *J. Chem. Theory Comput.*, vol. 15, no. 2, pp. 922–937, 2019.

[172] L. Gagliardi, D. G. Truhlar, G. Li Manni, R. K. Carlson, C. E. Hoyer, and J. L. Bao, “Multiconfiguration pair-density functional theory: A new way to treat strongly correlated systems,” *Acc. Chem. Res.*, vol. 50, no. 1, pp. 66–73, 2016.

[173] K. Doblhoff-Dier, J. Meyer, P. E. Hoggan, G.-J. Kroes, and L. K. Wagner, “Diffusion monte carlo for accurate dissociation energies of 3d transition metal containing molecules,” *J. Chem. Theory Comput.*, vol. 12, no. 6, pp. 2583–2597, 2016.

[174] H. G. A. Burton and A. J. W. Thom, "Reaching full correlation through nonorthogonal configuration interaction: A second-order perturbative approach," *J. Chem. Theory Comput.*, vol. 16, no. 9, pp. 5586–5600, 2020.

[175] C. J. Cramer, M. Włoch, P. Piecuch, C. Puzzarini, and L. Gagliardi, "Theoretical models on the cu_2o_2 torture track: Mechanistic implications for oxytyrosinase and small-molecule analogues," *J. Phys. Chem. A.*, vol. 110, no. 5, pp. 1991–2004, 2006.

[176] L. R. Maurer, M. Bursch, S. Grimme, and A. Hansen, "Assessing density functional theory for chemically relevant open-shell transition metal reactions," *J. Chem. Theory Comput.*, 2021.

[177] M. A. Iron and T. Janes, "Evaluating transition metal barrier heights with the latest density functional theory exchange–correlation functionals: The mobh35 benchmark database," *J. Phys. Chem. A*, vol. 123, no. 17, pp. 3761–3781, 2019.

[178] G. Bistoni, C. Riplinger, Y. Minenkov, L. Cavallo, A. A. Auer, and F. Neese, "Treating subvalence correlation effects in domain based pair natural orbital coupled cluster calculations: An out-of-the-box approach," *J. Chem. Theory Comput.*, vol. 13, no. 7, pp. 3220–3227, 2017.

[179] J. C. Snyder, M. Rupp, K. Hansen, K.-R. Müller, and K. Burke, "Finding density functionals with machine learning," *Phys. Rev. Lett.*, vol. 108, no. 25, p. 253 002, 2012.

[180] J. Schmidt, C. L. Benavides-Riveros, and M. A. Marques, "Machine learning the physical nonlocal exchange–correlation functional of density-functional theory," *J. Phys. Chem. Lett.*, vol. 10, no. 20, pp. 6425–6431, 2019.

[181] R. Nagai, R. Akashi, and O. Sugino, "Completing density functional theory by machine learning hidden messages from molecules," *Npj Comput. Mater.*, vol. 6, no. 1, pp. 1–8, 2020.

[182] S. Dick and M. Fernandez-Serra, "Machine learning accurate exchange and correlation functionals of the electronic density," *Nat. Commun.*, vol. 11, no. 1, pp. 1–10, 2020.

[183] R. Meyer, M. Weichselbaum, and A. W. Hauser, "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.*, vol. 16, no. 9, pp. 5685–5694, 2020.

[184] B. Kalita, L. Li, R. J. McCarty, and K. Burke, "Learning to approximate density functionals," *Acc. Chem. Res.*, vol. 54, no. 4, pp. 818–826, 2021.

[185] T Ito, H Takagi, S Ishibashi, T Ido, and S Uchida, “Normal-state conductivity between cuo$_2$ planes in copper oxide superconductors,” *Nature*, vol. 350, no. 6319, pp. 596–598, 1991.

[186] Y. J. Uemura *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.*, vol. 62, no. 19, p. 2317, 1989.

[187] B. D. Ravetz, A. B. Pun, E. M. Churchill, D. N. Congreve, T. Rovis, and L. M. Campos, “Photoredox catalysis using infrared light via triplet fusion upconversion,” *Nature*, vol. 565, no. 7739, p. 343, 2019.

[188] M. H. Shaw, J. Twilton, and D. W. MacMillan, “Photoredox catalysis in organic chemistry,” *The Journal of organic chemistry*, vol. 81, no. 16, pp. 6898–6926, 2016.

[189] N. A. Romero and D. A. Nicewicz, “Organic photoredox catalysis,” *Chemical reviews*, vol. 116, no. 17, pp. 10 075–10 166, 2016.

[190] J. Zhou, Q. Liu, W. Feng, Y. Sun, and F. Li, “Upconversion luminescent materials: Advances and applications,” *Chemical reviews*, vol. 115, no. 1, pp. 395–465, 2014.

[191] K. Miyata, F. S. Conrad-Burton, F. L. Geyer, and X.-Y Zhu, “Triplet Pair States in Singlet Fission,” 2019.

[192] A. Endo, M. Ogasawara, A. Takahashi, D. Yokoyama, Y. Kato, and C. Adachi, “Thermally activated delayed fluorescence from sn4+–porphyrin complexes and their application to organic light emitting diodes—a novel mechanism for electroluminescence,” *Advanced Materials*, vol. 21, no. 47, pp. 4802–4806, 2009.

[193] J. C. Goldschmidt and S. Fischer, “Upconversion for photovoltaics–a review of materials, devices and concepts for performance enhancement,” *Advanced Optical Materials*, vol. 3, no. 4, pp. 510–535, 2015.

[194] T. S. Atabaev and A. Molkenova, “Upconversion optical nanomaterials applied for photocatalysis and photovoltaics: Recent advances and perspectives,” *Frontiers of Materials Science*, vol. 13, no. 4, pp. 335–341, 2019.

[195] Y. Sasaki *et al.*, “Near-infrared optogenetic genome engineering based on photon-upconversion hydrogels,” *Angewandte Chemie International Edition*, vol. 58, no. 49, pp. 17 827–17 833, 2019.

[196] J. Zhao, S. Ji, and H. Guo, “Triplet–triplet annihilation based upconversion: From triplet sensitizers and triplet acceptors to upconversion quantum yields,” *Rsc Advances*, vol. 1, no. 6, pp. 937–950, 2011.

[197] R. Haruki, Y. Sasaki, K. Masutani, N. Yanai, and N. Kimizuka, “Leaping across the visible range: Near-infrared-to-violet photon upconversion employing a silyl-substituted anthracene,” *Chemical Communications*, 2020.

[198] N. Nishimura *et al.*, “Photon upconversion from near-infrared to blue light with tips-anthracene as an efficient triplet–triplet annihilator,” *ACS Materials Letters*, vol. 1, no. 6, pp. 660–664, 2019.

[199] K. Börjesson, P. Rudquist, V. Gray, and K. Moth-Poulsen, “Photon upconversion with directed emission,” *Nature communications*, vol. 7, p. 12 689, 2016.

[200] V. Gray *et al.*, “Photophysical characterization of the 9, 10-disubstituted anthracene chromophore and its applications in triplet–triplet annihilation photon upconversion,” *Journal of Materials Chemistry C*, vol. 3, no. 42, pp. 11 111–11 121, 2015.

[201] N. Yanai, M. Kozue, S. Amemori, R. Kabe, C. Adachi, and N. Kimizuka, “Increased vis-to-uv upconversion performance by energy level matching between a tadf donor and high triplet energy acceptors,” *Journal of Materials Chemistry C*, vol. 4, no. 27, pp. 6447–6451, 2016.

[202] W. Zhao and F. N. Castellano, “Upconverted emission from pyrene and di-tert-butylpyrene using ir(ppy)3 as triplet sensitizer,” *The Journal of Physical Chemistry A*, vol. 110, no. 40, pp. 11 440–11 445, 2006.

[203] V. Gray *et al.*, “Cds/zns core–shell nanocrystal photosensitizers for visible to uv upconversion,” *Chemical science*, vol. 8, no. 8, pp. 5488–5496, 2017.

[204] T. N. Singh-Rachford and F. N. Castellano, “Low power visible-to-uv upconversion,” *The Journal of Physical Chemistry A*, vol. 113, no. 20, pp. 5912–5917, 2009.

[205] L. Huang, E. Kakadiaris, T. Vaneckova, K. Huang, M. Vaculovicova, and G. Han, “Designing next generation of photon upconversion: Recent advances in organic triplet-triplet annihilation upconversion nanoparticles,” *Biomaterials*, 2019.

[206] A. B. Pun, L. M. Campos, and D. N. Congreve, “Tunable emission from triplet fusion upconversion in diketopyrrolopyrroles,” *Journal of the American Chemical Society*, vol. 141, no. 9, pp. 3777–3781, 2019.

[207] B. Ehrler *et al.*, “In situ measurement of exciton energy in hybrid singlet-fission solar cells,” *Nature communications*, vol. 3, p. 1019, 2012.

[208] A Völcker, H.-J. Adick, R Schmidt, and H.-D. Brauer, “Near-infrared phosphorescence emission of compounds with low-lying triplet states,” *Chemical physics letters*, vol. 159, no. 1, pp. 103–108, 1989.

[209] M. Fagnoni, “Modern Molecular Photochemistry of Organic Molecules. Von Nicholas J. Turro, V. Ramamurthy und Juan C. Scaiano.,” *Angewandte Chemie*, vol. 122, no. 38, pp. 6859–6860, 2010.

[210] J. Hachmann, J. J. Dorando, M. Avilés, and G. K.-L. Chan, “The radical character of the acenes: A density matrix renormalization group study,” *J. Chem. Phys.*, vol. 127, no. 13, p. 134 309, 2007.

[211] M. Rosenberg, C. Dahlstrand, K. Kilsa, and H. Ottosson, “Excited state aromaticity and antiaromaticity: Opportunities for photophysical and photochemical rationalizations,” *Chemical reviews*, vol. 114, no. 10, pp. 5379–5425, 2014.

[212] M. Bendikov, H. M. Duong, K. Starkey, K. Houk, E. A. Carter, and F. Wudl, “Oligoacenes: Theoretical prediction of open-shell singlet diradical ground states,” *J. Am. Chem. Soc.*, vol. 126, no. 24, pp. 7416–7417, 2004.

[213] A. Dreuw and M. Head-Gordon, “Single-reference ab initio methods for the calculation of excited states of large molecules,” *Chemical reviews*, vol. 105, no. 11, pp. 4009–4037, 2005.

[214] J. Lee and M. Head-Gordon, “Two single-reference approaches to singlet biradicaloid problems: Complex, restricted orbitals and approximate spin-projection combined with regularized orbital-optimized møller-plesset perturbation theory,” *The Journal of Chemical Physics*, vol. 150, no. 24, p. 244 106, 2019.

[215] Y. A. Bernard, Y. Shao, and A. I. Krylov, “General formulation of spin-flip time-dependent density functional theory using non-collinear kernels: Theory, implementation, and benchmarks,” *J. Chem. Phys.*, vol. 136, no. 20, p. 204 103, 2012.

[216] L. V. Slipchenko and A. I. Krylov, “Singlet-triplet gaps in diradicals by the spin-flip approach: A benchmark study,” *J. Chem. Phys.*, vol. 117, no. 10, pp. 4694–4708, 2002. eprint: `https://doi.org/10.1063/1.1498819`.

[217] P. Sharma, V. Bernales, S. Knecht, D. G. Truhlar, and L. Gagliardi, “Density matrix renormalization group pair-density functional theory (dmrg-pdft): Singlet-triplet gaps in polyacenes and polyacetylenes,” *Chem. Sci.*, vol. 10, pp. 1716–1723, 6 2019.

[218] P. M. Zimmerman, “Singlet–triplet gaps through incremental full configuration interaction,” *J. Phys. Chem. A*, vol. 121, no. 24, pp. 4712–4720, 2017.

[219] S. R. Yost and M. Head-Gordon, “Size consistent formulations of the perturb-then-diagonalize møller-plesset perturbation theory correction to non-orthogonal configuration interaction,” *J. Chem. Phys.*, vol. 145, no. 5, p. 054 105, 2016.

[220] Z. Lin and T. Van Voorhis, "Triplet tuning: A novel family of non-empirical exchange–correlation functionals," *Journal of chemical theory and computation*, vol. 15, no. 2, pp. 1226–1241, 2019.

[221] D. G. Liakos, Y. Guo, and F. Neese, "Comprehensive benchmark results for the domain based local pair natural orbital coupled cluster method (dlpno-ccsd (t)) for closed-and open-shell systems," *The Journal of Physical Chemistry A*, 2019.

[222] M. Sparta and F. Neese, "Chemical applications carried out by local pair natural orbital based coupled-cluster methods," *Chemical Society Reviews*, vol. 43, no. 14, pp. 5032–5041, 2014.

[223] J. L. Bao, A. Sand, L. Gagliardi, and D. G. Truhlar, "Correlated-participating-orbitals pair-density functional method and application to multiplet energy splittings of main-group divalent radicals," *J. Chem. Theory Comput.*, vol. 12, no. 9, pp. 4274–4283, 2016.

[224] X. Cui, A. Charaf-Eddin, J. Wang, B. Le Guennic, J. Zhao, and D. Jacquemin, "Perylene-derived triplet acceptors with optimized excited state energy levels for triplet–triplet annihilation assisted upconversion," *The Journal of organic chemistry*, vol. 79, no. 5, pp. 2038–2048, 2014.

[225] A. D. Laurent and D. Jacquemin, "Td-dft benchmarks: A review," *International Journal of Quantum Chemistry*, vol. 113, no. 17, pp. 2019–2039, 2013.

[226] K. Xu, J. Zhao, X. Cui, and J. Ma, "Switching of the triplet–triplet-annihilation upconversion with photoresponsive triplet energy acceptor: Photocontrollable singlet/triplet energy transfer and electron transfer," *The Journal of Physical Chemistry A*, vol. 119, no. 3, pp. 468–481, 2015.

[227] M. J. Peach, M. J. Williamson, and D. J. Tozer, "Influence of triplet instabilities in tddft," *Journal of chemical theory and computation*, vol. 7, no. 11, pp. 3578–3585, 2011.

[228] J. Autschbach and M. Srebro, "Delocalization error and 'functional tuning' in kohn–sham calculations of molecular properties," *Accounts of chemical research*, vol. 47, no. 8, pp. 2592–2602, 2014.

[229] T. Yanai, D. P. Tew, and N. C. Handy, "A new hybrid exchange–correlation functional using the coulomb-attenuating method (cam-b3lyp)," *Chemical Physics Letters*, vol. 393, no. 1-3, pp. 51–57, 2004.

[230] S. Hirata and M. Head-Gordon, "Time-dependent density functional theory within the tamm–dancoff approximation," *Chemical Physics Letters*, vol. 314, no. 3-4, pp. 291–299, 1999.

[231] C. Brueckner and B. Engels, “Benchmarking singlet and triplet excitation energies of molecular semiconductors for singlet fission: Tuning the amount of hf exchange and adjusting local correlation to obtain accurate functionals for singlet–triplet gaps,” *Chemical Physics*, vol. 482, pp. 319–338, 2017.

[232] B. Moore, H. Sun, N. Govind, K. Kowalski, and J. Autschbach, “Charge-transfer versus charge-transfer-like excitations revisited,” *Journal of chemical theory and computation*, vol. 11, no. 7, pp. 3305–3320, 2015.

[233] A. S. Gertsen, M. Koerstz, and K. V. Mikkelsen, “Benchmarking triplet–triplet annihilation photon upconversion schemes,” *Physical Chemistry Chemical Physics*, vol. 20, no. 17, pp. 12 182–12 192, 2018.

[234] S. Zhang, “Ab initio electronic structure calculations by auxiliary-field quantum monte carlo,” *Handbook of Materials Modeling: Methods: Theory and Modeling*, pp. 1–27, 2018.

[235] P. R. Kent *et al.*, “Qmcpack: Advances in the development, efficiency, and application of auxiliary field and real-space variational and diffusion quantum monte carlo,” *The Journal of chemical physics*, vol. 152, no. 17, p. 174 105, 2020.

[236] F. D. Malone, S. Zhang, and M. A. Morales, “Overcoming the memory bottleneck in auxiliary field quantum monte carlo simulations with interpolative separable density fitting,” *Journal of chemical theory and computation*, vol. 15, no. 1, pp. 256–264, 2018.

[237] T. N. Singh-Rachford, R. R. Islangulov, and F. N. Castellano, “Photochemical upconversion approach to broad-band visible light generation,” *The Journal of Physical Chemistry A*, vol. 112, no. 17, pp. 3906–3910, 2008.

[238] J. Yu *et al.*, “2, 1, 3-benzothiadiazole-5, 6-dicarboxylicimide-based polymer semiconductors for organic thin-film transistors and polymer solar cells,” *ACS applied materials & interfaces*, vol. 9, no. 48, pp. 42 167–42 178, 2017.

[239] N. Wang, Z. Chen, W. Wei, and Z. Jiang, “Fluorinated benzothiadiazole-based conjugated polymers for high-performance polymer solar cells without any processing additives or post-treatments,” *Journal of the American Chemical Society*, vol. 135, no. 45, pp. 17 060–17 068, 2013.

[240] B. A. Neto, P. H. Carvalho, and J. R. Correa, “Benzothiadiazole derivatives as fluorescence imaging probes: Beyond classical scaffolds,” *Accounts of chemical research*, vol. 48, no. 6, pp. 1560–1569, 2015.

[241] M. J. Edelmann *et al.*, “Dramatically enhanced fluorescence of heteroaromatic chromophores upon insertion as spacers into oligo (triacetylene) s,” *Helvetica chimica acta*, vol. 85, no. 7, pp. 2195–2213, 2002.

[242] S. Zhang, J. Carlson, and J. E. Gubernatis, "Constrained path monte carlo method for fermion ground states," *Phys. Rev. B*, vol. 55, pp. 7464–7477, 12 Mar. 1997.

[243] W. Purwanto, S. Zhang, and H. Krakauer, "Excited state calculations using phaseless auxiliary-field quantum monte carlo: Potential energy curves of low-lying c 2 singlet states," *The Journal of chemical physics*, vol. 130, no. 9, p. 094 107, 2009.

[244] E. J. Landinez Borda, J. Gomez, and M. A. Morales, "Non-orthogonal multi-slater determinant expansions in auxiliary field quantum monte carlo," *The Journal of chemical physics*, vol. 150, no. 7, p. 074 105, 2019.

[245] E. R. Sayfutyarova and S. Hammes-Schiffer, "Constructing molecular π-orbital active spaces for multireference calculations of conjugated systems," *Journal of chemical theory and computation*, vol. 15, no. 3, pp. 1679–1689, 2019.

[246] R. A. Kendall, T. H. Dunning Jr, and R. J. Harrison, "Electron affinities of the first-row atoms revisited. systematic basis sets and wave functions," *J. Chem. Phys.*, vol. 96, no. 9, pp. 6796–6806, 1992.

[247] Y. Yang, E. R. Davidson, and W. Yang, "Nature of ground and electronic excited states of higher acenes," *Proc. Natl. Acad. Sci.*, vol. 113, no. 35, E5098–E5107, 2016. eprint: `https://www.pnas.org/content/113/35/E5098.full.pdf`.

[248] Z. Fang, Z. Lee, K. A. Peterson, and D. A. Dixon, "Use of improved orbitals for ccsd(t) calculations for predicting heats of formation of group iv and group vi metal oxide monomers and dimers and ucl_6," *J. Chem. Theory Comput.*, vol. 12, no. 8, pp. 3583–3592, 2016.

[249] A. R. Dixon, D. Khuseynov, and A. Sanov, "Benzonitrile: Electron affinity, excited states, and anion solvation," *The Journal of chemical physics*, vol. 143, no. 13, p. 134 306, 2015.

[250] A. Darmanyan, "Experimental study of singlet-triplet energy transfer in liquid solutions," *Chemical physics letters*, vol. 110, no. 1, pp. 89–94, 1984.

[251] R. Ieuji, K. Goushi, and C. Adachi, "Triplet–triplet upconversion enhanced by spin–orbit coupling in organic light-emitting diodes," *Nature communications*, vol. 10, no. 1, pp. 1–10, 2019.

[252] J. D. Warren, V. J. Lee, and R. B. Angier, "Synthesis of 5, 8-dihydroxynaphtho [2, 3-c][1, 2, 5] thiadiazole-6, 9-dione and 6, 9-dihydroxybenzo [g] quinoxaline-5, 10-dione," *Journal of Heterocyclic Chemistry*, vol. 16, no. 8, pp. 1617–1624, 1979.

[253] F. S. Mancilha *et al.*, "Are molecular 5, 8-π-extended quinoxaline derivatives good chromophores for photoluminescence applications?" *European journal of organic chemistry*, vol. 2006, no. 21, pp. 4924–4933, 2006.

[254] K Lopata, R. Reslan, M Kowalska, D. Neuhauser, N. Govind, and K. Kowalski, "Excited-state studies of polyacenes: A comparative picture using eomccsd, cr-eomccsd (t), range-separated (lr/rt)-tddft, td-pm3, and td-zindo," *Journal of chemical theory and computation*, vol. 7, no. 11, pp. 3686–3693, 2011.

[255] S. Baluschev *et al.*, "Two pathways for photon upconversion in model organic compound systems," *Journal of applied physics*, vol. 101, no. 2, p. 023 101, 2007.

[256] V. A. Walters *et al.*, "Electronic structure of triplet states of zinc (ii) tetraphenylporphyrins," *The Journal of Physical Chemistry*, vol. 99, no. 4, pp. 1166–1171, 1995.

[257] Y. Y. Cheng *et al.*, "Entropically driven photochemical upconversion," *The Journal of Physical Chemistry A*, vol. 115, no. 6, pp. 1047–1053, 2011.

[258] J. Isokuortti *et al.*, "Endothermic and exothermic energy transfer made equally efficient for triplet–triplet annihilation upconversion," *The Journal of Physical Chemistry Letters*, vol. 11, no. 1, pp. 318–324, 2019.

[259] S. Tsuzuki and N. Sato, "Origin of attraction in chalgogen–nitrogen interaction of 1, 2, 5-chalcogenadiazole dimers," *The Journal of Physical Chemistry B*, vol. 117, no. 22, pp. 6849–6855, 2013.

[260] C. Gao *et al.*, "Triplet fusion upconversion using sterically protected 9, 10-diphenylanthracene as the emitter," *Physical Chemistry Chemical Physics*, vol. 22, no. 11, pp. 6300–6307, 2020.

[261] F. Deng, J. Blumhoff, and F. N. Castellano, "Annihilation limit of a visible-to-uv photon upconversion composition ascertained from transient absorption kinetics," *The Journal of Physical Chemistry A*, vol. 117, no. 21, pp. 4412–4419, 2013.

[262] P. Duan, N. Yanai, H. Nagatomi, and N. Kimizuka, "Photon upconversion in supramolecular gel matrixes: Spontaneous accumulation of light-harvesting donor–acceptor arrays in nanofibers and acquired air stability," *Journal of the American Chemical Society*, vol. 137, no. 5, pp. 1887–1894, 2015.

[263] X. Wang, R. Tom, X. Liu, D. Congreve, and N. Marom, "An energetics perspective on why there are so few triplet-triplet annihilation emitters," *Journal of Materials Chemistry C*, 2020.

[264] A. Jiménez-Urias *et al.*, "Synthesis and characterization of dumbbell-like btd-based derivatives to engineer organic building blocks in solid-state," *Journal of Molecular Structure*, vol. 1153, pp. 34 –41, 2018.

[265] T. Lei *et al.*, “Aceno[2,1,3]thiadiazoles for field-effect transistors: Synthesis and crystal packing,” *Organic Letters*, vol. 13, no. 10, pp. 2642–2645, 2011, PMID: 21510693. eprint: `https://doi.org/10.1021/ol200748c`.

[266] S. Zeng *et al.*, “D–π–a–π–d type benzothiadiazole–triphenylamine based small molecules containing cyano on the π-bridge for solution-processed organic solar cells with high open-circuit voltage,” *Chem. Commun.*, vol. 48, pp. 10 627–10 629, 86 2012.

[267] Y. Zhao *et al.*, “Direct conversion of benzothiadiazole to benzimidazole: New benzimidazole-derived metal–organic frameworks with adjustable honeycomb-like cavities,” *Chemistry – A European Journal*, vol. 25, no. 20, pp. 5246–5250, 2019. eprint: `https://chemistry-europe.onlinelibrary.wiley.com/doi/pdf/10.1002/chem.201805697`.

[268] M. Schreiber, M. R. Silva-Junior, S. P. Sauer, and W. Thiel, “Benchmarks for electronically excited states: Caspt2, cc2, ccsd, and cc3,” *The Journal of chemical physics*, vol. 128, no. 13, p. 134 110, 2008.

[269] L. Goerigk and S. Grimme, “Assessment of td-dft methods and of various spin scaled cis (d) and cc2 versions for the treatment of low-lying valence excitations of large organic dyes,” *The Journal of Chemical Physics*, vol. 132, no. 18, p. 184 103, 2010.

[270] R. J. Johnson III, *NIST Computational Chemistry Comparison and Benchmark Database*, http://cccbdb.nist.gov/, NIST Standard Reference Database Number 101, Release 17b, September 2015.

[271] H. Sun, Z. Hu, C. Zhong, X. Chen, Z. Sun, and J.-L. Brédas, “Impact of dielectric constant on the singlet–triplet gap in thermally activated delayed fluorescence materials,” *The journal of physical chemistry letters*, vol. 8, no. 11, pp. 2393–2398, 2017.

[272] D. Hashemi, X. Ma, R. Ansari, J. Kim, and J. Kieffer, “Design principles for the energy level tuning in donor/acceptor conjugated polymers,” *Physical Chemistry Chemical Physics*, vol. 21, no. 2, pp. 789–799, 2019.

[273] A. Mahajan and S. Sharma, “Taming the sign problem in auxiliary-field quantum monte carlo using accurate wave functions,” *Journal of Chemical Theory and Computation*, vol. 17, no. 8, pp. 4786–4798, 2021.

[274] S. Zhang, H. Krakauer, W. A. Al-Saidi, and M. Suewattana, “Quantum simulations of realistic systems by auxiliary fields,” *Computer physics communications*, vol. 169, no. 1-3, pp. 394–399, 2005.

[275] J. Lee, F. D. Malone, and D. R. Reichman, “The performance of phaseless auxiliary-field quantum monte carlo on the ground state electronic energy of benzene,” *The Journal of Chemical Physics*, vol. 153, no. 12, p. 126 101, 2020.

[276] W. Purwanto, S. Zhang, and H. Krakauer, “Auxiliary-field quantum monte carlo calculations of the molybdenum dimer,” *J. Chem. Phys.*, vol. 144, no. 24, p. 244 306, 2016.

[277] W. Purwanto, S. Zhang, and H. Krakauer, “An auxiliary-field quantum monte carlo study of the chromium dimer,” *J. Chem. Phys.*, vol. 142, no. 6, p. 064 302, 2015.

[278] A. Mahajan, J. Lee, and S. Sharma, “Selected configuration interaction wave functions in phaseless auxiliary field quantum monte carlo,” *arXiv preprint arXiv:2201.10458*, 2022.

[279] M. Motta and S. Zhang, “Computation of ground-state properties in molecular systems: Back-propagation with auxiliary-field quantum monte carlo,” *J. Chem. Theory Comput.*, vol. 13, no. 11, pp. 5367–5378, 2017.

[280] M. Motta, J. Shee, S. Zhang, and G. K.-L. Chan, “Efficient ab initio auxiliary-field quantum monte carlo calculations in gaussian bases via low-rank tensor decomposition,” *J. Chem. Theory Comput.*, vol. 15, no. 6, pp. 3510–3521, 2019.

[281] J. Lee and D. R. Reichman, “Stochastic resolution-of-the-identity auxiliary-field quantum monte carlo: Scaling reduction without overhead,” *The Journal of Chemical Physics*, vol. 153, no. 4, p. 044 131, 2020.

[282] A. Mahajan and S. Sharma, “Efficient local energy evaluation for multi-slater wave functions in orbital space quantum monte carlo,” *The Journal of Chemical Physics*, vol. 153, no. 19, p. 194 108, 2020.

[283] NVIDIA, *Cuda, release: 11.4.2*, 2021.

[284] D. J. Thouless, “Stability conditions and nuclear rotations in the hartree-fock theory,” *Nucl. Phys.*, vol. 21, pp. 225–232, 1960.

[285] N. Rom, E. Fattal, A. K. Gupta, E. A. Carter, and D. Neuhauser, “Shifted-contour auxiliary-field monte carlo for molecular electronic structure,” *The Journal of chemical physics*, vol. 109, no. 19, pp. 8241–8248, 1998.

[286] T. Booth and J. Gubernatis, “Monte carlo determination of multiple extremal eigenpairs,” *Physical Review E*, vol. 80, no. 4, p. 046 704, 2009.

[287] M. C. Buonaura and S. Sorella, “Numerical study of the two-dimensional heisenberg model using a green function monte carlo technique with a fixed number of walkers,” *Physical Review B*, vol. 57, no. 18, p. 11 446, 1998.

[288] F. Neese, F. Wennmohs, and A. Hansen, “Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method,” *The Journal of chemical physics*, vol. 130, no. 11, p. 114 108, 2009.

[289] J. Foster and S. Boys, "Canonical configurational interaction procedure," *Reviews of Modern Physics*, vol. 32, no. 2, p. 300, 1960.

[290] F. Neese, A. Hansen, F. Wennmohs, and S. Grimme, "Accurate theoretical chemistry with coupled pair models," *Accounts of chemical research*, vol. 42, no. 5, pp. 641–648, 2009.

[291] A. Karton, P. R. Schreiner, and J. M. Martin, "Heats of formation of platonic hydrocarbon cages by means of high-level thermochemical procedures," *Journal of computational chemistry*, vol. 37, no. 1, pp. 49–58, 2016.

[292] C. Riplinger and F. Neese, "An efficient and near linear scaling pair natural orbital based local coupled cluster method," *J. Chem. Phys.*, vol. 138, no. 3, p. 034 106, 2013.

[293] D. Datta, S. Kossmann, and F. Neese, "Analytic energy derivatives for the calculation of the first-order molecular properties using the domain-based local pair-natural orbital coupled-cluster theory," *The Journal of Chemical Physics*, vol. 145, no. 11, p. 114 101, 2016.

[294] H. Chen, G. C. Dismukes, and D. A. Case, "Resolving ambiguous protonation and oxidation states in the oxygen evolving complex of photosystem ii," *The Journal of Physical Chemistry B*, vol. 122, no. 37, pp. 8654–8664, 2018.

[295] J. L. Weber, H. Vuong, R. A. Friesner, and D. R. Reichman, "The design of new practical constraints in auxiliary-field quantum monte carlo," *arXiv preprint arXiv:2306.09207*, 2023.

[296] Y.-Y. He, M. Qin, H. Shi, Z.-Y. Lu, and S. Zhang, "Finite-temperature auxiliary-field quantum monte carlo: Self-consistent constraint and systematic approach to low temperatures," *Physical Review B*, vol. 99, no. 4, p. 045 108, 2019.

[297] M. Qin, "Self-consistent optimization of the trial wave-function in constrained path auxiliary field quantum monte carlo using mixed estimators," *arXiv preprint arXiv:2303.10301*, 2023.

[298] W. J. Huggins, B. A. O'Gorman, N. C. Rubin, D. R. Reichman, R. Babbush, and J. Lee, "Unbiasing fermionic quantum monte carlo with a quantum computer," *Nature*, vol. 603, no. 7901, pp. 416–420, 2022.

[299] J. Lee, H. Q. Pham, and D. R. Reichman, "Twenty years of auxiliary-field quantum monte carlo in quantum chemistry: An overview and assessment on main group chemistry and bond-breaking," *Journal of Chemical Theory and Computation*, vol. 18, no. 12, pp. 7024–7042, 2022.

[300] Z. Sukurma, M. Schlipf, M. Humer, A. Taheridehkordi, and G. Kresse, "Benchmark phaseless auxiliary-field quantum monte carlo method for small molecules," *arXiv preprint arXiv:2303.04256*, 2023.

[301] H. Shi and S. Zhang, “Symmetry in auxiliary-field quantum monte carlo calculations,” *Phys. Rev. B*, vol. 88, p. 125 132, 12 Sep. 2013.

[302] Z.-Y. Xiao, H. Shi, and S. Zhang, “Interfacing branching random walks with metropolis sampling: Constraint release in auxiliary-field quantum monte carlo,” *arXiv preprint arXiv:2305.09575*, 2023.

[303] J. J. Eriksen *et al.*, “The ground state electronic energy of benzene,” *The journal of physical chemistry letters*, vol. 11, no. 20, pp. 8922–8929, 2020.

[304] D. Ceperley and B. Alder, “Quantum monte carlo for molecules: Green’s function and nodal release,” *The Journal of chemical physics*, vol. 81, no. 12, pp. 5833–5844, 1984.