

Department of Quantum Physics
Faculty of Physics, Astronomy and Informatics
Nicolaus Copernicus University in Toruń

Advancement of the pair coupled cluster doubles-based molecular property calculations with quantum embedding

Rahul Chakraborty

Dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Physics

Thesis supervised by dr. hab. Paweł Tecmer, Prof. UMK
Department of Quantum Physics
Institute of Physics, Nicolaus Copernicus University in Toruń, Poland



**NICOLAUS COPERNICUS
UNIVERSITY
IN TORUŃ**

Faculty of Physics, Astronomy
and Informatics

Toruń 2026

Abstract

The works included in this thesis aim to extend the applications of pair-coupled cluster doubles (pCCD) theory. pCCD produces reliable results for strongly correlated chemical systems and performs better than other traditional quantum chemical methods. Still, achieving greater accuracy requires post-pCCD corrections, which steeply increase computational cost, limiting its widespread application. Quantum embedding approaches, which reserve accurate but high-cost methods for the active site of chemical change and approximated methods for other regions of a chemical system, can be used to circumvent this problem. In this context, two new pCCD-based embedding formulations have been implemented and tested. Both the pCCD-in-DFT and pCCD-in-pCCD methods produce reliable results across a wide range of intermolecular interactions. In particular, the latter method shows significantly better performance, as it does not depend on approximated density functionals. Additionally, estimation of molecular properties is also pertinent for practical applications with pCCD. As the traditional response-theory-based formulation for pCCD analyzed here did not meet the required level of accuracy, an alternate approach based on the expectation-value method was implemented for pCCD and post-pCCD methods. This approach produces excellent agreement with the references, at a much lower computational cost, as we bypass the need to solve the response equations. Thus, the pCCD-in-pCCD embedding method and the expectation-value-pCCD approach provide routes for large-scale simulations of challenging, complex systems.

Streszczenie

Badania naukowe przedstawione w niniejszej rozprawie doktorskiej mają na celu rozszerzenie zakresu zastosowań teorii opartych na metodzie sprzężonych klastrów z podwójnie wzbudzonymi parami – pCCD (ang. pair Coupled Cluster Doubles). Metoda pCCD dostarcza wiarygodnych wyników dla układów chemicznych charakteryzujących się silną korelacją elektronową i wykazuje wyższą wydajność obliczeniową niż wiele tradycyjnych metod chemii kwantowej. Jednak osiągnięcie wyższej dokładności wymaga wprowadzenia poprawek do modelu pCCD, co znacząco zwiększa nakład obliczeniowy i ogranicza jej powszechne zastosowanie w praktyce. Aby przezwyciężyć te ograniczenia, w pracy zastosowano podejścia oparte na osadzaniu kwantowym (ang. quantum embedding). Strategia ta polega na rezerwowaniu dokładnych, lecz kosztownych obliczeniowo metod dla aktywnego obszaru układu kwantowo-chemicznego, podczas gdy pozostałe części opisywane są za pomocą metod bardziej przybliżonych. W ramach niniejszej rozprawy zaimplementowano i przetestowano dwie nowe formułacje osadzania oparte na pCCD: metodę pCCD-in-DFT oraz pCCD-in-pCCD. Obie metody zapewniają wiarygodne wyniki w szerokim zakresie oddziaływań międzycząsteczkowych. W szczególności metoda pCCD-in-pCCD wykazuje znacznie lepszą wiarygodność do modelowania dużych układów, ponieważ nie opiera się na przybliżonych funkcjonalach gęstości elektronowej. Ponadto w pracy przeanalizowano obliczanie właściwości molekularnych w ramach metody pCCD. Ponieważ tradycyjne podejście oparte na teorii odpowiedzi (równania Lambda) nie spełniło oczekiwań pod względem dokładności, wdrożono alternatywne podejście wykorzystujące wartości oczekiwane dla metod pCCD i post-pCCD. Nowe podejście zapewnia doskonałą zgodność z danymi referencyjnymi przy istotnie niższych kosztach obliczeniowych, ponieważ eliminuje konieczność rozwiązywania równań odpowiedzi. W rezultacie metoda osadzania oparta na schemacie pCCD-in-pCCD oraz podejście oparte na wartościach oczekiwanych stanowią obiecujące ścieżki umożliwiające symulacje na dużą skalę trudnych, silnie skorelowanych układów chemicznych.

Dedication

To the people of India and Poland, whose taxes fund public education and scientific research.

Acknowledgements

I would like to express my gratitude for my doctoral supervisor, Prof. Paweł Tecmer, for his training, guidance, and motivation. I would especially acknowledge the freedom Prof. Tecmer gave me, which allowed me to explore my ideas, make mistakes, learn, unlearn, and repeat. I am also extremely grateful to Prof. Katharina Boguslawski, head of the PyBEST research group, for her guidance and support. I also had the privilege to scientifically collaborate with Dr. Matheus Morato F. de Moraes and Dr. Somayeh Ahmadkhani, former post-docs in our group, and from both of whom I learnt a ton. I would like to acknowledge the NCU University Center of Excellence for the financial support for my research. I also thank Prof. Wiesław Nowak and Agnieszka Górską-Pukownik from the AST doctoral school for all their kind help in conducting my research. I also thank Prof. Piotr Żuchowski and Joanna Kozłowska from the quantum physics department for their kind help on several occasions. I thank Dr. Artur Nowak and Dr. Aleksandra Leszczyk for all their help in my initial days in our research group.

Outside NCU, I must acknowledge my mentors in Algorithmiq, Dr. Adam Glos and Dr. Özlem Salehi, from whom I learnt a lot in the domain of quantum computing, which was completely new to me. I also owe a lot to Prof. Debashree Ghosh, with whom I spent my early research days at the National Chemical Laboratory, India, and who shaped my scientific thinking.

During this long and tenacious PhD journey, I was fortunate to have the support of my friends and family, near and far. First of all, I thank my partner, Sreejita Das, who gave me the courage to come back to academics again. Without her, probably this would not have been possible. I would also like to mention my mother, Keya Chakraborty, and my sister Priyanka Roy Chowdhury, who have been constant supports in my life. I was also lucky to have my friends cum colleagues Arameh Jahani, Saman Behjou, Julian Świerszyński, Aditi Singh, Gabriel Chirchir, Ram Dhari Pandey, Zahra Karimi, Matheus Witkowski, Emil Sujkowski, Vignesh Kumar, Garima Singh, Debleena Mandal, Dr. Asish Gupta, Dr. Justyna Gupta, and Dr. Gregory Gondek. Without them, my workdays, weekends, and conference

trips would not have been memories worth reliving. Also, to mention Dr. Sanjana Dutt and Diksha Singh, who were an integral part of my life in Toruń. A very special thanks to Dr. Marta Gałynska, who has been kind enough to act as my counselor, Polish translator, and go-to person to discuss any nonsensical research idea. A special shout-out to my friends from India, Dr. Adhidev Biswas, Soimantik Bhattacharya, Srijit Bhattacharyya, Dr. Parichita Saha, Subhanjan Ghosh, Dr. Samik Bose, Dr. Diksha Dhawan, Dr. Paulami Ghosh, Dr. Sayantan Acharya, and Dr. Amit Kumawat, who, despite being in different parts of India and the world, have stuck with me for all these years. With the risk of sounding melodramatic, I would like to mention the book ‘Biswasghatak’ (The Traitor) by Narayan Sanyal about the Manhattan Project. Perhaps every Bengali like me who has read the book has, at some point in life, wanted to become a rebellious physicist—just like Klaus Fuchs.

Last but not least, I would like to thank the local people of the lovely city of Toruń, who made me feel welcome and made my time here enjoyable.

Contents

1	Introduction	1
1.1	Premise	1
1.2	Theoretical background	3
1.2.1	Seniority number in quantum chemistry	6
1.2.2	Geminal-based methods	9
1.2.3	The pCCD ansatz	11
1.3	Dynamic correlation corrections on pCCD	15
1.3.1	Coupled cluster-based dynamic correlation corrections on pCCD	15
1.4	Excited states with pCCD	21
1.4.1	Modeling electron-pair excited states with EOM-pCCD	23
1.4.2	EOM-pCCD extensions for singly excited states	24
1.4.3	EOM-pCCD for general doubly excited states	25
1.5	Molecular property calculations with pCCD-based methods	28
1.5.1	Response approach for pCCD and post-pCCD methods	29
1.5.2	Expectation value based approach for pCCD and post-pCCD methods	33
1.6	Quantum chemical embedding methods	35
2	Summary of the publications	41
2.1	Published articles comprising the thesis	41
2.2	Other published articles	45
2.3	Author contribution statement	50
3	Conclusions	54
3.1	Summary of key findings	54
3.2	Outlook and future Work	60
	Appendix A Derivation of XCCSD density matrix	80
	Appendix B Publications	88
B.1	Static embedding with pair coupled cluster doubles based methods	89

B.2	Supporting information for: Static embedding with pair coupled cluster doubles based methods	101
B.3	Toward Reliable Dipole Moments without Single Excitations: The Role of Orbital Rotations and Dynamical Correlation	112
B.4	Supporting information for: Toward reliable dipole moments without single excitations: The role of orbital rotations and dynamical correlation	126
B.5	Expectation value-pCCD-based methods for single-electron properties	137
B.6	Supporting information for: Expectation value-pCCD-based methods for single-electron properties	157
B.7	Frozen density embedding with pCCD electron densities	173
B.8	Supporting information for: Frozen density embedding with pCCD electron densities	183

List of Figures

3.1	Spin-free vertical excitation energies of $\text{UO}_2\text{F}_4^{2-}$ obtained with EOM-fpLCCSD. UO_2^{2+} -in-pc and UO_2^{2+} -in-DFT indicate that the F are replaced by point-charges and embedding potential obtained from DFT-in-DFT, respectively. Schematic diagram on the left side for visual representation.	55
3.2	Pictorial description of FDE freeze-and-thaw cycles with pCCD densities. The green block represents the initial setup within the isolated subsystem environment. The red block signifies the final stage, where the converged pCCD wave function from the active subsystem is used to perform post-pCCD calculations, including fpCC, fpLCC, and subsequent EOM-fpLCCSD.	59

List of Tables

1.1	Details of Coupled Cluster (CC) models for dynamic correlation correction.	20
1.2	Details of Equation-of-Motion Pair Coupled Cluster Doubles (EOM-pCCD)-based excited-state methods.	27
2.1	Author contributions for the article Static embedding with pair coupled cluster doubles based methods , <i>Phys. Chem. Chem. Phys.</i> , 2023, 25, 25377 (DOI: 10.1039/d3cp02502k)	50
2.2	Author contributions for the article Toward Reliable Dipole Moments without Single Excitations: The Role of Orbital Rotations and Dynamical Correlation , <i>J. Chem. Theory Comput.</i> 2024, 20, 4689–4702, DOI: 10.1021/acs.jctc.4c00471	51
2.3	Author contributions for the article Expectation Value-pCCD-Based Methods for Single-Electron Properties , <i>J. Phys. Chem. A</i> , 2025, 129, 6713–6732, DOI: 10.1021/acs.jpca.5c03859.	52
2.4	Author contributions for the article Frozen density embedding with pCCD electron densities , arXiv:2604.14904, 2026, DOI: 10.48550/arXiv.2604.14904.	53

List of Abbreviations

1-ORDM	One-Orbital Reduced Density Matrix
1-RDM	One-Particle Reduced Density Matrix
2-ORDM	Two-Orbital Reduced Density Matrix
2-RDM	Two-Particle Reduced Density Matrix
3-RDM	Three-Particle Reduced Density Matrix
4-RDM	Four-Particle Reduced Density Matrix
AP1roG	Antisymmetric Product of 1-reference Orbital Geminals
APIG	Antisymmetric Product of Interacting Geminals
APSG	Antisymmetric Product of Strongly Orthogonal Geminals
BCH	Baker–Campbell–Hausdorff
CAS	Complete Active Space
CASCI	Complete Active Space Configuration Interaction
CASSCF	Complete Active Space Self-Consistent Field
CC	Coupled Cluster
CCD	Coupled Cluster Doubles
CCS	Coupled Cluster Singles
CCSD	Coupled Cluster Singles and Doubles
CCSD(T)	Coupled Cluster Singles and Doubles with Perturbative Triples
CCSDT	Coupled Cluster Singles, Doubles, and Triples
CCSDTQ	Coupled Cluster with Single, Double, Triple, and Quadruple
CI	Configuration Interaction
CI(D)	Configuration Interaction with perturbative Doubles
CID	Configuration Interaction Doubles
CIS	Configuration Interaction Singles
CISD	Configuration Interaction Singles and Doubles
CR-EOM-CCSD(T)	completely renormalized Equation-of-Motion Coupled Cluster Sin- gles and Doubles with perturbative Triples

DFT	Density Functional Theory
DKH	Douglas–Kroll–Hess
DMRG	Density Matrix Renormalization Group
DOCI	Doubly Occupied Configuration Interaction
EFP	Effective Fragment Potential
EOM	Equation-of-Motion
EOM-CC	Equation-of-Motion Coupled Cluster
EOM-CCSD	Equation-of-Motion Coupled Cluster Singles and Doubles
EOM-fpLCCSD	Equation-of-Motion Frozen Pair Linearized Coupled Cluster Singles and Doubles
EOM-pCCD	Equation-of-Motion Pair Coupled Cluster Doubles
EOM-pCCD-tCCSD	Equation-of-Motion pair Coupled Cluster Doubles tailored Coupled Cluster Singles and Doubles
FCI	Full Configuration Interaction
FDE	Frozen Density Embedding
fpCC	Frozen Pair Coupled Cluster
fpCCD	Frozen Pair Coupled Cluster Doubles
fpCCSD	Frozen Pair Coupled Cluster Singles and Doubles
fpLCC	Frozen Pair Linearized Coupled Cluster
fpLCCD	Frozen Pair Linearized Coupled Cluster Doubles
fpLCCSD	Frozen Pair Linearized Coupled Cluster Singles and Doubles
FSCC	Fock-space Coupled Cluster
HF	Hartree–Fock
IHFSCC	Intermediate Hamiltonian Fock-Space Coupled Cluster
LCC	Linearized Coupled Cluster
LR	Linear Response
LR-pCCD	Linear Response Pair Coupled Cluster Doubles
MO	Molecular Orbital
MP2	Second-Order Møller–Plesset Perturbation Theory
MRCI-SD	Multi-Reference Configuration Interaction with Single and Double
MRMP	multi-reference Møller–Plesset
oo-pCCD	Orbital-Optimized Pair Coupled Cluster Doubles
pCCD	Pair Coupled Cluster Doubles

PES	Potential Energy Surface
PNOF5	Piris Natural Orbital Functional 5
SVD	Singular Value Decomposition
tCC	Tailored Coupled Cluster
VEE	Vertical Excitation Energy
WFT	Wave Function Theory

Chapter 1

Introduction

1.1 Premise

Quantum chemistry, in today’s context, is primarily a search for a balance of accuracy and computational resource management. Electronic structure theory of chemical systems, in particular, has undergone a rapid growth over the last few decades, pushing the boundaries of theoretical and computational research toward real-world problems. The formulation of sophisticated methodologies has kept pace with the rapid evolution of computer hardware and algorithms. Deeper insights into the electronic structure of a chemical species can be gained now with the computation of spectroscopic parameters, reaction pathways, and other quantities related to the energetics of molecules, and also the response of atoms and molecules to external perturbations. Theories of electronic structure and dynamics are now routinely used in understanding experimental findings and have, in several instances, encouraged new experiments.

Solving the Schrödinger equation is the central objective of electronic structure theory, yet the impossibility of an analytical solution for many-body Hamiltonians necessitates a reliance on approximations. The exact Full Configuration Interaction (FCI) method is rarely viable for larger molecular systems, driving the ongoing search for theoretical frameworks that balance accuracy with computational feasibility. Thus, a hierarchy of theoretical methods has evolved in the last few decades with varying levels of fidelity and computational demands. While WFT-based methods like CC, Complete Active Space Self-Consistent Field (CASSCF) offer excellent reliability, they are constrained by their computational cost. Density Functional Theory (DFT), while computationally efficient, is limited by the approximations employed in the exchange–correlation functional. The stark contrast in computational demands between these methods suggests that treating an entire complex system at a high

level of theory is often impractical. However, in complex systems, especially in the condensed phase, our region of interest, where the ‘chemistry’ is taking place, is only a small part of a much larger structure. This observation underscores the ground for the development of quantum embedding methods, which partition a system to treat a chemically crucial fragment with a high-cost approach (such as Complete Active Space (CAS) or CC methods) while describing the surrounding environment with a more affordable, low-level theory like DFT or Hartree–Fock (HF).

In this thesis, novel quantum embedding strategies based on a new family of Wave Function Theory (WFT) methods, namely Pair Coupled Cluster Doubles (pCCD) and its several extensions, have been discussed. By its construction, the pCCD ansatz relies on electron-pair excitations, effectively focusing on the seniority-zero sector of the Hilbert space, which often provides the leading contribution to the correlation energy in strongly correlated systems. Thus, pCCD with its mean-field-like computational scaling, facilitates reliable simulation of strongly correlated systems, hitherto difficult to access by other standard variants of the CC theory. By integrating pCCD-based methods in quantum embedding frameworks, their applicability can be extended to even more challenging systems. The implementation and performance of a pCCD-in-DFT embedding approach have been shown. While this method showed reliable performance for the tested systems, any embedding approach relying on DFT is not free from the inherent limitations of approximated density functionals. Hence, a newer approach has been recently implemented where pCCD electron densities, which are computationally cheap to obtain, have been used to simulate the effect of the environment on our fragment of interest. This pCCD-in-pCCD embedding provides excellent agreement with the theoretical references for the tested systems. Alongside that, newer approaches for accessing one-electron properties, such as dipole and quadrupole moments, with pCCD-based methods have been discussed. Specifically, the expectation value-based property estimation using $\hat{S}$ -operator formalism has been implemented for pCCD, which enables the construction of density matrices with the already available cluster amplitudes, bypassing the traditional way of solving the computationally expensive $\hat{\Lambda}$ -equations.

Overall, the developments presented in this thesis demonstrate that pCCD-based methodologies can be systematically extended, enabling both efficient treatment of strongly correlated systems and the reliable evaluation of molecular properties. The proposed embedding frameworks and property formulations highlight the versatility of pCCD as a practical tool for large-scale simulations. These advances open avenues for further applications of pCCD-based methods in complex chemical environments and provide a foundation for future methodological improvements.

1.2 Theoretical background

The importance of an accurate treatment of the electron correlation for studying chemical systems is already well established [1, 2]. Electron correlation can be defined as the difference between the exact non-relativistic energy of the system and the energy calculated with the HF approximation, at the theoretical limit of an infinite basis set. Electron correlation is primarily classified into two types: static/strong/non-dynamic and dynamic/weak. Static/strong/non-dynamic correlation arises from near-degeneracies of electronic energy levels and must be accounted for to obtain a correct zeroth-order description that a single Slater determinant method, such as HF, can not capture. Whereas dynamic/weak correlation arises due to the instantaneous, short-range repulsive interactions between electrons as they move to avoid each other. Siegbahn [3] described the origins of the two different correlation effects as- “In many situations, it is further convenient to subdivide the correlation energy into two parts with different physical origins. For chemical reactions where bonds are broken and formed, and for most excited states, the major part of the correlation energy can be obtained by adding only a few extra configurations besides the Hartree–Fock configuration. This part of the correlation energy is due to near degeneracy between different configurations and has its origin quite often in artifacts of the Hartree–Fock approximation. The physical origin of the second part of the correlation energy is the dynamical correlation of the motion of the electrons and is therefore sometimes called the dynamical correlation energy. Since the Hamiltonian operator contains only one- and two-particle operators, this part of the correlation energy can be very well described by single and double replacements from the leading, near-degenerate, reference configurations.”

A balanced inclusion of both is needed in any quantum chemical study in order to arrive at reliable results. Generally, for quantum chemical methods, strong correlation is the primary target as it describes the essential qualitative aspects of the electronic structure of the chemical system. To qualitatively describe systems such as radicals, transition metal complexes, bond-dissociation processes, etc., any theoretical framework must explicitly account for strong correlation [4].

In Configuration Interaction (CI) theory [5], the correlated electronic wave function is expressed as a linear expansion over Slater determinants $\{|\Phi_n\rangle\}$ constructed from a reference orbital basis.

$$|\Psi_{\text{CI}}\rangle = c_0 |\Phi_0\rangle + \sum_n c_n |\Phi_n\rangle. \quad (1.1)$$

For an N -electron system, the CI expansion can equivalently be written in operator form as

$$|\Psi_{\text{CI}}\rangle = \left(c_0 + \hat{C}_1 + \hat{C}_2 + \cdots + \hat{C}_n \right) |\Phi_0\rangle, \quad n < N. \quad (1.2)$$

Here, $\hat{C}_n$ is the excitation operator that generates n -tuple excitations from the reference determinant $|\Phi_0\rangle$, and determinants up to the n -tuple excitation level are included. If all possible excitations are included, the CI expansion becomes the FCI wave function,

$$|\Psi_{\text{FCI}}\rangle = \left(c_0 + \hat{C}_1 + \hat{C}_2 + \hat{C}_3 + \cdots + \hat{C}_N \right) |\Phi_0\rangle. \quad (1.3)$$

The FCI wave function provides the exact solution of the on-relativistic electronic Schrödinger equation within a given one-electron basis. However, the factorial scaling of CI methods with system size renders them computationally infeasible for large systems. Interestingly, in most chemically relevant cases, the FCI expansion is effectively sparse in the determinant basis, as only a small subset of configurations contributes significantly to the wave function. This limitation motivates the development of more compact wave function parametrizations.

Generally, the CI expansion is truncated at some level of excitations, such as Configuration Interaction Singles and Doubles (CISD). However, this truncation introduces important theoretical deficiencies in the structure of the CI wave function. Size-extensivity refers to the correct scaling of the total energy with system size, such that for two non-interacting subsystems A and B , the energy is additive, $E(A+B) = E(A) + E(B)$. Size-consistency, closely related, requires that the method yield the correct dissociation limit, i.e., the energy of a combined system at infinite separation equals the sum of the energies of the individual fragments computed at the same level of theory. In particular, truncated CI methods violate both of these formal properties, leading to incorrect scaling of the total energy with system size and an incorrect description of dissociation into non-interacting fragments.

One of the most successful alternatives to the CI approach is the CC theory [6, 7], which replaces the linear CI expansion with an exponential ansatz,

$$|\Psi_{\text{CC}}\rangle = e^{\hat{T}} |\Phi_0\rangle, \quad (1.4)$$

where the cluster operator is defined as

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots. \quad (1.5)$$

In second quantization, the cluster operators for the different excitation levels can be expressed as

$$\hat{T}_1 = \sum_{i,a} t_i^a \hat{a}_a^\dagger \hat{a}_i, \quad (1.6)$$

$$\hat{T}_2 = \frac{1}{4} \sum_{i,j,a,b} t_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i. \quad (1.7)$$

and so on, where $i, j, k, \dots$ denote occupied orbitals and $a, b, c, \dots$ denote virtual orbitals. Here $\hat{a}_p^\dagger$ and $\hat{a}_p$ are the fermionic creation and annihilation operators, respectively. The creation operator $\hat{a}_p^\dagger$ creates an electron in spin-orbital p , while the annihilation operator $\hat{a}_p$ removes an electron from spin-orbital p . These operators satisfy the canonical anticommutation relations

$$\{\hat{a}_p, \hat{a}_q^\dagger\} = \delta_{pq}, \quad \{\hat{a}_p^\dagger, \hat{a}_q^\dagger\} = 0, \quad \{\hat{a}_p, \hat{a}_q\} = 0. \quad (1.8)$$

In contrast to truncated CI, CC methods satisfy both size-extensivity and size-consistency properties due to their exponential ansatz, which ensures proper factorization and additive scaling for non-interacting systems. Unlike CI, the exponential form generates higher-order excitations implicitly through the repeated action of the cluster excitation operators, allowing CC to efficiently approximate the 'glscfci wave function. In the limit of a complete excitation manifold,

$$e^{\hat{T}} |\Phi_0\rangle \equiv |\Psi_{\text{FCI}}\rangle, \quad (1.9)$$

i.e., CC theory becomes exact.

In practice, the expansion is truncated (e.g., Coupled Cluster Singles and Doubles (CCSD) with $\hat{T} \approx \hat{T}_1 + \hat{T}_2$) for computational limitations. Still, highly popular CC approaches such as CCSD and Coupled Cluster Singles and Doubles with Perturbative Triples (CCSD(T)) provide highly accurate results for many chemical systems [8].

However, single-reference CC theory relies on how good its single Slater determinant reference is and may fail when the wave function expansion becomes strongly multi-configurational, i.e., in the presence of strong/static correlation. This limitation motivates the search for alternative approaches, specifically curated to model static correlation, which will be discussed in the following sections.

1.2.1 Seniority number in quantum chemistry

The seniority number concept is quite well known in the nuclear physics field [9, 10]. In the case of fermionic systems, the seniority number counts the number of unpaired particles in a Slater determinant; equivalently, it is the number of singly occupied orbitals. A seniority zero determinant means all electrons are paired; in seniority two, we have two unpaired electrons by breaking one pair of electrons; in seniority four, two pairs are broken, and we get four unpaired electrons, so on and so forth. In nuclear physics, this concept, rooted in pairing correlations, is used to reduce the dimensionality of the nuclear shell-model calculations [11]. Similarly, in quantum chemistry, determinants in a multi-determinant wave function can likewise be organized by seniority, providing a convenient way to partition the Hilbert space.

We derive an expression for the size of each seniority sector and show that their sum reproduces the FCI dimension following Byatutas et al [12]. Consider a system with M spatial orbitals and N electrons of each spin (α, β) , assuming an even number of electrons. We work in the basis of Slater determinants and ignore spatial symmetry. The dimension of the FCI space is

$$D_{\text{FCI}} = \binom{M}{N}^2, \quad (1.10)$$

since α - and β -spin electrons can be independently distributed among M orbitals.

To partition the Hilbert space by seniority, we fix the number of unpaired electrons. For seniority $2k$, there are k unpaired α electrons and k unpaired β electrons. The α electrons can be arranged in $\binom{M}{N}$ ways. Among these, k electrons are chosen to remain unpaired, giving a factor $\binom{N}{k}$. The corresponding k β electrons must occupy orbitals not already occupied by α electrons, yielding $\binom{M-N}{k}$ possibilities.

Thus, the dimension of the seniority- $2k$ sector is

$$D_k = \binom{M}{N} \binom{N}{k} \binom{M-N}{k}. \quad (1.11)$$

Summing over all k ,

$$D = \sum_{k=0}^N \binom{M}{N} \binom{N}{k} \binom{M-N}{k}. \quad (1.12)$$

Using Vandermonde's identity

$$\binom{m+n}{P} = \sum_{r=0}^P \binom{m}{r} \binom{n}{P-r}, \quad (1.13)$$

we get

$$\binom{M}{N} = \sum_{r=0}^N \binom{N}{r} \binom{M-N}{N-r}. \quad (1.14)$$

Using the change of summation index $r = N - k$, we obtain

$$\sum_{k=0}^N \binom{N}{k} \binom{M-N}{k} = \sum_{k=0}^N \binom{N}{N-k} \binom{M-N}{k} = \binom{M}{N}. \quad (1.15)$$

Finally, we recover

$$D = \binom{M}{N}^2 = D_{\text{FCI}}. \quad (1.16)$$

We now consider the special case of seniority zero. In this case, all electrons are paired, i.e., each occupied spatial orbital contains both an α and a β electron. Consequently, the many-electron basis reduces to choosing which N spatial orbitals are doubly occupied out of the M available orbitals. The dimension of the seniority-zero space is therefore

$$D_0 = \binom{M}{N}. \quad (1.17)$$

Comparing with the FCI dimension from Eq. (1.10), we immediately observe that the seniority-zero space is parametrically much smaller. Specifically, the ratio of the two dimensions is $\frac{D_0}{D_{\text{FCI}}} = \frac{1}{\binom{M}{N}}$, which decreases combinatorially with system size.

To make the comparison of the dimensions more explicit, we analyze using Stirling's approximation. For large n ,

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n. \quad (1.18)$$

Applying this to the binomial coefficient $\binom{M}{N} = \frac{M!}{N!(M-N)!}$, we obtain

$$\begin{aligned} \binom{M}{N} &\approx \frac{\sqrt{2\pi M} \left(\frac{M}{e}\right)^M}{\sqrt{2\pi N} \left(\frac{N}{e}\right)^N \sqrt{2\pi(M-N)} \left(\frac{M-N}{e}\right)^{M-N}} \\ &= \frac{1}{\sqrt{2\pi M x(1-x)}} \exp[M \ln M - N \ln N - (M-N) \ln(M-N)], \end{aligned} \quad (1.19)$$

where $x = N/M$. Rewriting the exponent in terms of x ,

$$M \ln M - N \ln N - (M-N) \ln(M-N) = M[-x \ln x - (1-x) \ln(1-x)], \quad (1.20)$$

we arrive at

$$\binom{M}{N} \approx \frac{1}{\sqrt{2\pi M x(1-x)}} \exp[MH(x)], \quad (1.21)$$

where $H(x) = -x \ln x - (1-x) \ln(1-x)$ is the binary entropy function. Thus, the leading scaling is exponential: $\binom{M}{N} \sim e^{MH(N/M)}$. It follows that

$$D_0 \sim e^{MH(N/M)}, \quad D_{\text{FCI}} \sim e^{2MH(N/M)}, \quad (1.22)$$

demonstrating that the seniority-zero space is exponentially smaller than the full Hilbert space.

For a closed-shell system, a reference determinant for the ground state is often taken as the HF determinant where all electrons are paired. Hence, in such a case, the seniority-zero sector includes HF as well as all determinants with electron-pair excitations from the same spatial orbital to the same spatial orbital. Bytautas et al. investigated CI-type expansion of the wave function based on seniority numbers [12, 13]. Two schemes of constructing the CI wave function were compared- (a) based on the maximum excitation level from a single determinant reference, and (b) based on the allowed seniority numbers of the determinants in the wave function. They demonstrated that when a single-determinant reference (HF, in this case) provides a good zeroth-order description, i.e., the correlation energy is predominantly weak/dynamic, then standard excitation-based CI procedure is more efficient than seniority-based partitioning. However, in cases that are dominated by strong/static correlation, specifically, symmetric dissociation of N_2 , CO_2 , and H_8 chain in their work, the seniority-based approach outperformed the standard CI approach. With

an increase in the highest seniority number of the determinants allowed in the wave function, the obtained energies quickly converged with the FCI limits. Notably, even wave functions restricted to the seniority-zero sector reproduced qualitatively correct potential energy curves in these cases, especially with the use of spatial symmetry-broken molecular orbitals. These findings motivated the further development of seniority-based CC methods.

1.2.2 Geminal-based methods

Geminals are correlated two-electron functions that have long been used as a building block for electronic wave functions [14, 15, 16]. For a system with N electrons and K spatial orbitals in the given one-electron basis, a general geminal creation operator can be written in second-quantized form as

$$\phi_i^\dagger = \sum_{p < q}^{2K} C_{pq}^i \hat{a}_p^\dagger \hat{a}_q^\dagger. \quad (1.23)$$

Here, the coefficients C_{pq}^i describe how pairs of one-electron basis functions contribute to geminal i . A general geminal wave function is constructed as

$$|\Psi_{\text{Geminal}}\rangle = \prod_{i=1}^P \phi_i^\dagger |0\rangle = \prod_{i=1}^P \left(\sum_{p < q}^{2K} C_{pq}^i \hat{a}_p^\dagger \hat{a}_q^\dagger \right) |0\rangle,$$

where $|0\rangle$ denotes the vacuum state and $P = \frac{N}{2}$ is the number of electron-pairs.

Because fermionic creation operators satisfy the anticommutation relations

$$\hat{a}_p^\dagger \hat{a}_q^\dagger = -\hat{a}_q^\dagger \hat{a}_p^\dagger, \quad (1.24)$$

we get

$$\phi_i^\dagger = \sum_{p,q} C_{pq}^i \hat{a}_p^\dagger \hat{a}_q^\dagger = \frac{1}{2} \sum_{p,q} (C_{pq}^i - C_{qp}^i) \hat{a}_p^\dagger \hat{a}_q^\dagger, \quad (1.25)$$

which shows that only the antisymmetric part of C_{pq}^i contributes. Hence, without loss of generality,

$$C_{pq}^i = -C_{qp}^i. \quad (1.26)$$

Because the coefficient matrix C_{pq}^i is antisymmetric, it can be brought to canonical

(pairing) form by a unitary transformation of the one-electron basis, resulting in independent 2×2 skew-symmetric blocks. In this *natural* representation [17], the geminal operator becomes

$$\phi_i^\dagger = \sum_p^K C_p^i \hat{a}_p^{(i)\dagger} \hat{a}_{\bar{p}}^{(i)\dagger}, \quad (1.27)$$

leading to

$$|\Psi_{\text{Geminal}}\rangle = \prod_{i=1}^P \left(\sum_p^K C_p^i \hat{a}_p^{(i)\dagger} \hat{a}_{\bar{p}}^{(i)\dagger} \right) |0\rangle. \quad (1.28)$$

It should be noted that the pairing index $\bar{p}$ does not, in general, denote a fixed spin counterpart (i.e., an α/β partner of the same spatial orbital), but rather a partner spin-orbital defined by the canonical transformation in the one-electron basis. A strict identification with spin-up and spin-down pairs arises only in a spin-restricted, singlet-coupled (seniority-zero) representation of the geminals. For the remaining discussion, we adopt this spin-restricted framework, such that each pair corresponds to electrons occupying the same spatial orbital with opposite spin, i.e., $(i\alpha, i\beta)$ pairing.

In Eq. (1.28), $\hat{a}_p^{(i)\dagger}$ indicates that orbital p belongs to the one-electron subspace $\mathcal{H}_i$ associated with geminal i . In the most general case, no restriction is imposed on these subspaces, and all geminals share the same one-electron basis. This leads to the Antisymmetric Product of Interacting Geminals (APIG) [18],

$$|\Psi_{\text{APIG}}\rangle = \prod_{i=1}^P \left(\sum_p^K C_p^i \hat{a}_p^\dagger \hat{a}_{\bar{p}}^\dagger \right) |0\rangle. \quad (1.29)$$

In this case, the index p runs over the same set of orbitals for every geminal i , so different electron pairs can occupy the same orbitals. This allows for full inter-geminal (inter-pair) correlation, but leads to a steep computational cost.

A simpler approximation is obtained by enforcing strong orthogonality between geminals, $\mathcal{H}_i \cap \mathcal{H}_j = \emptyset$ ($i \neq j$), which yields the Antisymmetric Product of Strongly Orthogonal Geminals (APSG) [19],

$$|\Psi_{\text{APSG}}\rangle = \prod_{i=1}^P \left(\sum_{p \in \mathcal{H}_i} C_p^i \hat{a}_p^\dagger \hat{a}_{\bar{p}}^\dagger \right) |0\rangle. \quad (1.30)$$

Here, the index p is restricted to orbitals within the subspace $\mathcal{H}_i$ of geminal i , meaning that each geminal has its own distinct set of orbitals. This reduces the computational complexity, but neglects the correlation between different electron pairs, which severely limits the quality of the geminal wave function.

To balance accuracy and efficiency, Limacher *et al.* introduced the Antisymmetric Product of 1-reference Orbital Geminals (AP1roG) [20], in which each geminal has only one pair from an occupied spatial orbital in the reference determinant and includes contributions from virtual orbitals:

$$\phi_i^\dagger = \hat{a}_i^\dagger \hat{a}_i^\dagger + \sum_a^{\text{virt}} C_i^a \hat{a}_a^\dagger \hat{a}_a^\dagger. \quad (1.31)$$

The wave function constructed from these geminals reads as

$$|\Psi_{\text{AP1roG}}\rangle = \prod_{i=1}^P \left(\hat{a}_i^\dagger \hat{a}_i^\dagger + \sum_a^{\text{virt}} C_i^a \hat{a}_a^\dagger \hat{a}_a^\dagger \right) |0\rangle. \quad (1.32)$$

The AP1roG scales as $\mathcal{O}(P^2(K - P)^2)$. Along with its favorable scaling, AP1roG also provides remarkable accuracy. For the small systems studied in Ref. [20] and in Ref. [21], including the strongly correlated case of hydrogen atom chains, AP1roG yields energies that are virtually indistinguishable from those obtained with Doubly Occupied Configuration Interaction (DOCI), which is essentially FCI restricted to seniority-zero determinants, a factorial scaling method that serves as a reference model for systems with strong electron correlation. Though achieving this level of agreement required the use of orbitals optimized at the DOCI level. Using the optimized orbitals, AP1roG also becomes size-consistent, as shown in the same work.

1.2.3 The pCCD ansatz

The similarity between the geminal wave functions and the exponential form of the CC ansatz was already known. For example, a simple derivation showing the connection between APSG and Coupled Cluster Doubles (CCD) can be found in [22].

To that end, the AP1roG wave function can also be expressed in an exponential form, culminating eventually as the pCCD ansatz.

$$\begin{aligned}
 |\Psi_{\text{AP1roG}}\rangle &= \prod_{i=1}^P \left(\hat{a}_i^\dagger \hat{a}_i^\dagger + \sum_a^{\text{virt}} C_i^a \hat{a}_a^\dagger \hat{a}_a^\dagger \right) |0\rangle \\
 &= \exp \left(\sum_i^{\text{occ}} \sum_a^{\text{virt}} t_{ii}^{aa} \hat{a}_a^\dagger \hat{a}_a^\dagger \hat{a}_i \hat{a}_i \right) |\Phi_0\rangle = \exp(\hat{T}_p) |\Phi_0\rangle = |\Psi_{\text{pCCD}}\rangle.
 \end{aligned} \tag{1.33}$$

where $\hat{T}_p$ is the cluster operator describing electron-pair excitations from the same spatial orbital. The operators $\hat{a}_p^\dagger$ and $\hat{a}_{\bar{p}}^\dagger$ ($\hat{a}_p$ and $\hat{a}_{\bar{p}}$) denote fermionic creation (annihilation) operators for α -spin orbital p and its corresponding β -spin orbital $\bar{p}$, respectively. The coefficients t_{ii}^{aa} are the pCCD cluster amplitudes associated with exciting an electron pair from the occupied orbital pair $(i, \bar{i})$ to the virtual orbital pair $(a, \bar{a})$, and $|\Phi_0\rangle$ denotes the reference wave function. These cluster amplitudes of $\hat{T}_p$ are denoted by double indices as t_{ii}^{aa} to differentiate from the standard way of writing the single excitation amplitudes. The equivalence between the energy and amplitude equations of AP1roG and pCCD can be very easily established [23].

As discussed for AP1roG, size-consistency can be achieved here when the orbital basis is optimized. Unlike Limacher et al., [20] who used orbitals optimized at DOCI-level, using orbitals optimized at the pCCD level has been suggested [24, 25]. While non-variational orbital optimization protocols have been developed for pCCD [26, 27], the works in this thesis have employed the variational approach. The variational orbital optimization approach in this case is essentially similar to orbital optimized CCD by Bozkaya et al. [28]

The Orbital-Optimized Pair Coupled Cluster Doubles (oo-pCCD) method determines the optimal molecular orbitals by enforcing stationarity of an energy Lagrangian with respect to cluster amplitudes, Lagrange multipliers, and orbital rotation parameters. The Lagrangian is defined as

$$\mathcal{L} = \langle \Phi_0 | (1 + \hat{\Lambda}_p) \hat{\mathcal{H}}_\kappa | \Phi_0 \rangle, \tag{1.34}$$

where the similarity-transformed Hamiltonian in the rotated orbital basis is $\hat{\mathcal{H}}_\kappa = e^{-\hat{T}_p} \hat{H}_\kappa e^{\hat{T}_p}$, and $\hat{H}_\kappa = e^{\hat{\kappa}} \hat{H} e^{-\hat{\kappa}}$.

The cluster excitation and de-excitation operators are restricted to electron pairs:

$$\hat{T}_p = \sum_{ia} t_{ii}^{aa} \hat{\tau}_{ii}^{a\bar{a}}, \quad \hat{\Lambda}_p = \sum_{ia} \lambda_{ii}^{aa} \hat{\tau}_{a\bar{a}}^{\bar{i}i}. \tag{1.35}$$

where $\hat{\tau}_{ii}^{a\bar{a}} = \hat{a}_a^\dagger \hat{a}_{\bar{a}}^\dagger \hat{a}_{\bar{i}} \hat{a}_i$. Orbital rotations are parameterized by a unitary transforma-

tion generated by the anti-Hermitian operator

$$\hat{\kappa} = \sum_{p>q} \kappa_{pq} (\hat{a}_p^\dagger \hat{a}_q - \hat{a}_q^\dagger \hat{a}_p). \quad (1.36)$$

In the oo-pCCD framework, orbital rotations in all blocks (occupied–occupied, virtual–virtual, and occupied–virtual) are included. Unlike conventional CCD, where rotations within the occupied or virtual subspaces are redundant due to wave function invariance, these rotations affect the pCCD wave function and must therefore be optimized. At each iteration, the Lagrangian is evaluated in the current orbital basis (corresponding to $\kappa = 0$ with respect to the current reference). Stationarity with respect to the Lagrange multipliers $\{\lambda_{ii}^{aa}\}$ yields the pCCD amplitude equations:

$$\frac{\partial \mathcal{L}}{\partial \lambda_{ii}^{aa}} = \langle \Phi_{ii}^{a\bar{a}} | \hat{\mathcal{H}}_\kappa | \Phi_0 \rangle = 0. \quad (1.37)$$

Stationarity with respect to the cluster amplitudes $\{t_{ii}^{aa}\}$ gives the equations for the Lagrange multipliers:

$$\frac{\partial \mathcal{L}}{\partial t_{ii}^{aa}} = \langle \Phi_0 | (1 + \hat{\Lambda}_p) \left[\hat{\mathcal{H}}_\kappa, \hat{\tau}_{ii}^{a\bar{a}} \right] | \Phi_0 \rangle = 0, \quad (1.38)$$

where all quantities are evaluated at the current orbital space. The orbital gradient is defined as the derivative of the Lagrangian with respect to the rotation parameters and is evaluated at $\kappa = 0$:

$$g_{pq} = \left. \frac{\partial \mathcal{L}}{\partial \kappa_{pq}} \right|_{\kappa=0} = \langle \Phi_0 | (1 + \hat{\Lambda}_p) \left[\hat{\mathcal{H}}_\kappa, \hat{E}_{pq}^- \right] | \Phi_0 \rangle, \quad (1.39)$$

where

$$\hat{E}_{pq}^- = \hat{a}_p^\dagger \hat{a}_q - \hat{a}_q^\dagger \hat{a}_p. \quad (1.40)$$

To update the orbitals, the Lagrangian is expanded to second order around $\kappa = 0$:

$$\mathcal{L}(\kappa) \approx \mathcal{L}^{(0)} + \mathbf{g}^\dagger \kappa + \frac{1}{2} \kappa^\dagger \mathbf{A} \kappa, \quad (1.41)$$

where $\mathbf{g}$ is the orbital gradient and $\mathbf{A}$ is the orbital Hessian with elements

$$A_{pq,rs} = \left. \frac{\partial^2 \mathcal{L}}{\partial \kappa_{pq} \partial \kappa_{rs}} \right|_{\kappa=0}. \quad (1.42)$$

The Newton–Raphson update for the orbital rotation parameters is then given by

$$\kappa = -\mathbf{A}^{-1}\mathbf{g}. \quad (1.43)$$

The unitary transformation matrix is approximated to second order as

$$\mathbf{U} \approx \mathbf{1} - \kappa + \frac{1}{2}\kappa^2, \quad (1.44)$$

followed by re-orthogonalization. These steps are iterated until convergence in both the energy and the orbital gradient is achieved.

As can be seen, orbital optimization in pCCD increases the computational cost from $\mathcal{O}(N^4)$ to $\mathcal{O}(N^6)$, as the evaluation of the orbital gradient and Hessian involves contractions over all orbital pairs, thereby removing the pair restriction that makes pCCD computationally efficient. By introducing suitable intermediates, such as factorizations based on density fitting [29] or Cholesky decomposition [30], the most expensive contractions can be reorganized, reducing the formal scaling from $\mathcal{O}(N^6)$ to $\mathcal{O}(N^5)$.

For many molecular systems, oo-pCCD has been shown to closely reproduce DOCI results [23, 31]. For a strongly correlated system like H₅₀-chain, oo-pCCD was found to outperform Second-Order Møller–Plesset Perturbation Theory (MP2) and DFT, providing qualitatively correct Potential Energy Surface (PES) and reliable spectroscopic constants using Density Matrix Renormalization Group (DMRG) as reference [24]. Further work with oo-pCCD by Tecmer et al. [32] revealed the ability of oo-pCCD to reliably describe the electronic structure of closed-shell actinide compounds.

1.3 Dynamic correlation corrections on pCCD

As has been discussed so far, the pCCD wave function reliably depicts static electron correlation effects. However, a large part of the dynamic correlation effect, which is attributed to excitations beyond electron pairs, is excluded in the model. The inclusion of dynamic correlation effects is indispensable for attaining quantitative accuracy in energies and reliable spectroscopic properties. Various Post-pCCD *a posteriori* corrections have been strategized to circumvent this problem. DFT [33, 34] and CI-based [35] strategies have been tested for the purpose of adding dynamic correlation to pCCD. Perturbation theory-based correction approaches were implemented in this context [36]. The performance of the perturbative corrections methods has been analyzed for homonuclear and heteronuclear diatomics in Ref. [37]. However, most of the perturbative models exhibited numerical divergences in potential energy surfaces of diatomics, particularly near the dissociation region, and produced unreliable dissociation energies. These discrepancies lead us to the need for CC theory-based dynamic correlation corrections. The works comprised in this thesis have used CC-based dynamic correlation corrections, which will be discussed in detail in the next part.

1.3.1 Coupled cluster-based dynamic correlation corrections on pCCD

The basic idea of adding correlation correction in the pCCD model revolves around adding more excitations to the reference wave function. In the traditional CC studies employing the usual truncated cluster operator (CCSD generally), this inclusion of higher excitations has been done by obtaining a subset of cluster amplitudes associated with higher excitations using independent external models. Paldus and co-workers made extensive development of this approach, termed the externally corrected (ec)-CC [38, 39, 40, 41, 42]. In the same context, Stolarczyk proposed to obtain the triple and quadruple excitation amplitudes from CASSCF wave function and add them to the CCSD model using the Coupled Cluster with Single, Double, Triple, and Quadruple (CCSDTQ) framework [43]. This strategy was implemented by Peris et al [44]. Paldus and Li also used Multi-Reference Configuration Interaction with Single and Double (MRCI-SD) to extract the triple and quadruple excitation amplitudes. The method, given the moniker reduced multireference CCSD (RMR CCSD), performed better than the CASSCF-based ec-CC for bond-breaking cases [45, 46, 47, 48]. Recent work by Aroeira et al. explores CC externally corrected by triples and quadruples amplitudes obtained from adaptive CI [49]. Jankowski et al. pursued a ‘split-amplitude’ approach, expressing each type of cluster amplitude

as a sum of amplitudes obtained from two sources [50, 51]. Methods like MP2, Configuration Interaction Doubles (CID), and MRCI-SD were used alongside linearized CCSD, which the authors termed *almost linear* CC (AL-CC) in their case, as the linearized amplitude equations depended on amplitudes obtained from the external methods.

A similar approach is the Tailored Coupled Cluster (tCC) approach. In tCC, the excitation space is partitioned [52, 53, 54] by factorizing the cluster operator as $\hat{T} = \hat{T}_{\text{ext}} + \hat{T}_{\text{int}}$. Utilizing the exponential form of the CC ansatz and the Baker–Campbell–Hausdorff (BCH) expansion, the tCC wave function can then be expressed as (provided $\hat{T}_{\text{int}}$ and $\hat{T}_{\text{ext}}$ commute)

$$|\Psi_{\text{tCC}}\rangle = e^{\hat{T}_{\text{ext}}} e^{\hat{T}_{\text{int}}} |\Phi_0\rangle. \quad (1.45)$$

First, $\hat{T}_{\text{int}}$, describing the strong correlation effect, is applied to the reference Slater determinant $|\Phi_0\rangle$. Subsequently, $\hat{T}_{\text{ext}}$ is applied keeping the amplitudes of $\hat{T}_{\text{int}}$ constant to add additional excitations and incorporate the effect of dynamic correlation in the final wave function. For example, Kinoshita et al. implemented a Complete Active Space Configuration Interaction (CASSCI)-based tCC [55, 56]. Melnichuk et al. tried to mitigate the errors in tCC due to the small active space of the CI step by expanding the active space for a Coupled Cluster Singles, Doubles, and Triples (CCSDT) calculation and keeping the singles and doubles amplitudes obtained at this step frozen while running CCSD for the full orbital space. They called this method ‘extended tCCSD’ [57, 58]. Feldmann et al. implemented an iterative scheme of improving the partitioning of the excitation space, by using similarity transformation of the Hamiltonian of the CASSCI calculation with the external amplitudes coming from the CCSD calculations [59]. CC tailored by matrix product states as in DMRG (DMRG-tCC) has also been implemented in recent years, which exploits the capacity of DMRG to deal with larger active spaces compared to CASSCI or CASSCF [60, 61, 62].

A similar tCC approach can be conceptualized for pCCD, where the cluster pair-excitation operator $\hat{T}_{\text{p}}$ takes care of the strong correlation, i.e, it acts as $\hat{T}_{\text{int}}$. Such an approach is termed as Frozen Pair Coupled Cluster (fpCC) [23], where the fpCC ansatz is expressed as

$$|\Psi_{\text{fpCC}}\rangle = e^{\hat{T}_{\text{ext}}} e^{\hat{T}_{\text{p}}} |\Phi_0\rangle. \quad (1.46)$$

In Eq. (1.46), $\hat{T}_{\text{ext}}$ is the cluster operator which acts upon the pCCD wave function and generates single and unpaired double excitations. The amplitudes of $\hat{T}_{\text{p}}$, which

acts as $\hat{T}_{\text{int}}$ in this case, are kept constant (not reoptimized in presence of the $\hat{T}_{\text{ext}}$ amplitudes), and thus the method is denoted as ‘frozen pair’. The amplitudes of $\hat{T}_{\text{ext}}$ can be obtained by solving the following equation

$$\langle \Phi_\nu | e^{-\hat{T}_p} e^{-\hat{T}_{\text{ext}}} \hat{H}_N e^{\hat{T}_{\text{ext}}} e^{\hat{T}_p} | \Phi_0 \rangle = 0. \quad (1.47)$$

Here, $\hat{H}_N = \hat{H} - \langle \Phi_0 | \hat{H} | \Phi_0 \rangle$ is the normal ordered form of the electronic Hamiltonian $\hat{H}$ and $\langle \Phi_\nu |$ is a Slater determinant from the projection space $\{\nu\}$ created by $\hat{T}_{\text{ext}}$ acting upon the pCCD wave function. The cluster operator $\hat{T}_{\text{ext}}$ is put under the additional constraint that it generates states orthogonal to the pCCD wave function, i.e.,

$$\langle \Psi_{\text{pCCD}} | \hat{T}_{\text{ext}} | \Psi_{\text{pCCD}} \rangle = 0. \quad (1.48)$$

From this, we can express the similarity-transformed Hamiltonian of fpCC as

$$\begin{aligned} e^{-\hat{T}_p} e^{-\hat{T}_{\text{ext}}} \hat{H}_N e^{\hat{T}_{\text{ext}}} e^{\hat{T}_p} &= \hat{H}_N + [\hat{H}_N, \hat{T}_{\text{ext}}] + \frac{1}{2} [[\hat{H}_N, \hat{T}_{\text{ext}}], \hat{T}_{\text{ext}}] + \frac{1}{6} [[[[\hat{H}_N, \hat{T}_{\text{ext}}], \hat{T}_{\text{ext}}], \hat{T}_{\text{ext}}], \hat{T}_{\text{ext}}] \\ &\quad + \frac{1}{24} [[[[[\hat{H}_N, \hat{T}_{\text{ext}}], \hat{T}_{\text{ext}}], \hat{T}_{\text{ext}}], \hat{T}_{\text{ext}}], \hat{T}_{\text{ext}}] \\ &\quad + [\hat{H}_N, \hat{T}_p] + [[\hat{H}_N, \hat{T}_{\text{ext}}], \hat{T}_p] + \frac{1}{2} [[[[\hat{H}_N, \hat{T}_{\text{ext}}], \hat{T}_{\text{ext}}], \hat{T}_p] \\ &\quad + \frac{1}{2} [[\hat{H}_N, \hat{T}_p], \hat{T}_p]. \end{aligned} \quad (1.49)$$

Following the spin-free (or spin-adapted) CC formulation, where all cluster operators are written in terms of singlet excitation operators $\hat{E}_{ai}$, the cluster operators for single excitation and unpaired double excitations are defined as

$$\hat{T}_1 = \sum_i^{n_{\text{occ}}} \sum_a^{n_{\text{virt}}} t_i^a \hat{E}_{ai}, \quad (1.50)$$

and

$$\hat{T}'_2 = (\hat{T}_2 - \hat{T}_p) = \frac{1}{2} \sum_{i,j}^{n_{\text{occ}}} \sum_{a,b}^{n_{\text{virt}}} t_{ij}^{ab} \hat{E}_{ai} \hat{E}_{bj} - \frac{1}{2} \sum_i^{n_{\text{occ}}} \sum_a^{n_{\text{virt}}} t_{ii}^{aa} \hat{E}_{ai} \hat{E}_{ai}, \quad (1.51)$$

respectively, where

$$\hat{E}_{ai} = \hat{a}_a^\dagger \hat{a}_i + \hat{a}_a^\dagger \hat{a}_{\bar{i}}. \quad (1.52)$$

Depending upon the choice of $\hat{T}_{\text{ext}}$, we can have two versions of fpCC: Frozen Pair Coupled Cluster Doubles (fpCCD), where $\hat{T}_{\text{ext}}$ is restricted to only unpaired doubles excitations ($\hat{T}_{\text{ext}} = \hat{T}'_2$) and Frozen Pair Coupled Cluster Singles and Doubles (fpCCSD) where we also include singles excitations ($\hat{T}_{\text{ext}} = \hat{T}_1 + \hat{T}'_2$). Hence, fpCC differs from the ec-CC methods also in the sense that we do not consider triple or quadruple excitations, but rather improve the wave function description by adding higher-seniority determinants using single and unpaired double excitations only. The fpCC approach should be considered as an easily implementable alternative to combine the reliable description of strong correlation by pCCD with the capability of traditional coupled cluster to also describe dynamic correlation. Leszczyk et al. [63] tested fpCC for a set of main group homonuclear and heteronuclear diatomics. Their numerical results show that fpCCD and fpCCSD provided reliable descriptions of potential energy surfaces and spectroscopic constants that closely match multireference references, showing almost the same or slightly lower average errors than CCD and CCSD. fpCCSD was also found to be devoid of the unphysical cusp in the potential energy curve of ethylene twist commonly encountered by CCD and CCSD. However, it should also be noted that the performance of fpCC is strongly limited by how good the pCCD itself is as the strong correlation part. In specific cases, where pCCD provides only a limited description of strong correlation, fpCC also breaks down, as shown in the case of the triple bond-breaking of N_2 in [23] or the hextuple bond-breaking of Cr_2 in [63].

Another simplified way of dynamic correlation correction on pCCD is the Frozen Pair Linearized Coupled Cluster (fpLCC) ansatz [64]. Linearized Coupled Cluster (LCC) corrections have been previously used as a dynamic correlation correction to the APSG wave function [65]. Here, the cluster operator is linearized, and then this acts upon the $|\Psi_{\text{pCCD}}\rangle$ reference wave function.

$$|\Psi_{\text{fpLCC}}\rangle \approx (1 + \hat{T}_{\text{ext}}) |\Psi_{\text{pCCD}}\rangle. \quad (1.53)$$

The associated BCH expansion is terminated at the second term, ignoring all higher-order non-linear terms in $\hat{T}_{\text{ext}}$,

$$e^{-\hat{T}_{\text{ext}}} \hat{H}_N e^{\hat{T}_{\text{ext}}} = \hat{H}_N + [\hat{H}_N, \hat{T}_{\text{ext}}] + \frac{1}{2} [[\hat{H}_N, \hat{T}_{\text{ext}}], \hat{T}_{\text{ext}}] + \dots \approx \hat{H}_N + [\hat{H}_N, \hat{T}_{\text{ext}}]. \quad (1.54)$$

Consequently, the fpLCC energy equation simplifies as

$$(\hat{H}_N + [\hat{H}_N, \hat{T}_{\text{ext}}]) |\Psi_{\text{pCCD}}\rangle = E |\Psi_{\text{pCCD}}\rangle. \quad (1.55)$$

The energy can then be obtained by multiplying Eq. (1.56) from the left by $\langle\Phi_0|$ and using intermediate normalization.

$$\langle\Phi_0| (\hat{H}_N + [\hat{H}_N, \hat{T}_{\text{ext}}]) |\Psi_{\text{pCCD}}\rangle = E. \quad (1.56)$$

The cluster amplitudes t_ν can be determined by multiplying Eq. (1.54) from the left by $\langle\Phi_\nu|$, which yields

$$\langle\Phi_\nu| (\hat{H}_N + [\hat{H}_N, \hat{T}_{\text{ext}}]) |\Psi_{\text{pCCD}}\rangle = 0. \quad (1.57)$$

Similar to fpCC, here also the projection space $\{\nu\}$ depends on the choice of $\hat{T}_{\text{ext}}$, which gives rise to two different versions of fpLCC. We can include only unpaired doubles excitations (Frozen Pair Linearized Coupled Cluster Doubles (fpLCCD)) or both unpaired doubles as well as singles excitations (Frozen Pair Linearized Coupled Cluster Singles and Doubles (fpLCCSD)). A point to note is that fpLCC was also denoted previously as AP1roG-LCC and pCCD-LCC in the literature.

Using $|\Psi_{\text{pCCD}}\rangle$ as reference brings additional terms in the fpLCC amplitude equations compared to the standard single-reference LCC approach which uses the HF determinant as reference [66, 67, 68]. This is also evident if we expand Eq. (1.55) as

$$(\hat{H}_N + [\hat{H}_N, \hat{T}_{\text{ext}}]) e^{\hat{T}_p} |\Phi_0\rangle = E e^{\hat{T}_p} |\Phi_0\rangle, \quad (1.58)$$

and left multiply with $e^{-\hat{T}_p}$

$$e^{-\hat{T}_p} (\hat{H}_N + [\hat{H}_N, \hat{T}_{\text{ext}}]) e^{\hat{T}_p} |\Phi_0\rangle = E |\Phi_0\rangle. \quad (1.59)$$

This can then be used to rewrite Eq. (1.57) as

$$\langle\Phi_\nu| \left(\hat{H}_N + [\hat{H}_N, \hat{T}_{\text{ext}}] + [\hat{H}_N, \hat{T}_p] + [[\hat{H}_N, \hat{T}_{\text{ext}}], \hat{T}_p] + \frac{1}{2} [[\hat{H}_N, \hat{T}_p], \hat{T}_p] \right) |\Phi_0\rangle = 0. \quad (1.60)$$

This shows the coupling term between the pair doubles, and singles and/or unpaired

doubles excitations. However, fpLCC has fewer terms than fpCC in the amplitude equations because we exclude the terms non-linear in $\hat{T}_{\text{ext}}$, which are retained in fpCC as shown in Eq. (1.49). Numerical results for several molecules such as C_2 , F_2 , the first excited state of BN, UO_2^{2+} (symmetric dissociation) obtained by Boguslawski et al. in [64] show the capability of fpLCCSD to produce reliable spectroscopic constants and PES when compared with CCSD(T) and MRCI-SD reference results. However, numerical divergences were observed for fpLCC in some cases, such as for CO [63]. This drawback can be remedied if non-linear terms are added in the amplitude equations, ultimately resulting in the fpCC framework.

Both fpCC and fpLCC corrections have the same $\mathcal{O}(o^2v^4)$ scaling as standard CCD/CCSD (where o and v denote the number of occupied and virtual molecular orbitals, respectively), although their computational pre-factors differ. Table 1.1 summarizes these details for all the CC-based corrections discussed in this section.

Table 1.1: Details of CC models for dynamic correlation correction.

Method	CC operator	Reference $ \Psi_0\rangle$	Scaling	Memory
pCCD	$e^{\hat{T}_p}$	$ \Phi_0\rangle$	$\mathcal{O}(o^2v^2)$	$\mathcal{O}(N^3)$
fpCCD	$e^{\hat{T}'_2}$	$ \Psi_{\text{pCCD}}\rangle$	$\mathcal{O}(o^2v^4)$	$\mathcal{O}(v^4)$
fpCCSD	$e^{\hat{T}_1 + \hat{T}'_2}$	$ \Psi_{\text{pCCD}}\rangle$	$\mathcal{O}(o^2v^4)$	$\mathcal{O}(v^4)$
fpLCCD	$1 + \hat{T}'_2$	$ \Psi_{\text{pCCD}}\rangle$	$\mathcal{O}(o^2v^4)$	$\mathcal{O}(v^4)$
fpLCCSD	$1 + \hat{T}_1 + \hat{T}'_2$	$ \Psi_{\text{pCCD}}\rangle$	$\mathcal{O}(o^2v^4)$	$\mathcal{O}(v^4)$

Reference wave functions (Ψ_0), the employed CC operators, and formal computational scalings of all the different flavors of CC-based dynamic correlation corrections discussed in this section. For

a schematic of the full computational costs, pCCD is also included here. o and v denote the number of occupied and virtual molecular orbitals, respectively, and $N = o + v$. $\hat{T}'_2$ is the cluster operator for unpaired double excitations, i.e., $\hat{T}'_2 = \hat{T}_2 - \hat{T}_p$. It is to be noted that $|\Psi_0\rangle$ for pCCD, i.e., $|\Phi_0\rangle$ can be the HF determinant or a single Slater determinant reference function for pCCD constructed from the optimized pCCD natural orbitals, i.e., in the oo-pCCD approach. Memory indicates the order of the storage bottleneck, excluding the solver-dependent prefactor.

1.4 Excited states with pCCD

Modeling the electronic structures of excited states is generally more demanding compared to the ground state [69, 70]. Handling different excited states of a chemical system, each characterized by varying degrees of static and dynamic electron correlation, introduces additional challenges. Several promising quantum chemistry approaches suitable for excited states have emerged in the last two decades or so. In the CC domain, Fock-space Coupled Cluster (FSCC) [71, 72] and its formulation based on the intermediate Hamiltonian (designed to overcome the intruder state problem) [73, 74, 75] methods have seen major developments for excited state studies. Parallely, the Equation-of-Motion (EOM) formalism has gained popularity for the same purpose [76, 77, 78]. Recent developments have taken place where pCCD and post-pCCD dynamic correlation corrected methods have been used to model the ground state for the EOM framework. Kossoski et al. recently implemented a state-specific orbital optimization protocol for pCCD and estimated the excitation energies from a Δ oo-pCCD approach [79]. The works included in this thesis have extensively used these EOM-based approaches. Hence, in this section, first, the basic formalism of EOM will be briefly discussed below. Then, the theoretical concept of EOM-pCCD will be presented. Thereafter, the methods to access singly excited states will be discussed.

In the standard EOM approach, the excited state wave function $|\Psi_k\rangle$ is expressed using a CI-type excitation operator,

$$\hat{R} = \sum_{\mu} r_{\mu} \hat{\tau}_{\mu}, \quad (1.61)$$

which acts on the CC reference ground state wave function $|\Psi_{\text{gs}}\rangle = e^{\hat{T}} |\Phi_0\rangle$ to generate the excited state,

$$|\Psi_k\rangle = \hat{R} |\Psi_{\text{gs}}\rangle = \sum_{\mu} r_{\mu} \hat{\tau}_{\mu} e^{\hat{T}} |\Phi_0\rangle. \quad (1.62)$$

Here, the sum index μ runs over all excitations present in the cluster operator, including the identity operator $\hat{\tau}_0$ (with corresponding coefficient $r_0^{(k)}$), which describes the weightage of the ground-state in the targeted excited state. In general, $\hat{\tau}_{\mu}$ denotes an excitation operator that promotes electrons from occupied to virtual orbitals, generating the excited determinants spanning the EOM-CC configuration space. They can be defined as

$$\hat{\tau}_1 = \sum_{i,a} \hat{a}_a^\dagger \hat{a}_i, \quad \hat{\tau}_2 = \frac{1}{4} \sum_{ij,ab} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i, \quad (1.63)$$

and so on. The excited state energies can be obtained from the corresponding equation,

$$\hat{H}_N |\Psi_k\rangle = \hat{H}_N \hat{R} e^{\hat{T}} |\Phi_0\rangle = E_k \hat{R} e^{\hat{T}} |\Phi_0\rangle = E_k |\Psi_k\rangle, \quad (1.64)$$

where $\hat{H}_N = \hat{H} - \langle \Phi_0 | \hat{H} | \Phi_0 \rangle$ is the normal ordered form of the Hamiltonian and E_k is the total electronic energy of the k -th excited state. Multiplying Eq. (1.64) with $e^{-\hat{T}}$ from left and using the fact that $\hat{R}$ and $\hat{T}$ commute, we get

$$e^{-\hat{T}} \hat{H}_N e^{\hat{T}} \hat{R} |\Phi_0\rangle = E_k \hat{R} |\Phi_0\rangle. \quad (1.65)$$

Subtracting the corresponding equation of the CC ground state and using the similarity-transformed Hamiltonian $\hat{\mathcal{H}}_N = e^{-\hat{T}} \hat{H}_N e^{\hat{T}}$, we arrive at the EOM equation,

$$[\hat{\mathcal{H}}_N, \hat{R}] |\Phi_0\rangle = \omega_k \hat{R} |\Phi_0\rangle, \quad (1.66)$$

where $\omega_k = (E_k - E_0)$ are the excitation energies with respect to the CC ground state $|\Psi_{\text{gs}}\rangle$. These excitation energies are calculated from

$$\begin{bmatrix} 0 & \langle \Phi_0 | \hat{\mathcal{H}}_N | \Phi_\mu \rangle \\ 0 & \langle \Phi_\nu | [\hat{\mathcal{H}}_N, \hat{\tau}_\mu] | \Phi_0 \rangle \end{bmatrix} \begin{bmatrix} r_0^{(k)} \\ r_\mu^{(k)} \end{bmatrix} = \omega_k \begin{bmatrix} r_0^{(k)} \\ r_\mu^{(k)} \end{bmatrix}. \quad (1.67)$$

Here, $|\Phi_\mu\rangle = \hat{\tau}_\mu |\Phi_0\rangle$ and $|\Phi_\nu\rangle = \hat{\tau}_\nu |\Phi_0\rangle$ denote excited Slater determinants generated from the reference determinant $|\Phi_0\rangle$. The coefficients $r_0^{(k)}$ and $r_\mu^{(k)}$ are the components of the right EOM eigenvector for the k -th excited state, where $r_0^{(k)}$ corresponds to the reference component and $r_\mu^{(k)}$ to the excited configurations. The excitation energy $\omega_k = E_k - E_0$ is obtained as an eigenvalue of the non-Hermitian matrix. If only excitation energies are of interest, solving the right-eigenvalue problem in Eq. (1.67) is sufficient. However, transition properties will require both left and right eigenvectors, which are different in this case, as $\hat{\mathcal{H}}_N$ is non-Hermitian.

It should be noted that the Equation-of-Motion Coupled Cluster (EOM-CC) formalism is formally exact, since the similarity transformation here does not alter the eigenvalues of the full Hamiltonian. In practical applications, however, both the

cluster operator $\hat{T}$ and the excitation operator $\hat{R}$ must be truncated to a manageable level of excitation, thereby introducing an approximation to the theory [78].

1.4.1 Modeling electron-pair excited states with EOM-pCCD

The simplest way of modeling the excited states with pCCD as the ground state counterpart is the EOM-pCCD model. In that case, the cluster operator $\hat{T}$ as well as the excitation operator of EOM $\hat{R}$ is restricted to electron-pair excitations only. In this case, Eq. (1.61) modifies to

$$\hat{R}_p = \hat{R}_0 + \sum_{ia} r_{ii}^{aa} \hat{\tau}_{ii}^{a\bar{a}}, \quad (1.68)$$

where $\hat{R}_0 = r_0 \hat{\tau}_0$, and $\hat{\tau}_{ii}^{a\bar{a}} = \hat{a}_a^\dagger \hat{a}_{\bar{a}}^\dagger \hat{a}_{\bar{i}} \hat{a}_i$ denotes pair excitations. In the spin-free CC formalism, this pair excitation operator can also be represented as $\hat{\tau}_{ii}^{a\bar{a}} = \frac{1}{2} \hat{E}_{ai} \hat{E}_{ai}$, where $\hat{E}_{ai}$ is given by Eq. (1.52). Consequently, Eq. (1.68) can also be written as

$$\hat{R}_p = \hat{R}_0 + \frac{1}{2} \sum_{ia} r_{ii}^{aa} \hat{E}_{ai} \hat{E}_{ai}. \quad (1.69)$$

The corresponding EOM-pCCD equation modifies to,

$$\left[\hat{\mathcal{H}}_N^{(p)}, \hat{R}_p \right] |\Phi_0\rangle = \omega_p \hat{R}_p |\Phi_0\rangle. \quad (1.70)$$

Here, $\hat{\mathcal{H}}_N^{(p)}$ is the similarity-transformed Hamiltonian for pCCD, which can be expressed as

$$\hat{\mathcal{H}}_N^{(p)} = e^{-\hat{T}_p} \hat{H}_N e^{\hat{T}_p} = \hat{H}_N + [\hat{H}_N, \hat{T}_p] + \frac{1}{2} [[\hat{H}_N, \hat{T}_p], \hat{T}_p]. \quad (1.71)$$

The main limitation of EOM-pCCD is that it can only target pair-excited states, which limits its applicability in general cases. Description of these pair-excited states in EOM-pCCD also suffers from the lack of relaxation effects, due to the exclusion of single excitations in the model. For example, EOM-pCCD overestimates the Vertical Excitation Energy (VEE) for the $2^1A_g^-$ state of different all-trans polyenes from C_4H_6 to $C_{14}H_{16}$ relative to the corresponding multi-reference Møller–Plesset (MRMP) excitation energies by as high as ≈ 4 eV and by ≈ 1.5 eV relative to Configuration Interaction with perturbative Doubles (CI(D)) [80, 81].

1.4.2 EOM-pCCD extensions for singly excited states

The simplest modification of EOM-pCCD to include single excitations is introduced in a way similar to the Configuration Interaction Singles (CIS) method [82]. In this approach, only the $\hat{R}$ operator of Eq. (1.68) is modified to include single excitations with respect to the pCCD reference determinant,

$$\hat{R}_{ps} = \hat{R}_0 + \sum_{ia} r_i^a \hat{\tau}_i^a + \sum_{ia} r_{ii}^{aa} \hat{\tau}_{ii}^{a\bar{a}}, \quad (1.72)$$

where $\hat{\tau}_i^a = \hat{a}_a^\dagger \hat{a}_i$ is single excitation operator (which corresponds to $\hat{E}_{ai}$ in the spin-free formalism). The corresponding excitation energies are obtained by solving the EOM equation

$$\left[\hat{\mathcal{H}}_N^{(p)}, \hat{R}_{ps} \right] |\Phi_0\rangle = \omega_{ps} \hat{R}_{ps} |\Phi_0\rangle. \quad (1.73)$$

Here, $\hat{\mathcal{H}}_N^{(p)}$ is the similarity-transformed Hamiltonian of pCCD as defined in Eq. (1.71). This model is denoted as EOM-pCCD+S. We should emphasize that EOM-pCCD+S uses the pCCD wave function as the reference state while CIS uses a single Slater determinant as reference. The electron correlation effects at pCCD level are included in EOM-pCCD+S through the $\hat{T}_p$ operator in the similarity transformed Hamiltonian.

Results from [80, 81] show that EOM-pCCD+S overestimates vertical excitation energies of the two lowest-lying excited states ($\sigma_u \rightarrow \delta_u$ and $\sigma_u \rightarrow \phi_u$) transitions in UO_2^{2+} by ≈ 0.5 eV, while the respective adiabatic excitation energies differ by ≈ 0.7 eV. CIS performs better in this case than EOM-pCCD+S and deviates from the associated Equation-of-Motion Coupled Cluster Singles and Doubles (EOM-CCSD) values by approximately 0.3 eV (vertical excitations) to 0.6 eV (adiabatic excitations). However, EOM-pCCD+S provides better equilibrium U–O bond lengths in excited states, differing from EOM-CCSD values by ≈ 0.06 Å compared to ≈ 0.08 Å in CIS. For all-trans polyenes from C_4H_6 to $\text{C}_{14}\text{H}_{16}$, the vertical excitation energies of the $2^1A_g^-$ states remain ≈ 1 -2 eV away from the corresponding MRMP values, while the same for the $1^1B_u^+$ state show much smaller differences from MRMP.

It should be noted here that the EOM-pCCD+S model does not remain size-consistent as the single excitations are introduced *a posteriori* only in the $\hat{R}$ operator and not in the cluster operator.

$$\begin{bmatrix} 0 & \langle \Phi_0 | \hat{\mathcal{H}}_N^{(p)} | \Phi_j^b \rangle & \langle \Phi_0 | \hat{\mathcal{H}}_N^{(p)} | \Phi_{ii}^{a\bar{a}} \rangle \\ \langle \Phi_i^a | \hat{\mathcal{H}}_N^{(p)} | \Phi_0 \rangle & \langle \Phi_i^a | \hat{\mathcal{H}}_N^{(p)} | \Phi_j^b \rangle & \langle \Phi_i^a | \hat{\mathcal{H}}_N^{(p)} | \Phi_{ii}^{a\bar{a}} \rangle \\ 0 & \langle \Phi_{ii}^{a\bar{a}} | \hat{\mathcal{H}}_N^{(p)} | \Phi_j^b \rangle & \langle \Phi_{ii}^{a\bar{a}} | \hat{\mathcal{H}}_N^{(p)} | \Phi_{ii}^{a\bar{a}} \rangle \end{bmatrix}$$

Terms such as $\langle \Phi_0 | \hat{\mathcal{H}}_N^{(p)} | \Phi_j^b \rangle$ are not equal to 0 in EOM-pCCD+S, as $\hat{T}_1$ is not present in the cluster operator. The size-intensivity can be restored by including single excitations in the reference wave function. If the EOM model is restricted to single and electron-pair excitations, we have to perform an *a posteriori* Coupled Cluster Singles (CCS) calculation by applying $\hat{T}_1$ on top of pCCD (denoted as pCCD-CCS). The corresponding EOM equation then reads as

$$\left[\hat{\mathcal{H}}_N^{(ps)}, \hat{R}_{ps} \right] |\Phi\rangle = \omega_{ps} \hat{R}_{ps} |\Phi\rangle, \quad (1.74)$$

where $\hat{\mathcal{H}}_N^{(ps)}$ is the similarity-transformed Hamiltonian of pCCD-CCS. Since single excitations are now included in the reference wavefunction, the first column of the effective Hamiltonian vanishes, bringing back size-consistency. It is important to note that the similarity-transformed Hamiltonian $\hat{\mathcal{H}}_N^{(ps)}$ contains contributions from both the single and electron-pair cluster operators in this case.

$$\begin{aligned} \hat{\mathcal{H}}_N^{(ps)} &= e^{-(\hat{T}_1 + \hat{T}_p)} \hat{H}_N e^{(\hat{T}_p + \hat{T}_1)} \\ &= \hat{H}_N + [\hat{H}_N, \hat{T}_p] + \frac{1}{2} [[\hat{H}_N, \hat{T}_p], \hat{T}_p] \\ &\quad + [\hat{H}_N, \hat{T}_1] + [[\hat{H}_N, \hat{T}_p], \hat{T}_1] \\ &\quad + \frac{1}{2} [[\hat{H}_N, \hat{T}_1], \hat{T}_1] + \frac{1}{2} [[[\hat{H}_N, \hat{T}_p], \hat{T}_1], \hat{T}_1] \\ &\quad + \frac{1}{6} [[[\hat{H}_N, \hat{T}_1], \hat{T}_1], \hat{T}_1] \\ &\quad + \frac{1}{24} [[[[\hat{H}_N, \hat{T}_1], \hat{T}_1], \hat{T}_1], \hat{T}_1] \end{aligned} \quad (1.75)$$

For matrix elements involving electron-pair excitations, the dominant contributions arise from the pair cluster operator $\hat{T}_p$, leading to a form similar to $\hat{\mathcal{H}}_N^{(p)}$ defined in Eq. (1.71). In contrast, matrix elements involving single excitations receive contributions from both $\hat{T}_1$ and $\hat{T}_p$, as well as their mixed commutators.

1.4.3 EOM-pCCD for general doubly excited states

By construction, non-pair doubly excited states are not accessible in the methods discussed so far. To target general doubly excited states, post-pCCD methods must

be considered that include the $\hat{T}'_2$ cluster operator in the coupled-cluster correction. In this context, the fpLCCSD method has been combined with the EOM framework to model excited states [83]. Recently, Bartlett and coworkers have also implemented an Equation-of-Motion pair Coupled Cluster Doubles tailored Coupled Cluster Singles and Doubles (EOM-pCCD-tCCSD) approach, which uses pCCD-tailored CCSD, which is essentially fpCCSD, as the ground state [84]. In the next part, the Equation-of-Motion Frozen Pair Linearized Coupled Cluster Singles and Doubles (EOM-fpLCCSD) method will be discussed, which has been used in the works comprised in this thesis.

Since the cluster operator of fpLCCSD contains both single and double (i.e., both electron-pair and non-pair) excitations, the $\hat{R}$ operator has the general form

$$\hat{R}_{\text{sd}} = \hat{R}_0 + \sum_{ia} r_i^a \hat{\tau}_i^a + \frac{1}{4} \sum_{ij} \sum_{ab} r_{ij}^{ab} \hat{\tau}_{ij}^{ab}. \quad (1.76)$$

Here $\hat{\tau}_{ij}^{ab} = \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i$ denotes general double excitations. In the spin-free formulation, the corresponding double-excitation part of the EOM operator can be written as

$$\hat{R}_2 = \frac{1}{2} \sum_{i,j} \sum_{a,b} r_{ij}^{ab} \hat{E}_{ai} \hat{E}_{bj}.$$

Substituting this expression for the $\hat{R}$ operator into the general EOM Eq. (1.66), the EOM-fpLCCSD equation becomes

$$\left[\hat{\mathcal{H}}_N^{(\text{sd})}, \hat{R}_{\text{sd}} \right] |\Phi_0\rangle = \omega_{\text{sd}} \hat{R}_{\text{sd}} |\Phi_0\rangle. \quad (1.77)$$

Similar to EOM-pCCD-CCS, different expressions for the similarity-transformed Hamiltonian associated with electron-pair and non-pair excitations are used. Specifically, for non-pair excitations, we have

$$\begin{aligned} \hat{\mathcal{H}}_N^{(np)} &= e^{-(\hat{T}_{np} + \hat{T}_p)} \hat{H}_N e^{(\hat{T}_p + \hat{T}_{np})} \\ &\approx \hat{H}_N + [\hat{H}_N, \hat{T}_{np}] + [\hat{H}_N, \hat{T}_p] \\ &\quad + [[\hat{H}_N, \hat{T}_{np}], \hat{T}_p] + \frac{1}{2} [[\hat{H}_N, \hat{T}_p], \hat{T}_p] \end{aligned} \quad (1.78)$$

Table 1.2 enumerates the different excitation operators and computational requirements of the EOM-pCCD methods discussed here. The performance of the three models with single excitations presented in this section was analyzed with respect to

the electronic structures of excited states of some f0 actinide species such as ThO, ThS, UO_2^{2+} , NUO^+ , CUO , NUN , PaO_2^+ , and ThO_2 molecules in Ref. [85]. EOM-fpLCCSD showed the lowest statistical errors compared to completely renormalized Equation-of-Motion Coupled Cluster Singles and Doubles with perturbative Triples (CR-EOM-CCSD(T)) reference data for excitation energies (vertical and adiabatic), giving very similar results as EOM-CCSD. The other two simpler single extension models perform rather poorly in these cases. For example, for the adiabatic excitation energies of UO_2^{2+} for states $2^1\Phi_g$, $1^1\Pi_g$, $1^1\Gamma_g$, EOM-pCCD+S and EOM-pCCD-CCS show difference of ≈ 3.3 eV from CR-EOM-CCSD(T). These numbers are drastically improved upon using EOM-fpLCCSD, even outperforming EOM-CCSD. In general, EOM-pCCD+S and EOM-pCCD-CCS overestimate adiabatic excitation energies by up to 3.5 eV for UO_2^{2+} , up to 2.5 eV for NUN , up to 1.2 eV for PaO_2^+ , and up to 2.3 eV for ThO_2 . However, spectroscopic constants (r_e , the equilibrium bond length, and ω_e , the harmonic vibrational frequency) obtained with EOM-fpLCCSD show higher error than EOM-CCSD. In general, EOM-fpLCCSD overestimates r_e typically by ≈ 0.03 Å or less and underestimates ω_e . Specifically, for ThO and ThS, the spectroscopic constants obtained with EOM-fpLCCSD are even worse than both EOM-pCCD+S and EOM-pCCD-CCS. To summarize, while EOM-fpLCCSD provides excitation energies in excellent agreement with reference values, the spectroscopic constants obtained with the same are not always reliable.

Table 1.2: Details of EOM-pCCD-based excited-state methods.

Method	CC operator	$\hat{R}$	Scaling	Memory
EOM-pCCD	$e^{\hat{T}_p}$	$\hat{R}_p$	$\mathcal{O}(o^2v^2)$	$\mathcal{O}(N^3)$
EOM-pCCD+S	$e^{\hat{T}_p}$	$\hat{R}_p + \hat{R}_1$	$\mathcal{O}(o^2v^3)$	$\mathcal{O}(o^2v^2)$
EOM-pCCD-CCS	$e^{\hat{T}_1 + \hat{T}_p}$	$\hat{R}_p + \hat{R}_1$	$\mathcal{O}(o^2v^3)$	$\mathcal{O}(o^2v^2)$
EOM-pCCD-LCCSD	$(1 + \hat{T}_1 + \hat{T}_2')e^{\hat{T}_p}$	$\hat{R}_1 + \hat{R}_2$	$\mathcal{O}(o^2v^4)$	$\mathcal{O}(v^4)$

$\hat{R}_p$, $\hat{R}_1$, and $\hat{R}_2$ are the corresponding excitation operators in the EOM ansatz for pair, single, and general double excitations, respectively. The $\hat{R}$ operator for all of these methods contains $\hat{R}_0 = r_0\hat{\tau}_0$ to account for the ground state contribution. o and v denote the number of occupied and virtual molecular orbitals, and $N = o + v$. Scaling and Memory indicate the formal computational cost and the total storage bottleneck (including integrals and intermediates), respectively.

1.5 Molecular property calculations with pCCD-based methods

For practical real-world purposes, quantum chemistry needs to go beyond calculating the energy of a chemical system by solving the electronic Schrödinger equation. It is essential to compute other quantities that characterize the atomic or molecular systems of our interest. Accurate theoretical estimation of structural, spectroscopic, and electronic properties—such as multipole moments and polarizabilities—is vital for validating theoretical models and establishing practical applications of computational chemistry.

Molecular properties associated with a given electronic state, such as dipole moment, polarizability, etc., describe the response of a molecular system to an external perturbation. Consider a molecule in an external electric field $\mathcal{E}$. The energy can be expanded in a Taylor series as (assuming $\mathcal{E}$ to be a weak perturbation)

$$E(\mathcal{E}) = E(0) + \left. \frac{dE}{d\mathcal{E}} \right|_{\mathcal{E}=0} \mathcal{E} + \frac{1}{2} \left. \frac{d^2E}{d\mathcal{E}^2} \right|_{\mathcal{E}=0} \mathcal{E}^2 + \dots \quad (1.79)$$

Here, the first-order interaction with $\mathcal{E}$ is the molecular dipole moment μ , the second-order interaction is the molecular polarizability α , etc. However, for non-variational wave functions, such as in the CC theory, the Hellmann–Feynman theorem does not generally hold. Consequently, properties cannot be evaluated as simple expectation values of the corresponding operators, i.e.,

$$\mu \neq \langle \Psi | \hat{\mu} | \Psi \rangle, \quad (1.80)$$

since in such a case, this expression does not correspond to the derivative of the energy with respect to the external perturbation. This is true for pCCD as well. To obtain molecular properties within CC theory, one instead employs the *response* function approach [86, 87], which essentially simplifies to the Lagrangian formalism of Jørgensen and Helgaker in the case of time-independent static properties [88, 89]. The CC Lagrangian is defined as

$$\mathcal{L} = \langle \Phi_0 | (1 + \hat{\Lambda}) \bar{H} | \Phi_0 \rangle, \quad (1.81)$$

where $\bar{H} = e^{-\hat{T}} \hat{H} e^{\hat{T}}$ is the similarity-transformed Hamiltonian and $\hat{\Lambda}$ is the de-excitation operator. Molecular properties are then obtained as derivatives of the Lagrangian with respect to the external perturbation. For example, the dipole

moment is given by

$$\mu = - \left. \frac{\partial \mathcal{L}}{\partial \mathcal{E}} \right|_{\mathcal{E}=0} = \langle \Phi_0 | (1 + \hat{\Lambda}) \bar{\mu} | \Phi_0 \rangle, \quad (1.82)$$

where $\bar{\mu} = e^{-\hat{T}} \hat{\mu} e^{\hat{T}}$ is the similarity-transformed dipole operator.

This framework naturally leads to the concept of response density matrices [90, 91]. Specifically, the One-Particle Reduced Density Matrix (1-RDM) can be identified as the first-order response of the CC Lagrangian with respect to the one-electron Hamiltonian elements h_{pq} ,

$$\gamma_q^p = \frac{\partial \mathcal{L}}{\partial h_{qp}}. \quad (1.83)$$

Evaluating this derivative within the CC Lagrangian framework yields

$$\gamma_q^p = \langle \Phi_0 | (1 + \hat{\Lambda}) e^{-\hat{T}} \hat{a}_p^\dagger \hat{a}_q e^{\hat{T}} | \Phi_0 \rangle, \quad (1.84)$$

which defines the relaxed (response-corrected) 1-RDM in CC theory.

Using this quantity, molecular one-electron properties can be obtained easily as contractions of the 1-RDM with the corresponding property integrals. For example, the dipole moment becomes

$$\mu = \sum_{pq} \gamma_q^p \mu_{pq}, \quad (1.85)$$

where μ_{pq} are the dipole integrals in the molecular orbital basis.

1.5.1 Response approach for pCCD and post-pCCD methods

The 1-RDM in the pCCD framework is defined as [23]

$${}^{\text{pCCD}}\gamma_q^p = \langle \Phi_0 | (1 + \hat{\Lambda}) e^{-\hat{T}_p} \hat{a}_p^\dagger \hat{a}_q e^{\hat{T}_p} | \Phi_0 \rangle, \quad (1.86)$$

where the cluster excitation and de-excitation operators are given by

$$\hat{T}_p = \sum_{ia} t_{ii}^{aa} \hat{a}_a^\dagger \hat{a}_a^\dagger \hat{a}_{\bar{i}} \hat{a}_i, \quad \hat{\Lambda}_p = \sum_{ia} \lambda_{ii}^{aa} \hat{a}_i^\dagger \hat{a}_{\bar{i}}^\dagger \hat{a}_{\bar{a}} \hat{a}_a. \quad (1.87)$$

Here, indices p, q, r, s denote general spin-orbital (or spatial orbital) labels, whereas the subscript p in $\hat{T}_p$ refers to the pair excitations. The similarity-transformed operator is expanded using the BCH expansion is

$$\bar{X} = e^{-\hat{T}_p} \hat{X} e^{\hat{T}_p} = \hat{X} + [\hat{X}, \hat{T}_p] + \frac{1}{2} [[\hat{X}, \hat{T}_p], \hat{T}_p] + \dots \quad (1.88)$$

For the one-electron operator $\hat{X} = \hat{a}_p^\dagger \hat{a}_q$, the expansion truncates at first order due to the structure of $\hat{T}_p$,

$$e^{-\hat{T}_p} \hat{a}_p^\dagger \hat{a}_q e^{\hat{T}_p} = \hat{a}_p^\dagger \hat{a}_q + [\hat{a}_p^\dagger \hat{a}_q, \hat{T}_p]. \quad (1.89)$$

Substituting into the definition of γ_q^p gives

$$\gamma_q^p = \langle \Phi_0 | (1 + \hat{\Lambda}_p) \left(\hat{a}_p^\dagger \hat{a}_q + [\hat{a}_p^\dagger \hat{a}_q, \hat{T}_p] \right) | \Phi_0 \rangle. \quad (1.90)$$

The first term yields

$$\langle \Phi_0 | \hat{a}_p^\dagger \hat{a}_q | \Phi_0 \rangle = \delta_{pq} n_p, \quad (1.91)$$

where $n_p = 2$ for occupied spatial orbitals and $n_p = 0$ for virtual orbitals in the closed-shell, spin-summed formulation. The second term contributes only through contractions involving $\hat{\Lambda}_p$,

$$\langle \Phi_0 | \hat{\Lambda}_p [\hat{a}_p^\dagger \hat{a}_q, \hat{T}_p] | \Phi_0 \rangle. \quad (1.92)$$

Using the pair structure of $\hat{T}_p$ and $\hat{\Lambda}_p$, only fully contracted terms survive. Due to the seniority-zero restriction, all off-diagonal elements vanish

$$\gamma_q^p = \delta_{pq} \gamma_p^p. \quad (1.93)$$

Evaluating the remaining contractions yields the diagonal elements as follows.

Occupied orbitals ($p = i$):

$$\gamma_i^i = 2 \left(1 - \sum_a t_{ii}^{aa} \lambda_{ii}^{aa} \right). \quad (1.94)$$

Virtual orbitals ($p = a$):

$$\gamma_a^a = 2 \sum_i t_{ii}^{aa} \lambda_{ii}^{aa}. \quad (1.95)$$

Thus, the pCCD 1-RDM is diagonal in the orbital basis. In a similar way, the Two-Particle Reduced Density Matrix (2-RDM) can also be obtained as

$$\text{pCCD} \Gamma_{pq}^{rs} = \langle \Phi_0 | (1 + \hat{\Lambda}_p) e^{-\hat{T}_p} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r e^{\hat{T}_p} | \Phi_0 \rangle, \quad (1.96)$$

which also has a sparse structure for pCCD. Expressions of the non-zero terms here can be found in Henderson et al. [23].

A particular point to note is that the pCCD one- and two-particle density matrices are also used in the orbital optimization procedure. Consequently, their evaluation does not introduce additional computational cost. Specifically, the pCCD 1-RDM is diagonal in the natural orbital basis, while the 2-RDM retains a highly sparse structure due to the seniority-zero restriction of the wave function. This sparsity significantly simplifies the evaluation of orbital gradients and Hessian elements. In parallel, this also facilitates easy estimation of one-electron properties with pCCD.

Through the density matrices, electron correlation effects can also be depicted. A quantitative measure of orbital interactions can be obtained from the single-orbital entropy $s_i^{(1)}$ and the two-orbital entropy $s_{ij}^{(2)}$ [92, 93]. The single-orbital entropy $s_i^{(1)}$ captures how strongly orbital i is entangled with the rest of the system and is defined as

$$s_i^{(1)} = - \sum_{\alpha=1}^4 \omega_{\alpha,i} \ln \omega_{\alpha,i}, \quad (1.97)$$

where $\omega_{\alpha,i}$ are the eigenvalues of the One-Orbital Reduced Density Matrix (1-ORDM). The dimension of the 1-ORDM matches that of the one-orbital Fock space and is 4 for spatial orbitals (unoccupied, singly occupied with α or β , and doubly occupied). Analogously, the two-orbital entropy $s_{ij}^{(2)}$ measures the joint correlation of a pair of orbitals i and j with the remaining system as

$$s_{ij}^{(2)} = - \sum_{\alpha=1}^{16} \omega_{\alpha,ij} \ln \omega_{\alpha,ij}, \quad (1.98)$$

where $\omega_{\alpha,ij}$ are the eigenvalues of the Two-Orbital Reduced Density Matrix (2-ORDM). In contrast to n -particle RDMs that are defined for a constant number of particles, the one- and two-orbital RDMs are constructed from contributions of

the 1-, 2-, 3-, and 4-particle RDMs. Expressions for Three-Particle Reduced Density Matrix (3-RDM) and Four-Particle Reduced Density Matrix (4-RDM) obtained with pCCD can be found in Ref. [94, 95]. The total correlation between orbitals i and j is given by the orbital-pair mutual information

$$I_{ij} = s_i^{(1)} + s_j^{(1)} - s_{ij}^{(2)}, \quad i \neq j. \quad (1.99)$$

However, analysis of electron correlation using single-orbital entropy and orbital-pair mutual information based on pCCD density matrices shows that the pCCD ansatz tends to overestimate correlation effects, even though it yields reasonably accurate energies [96]. Thus, it is a pertinent question whether *a posteriori* corrections on pCCD can cure this issue. Among such methods, the linearized CC correction-based fpLCC method has been tested for this purpose [97]. The response 1-RDM for the fpLCC wave function can be constructed as the sum of the reference pCCD response contribution and the additional correlation correction arising from the LCC correction on pCCD. The pCCD contribution is obtained from the reference response 1-RDM, while the LCC correction is evaluated using the corresponding $\hat{\Lambda}$ -equations.

$${}^{\text{LCC}}\gamma_q^p = \langle \Phi_0 | (1 + \hat{\Lambda}'_{\text{LCC}}) \left\{ e^{-\hat{T}_p - \hat{T}'} \hat{a}_p^\dagger \hat{a}_q e^{\hat{T}_p + \hat{T}'} \right\}_{L'} | \Phi_0 \rangle, \quad (1.100)$$

Here, $\hat{T}'$ denotes the additional LCC excitation amplitudes. The cluster de-excitation operator in this case is defined as

$$\hat{\Lambda}'_n = \frac{1}{(n!)^2} \sum_{ij\dots} \sum'_{ab\dots} \hat{\lambda}_{ab\dots}^{ij\dots} \hat{a}_i^\dagger \hat{a}_a \hat{a}_j^\dagger \hat{a}_b \dots, \quad (1.101)$$

where the prime indicates that pure electron-pair de-excitations are excluded, as they do not contribute to the LCC equations. In Eq. (1.100), only terms up to linear order in $\hat{T}'$ are retained, as indicated by the subscript L' . Here also, $\hat{T}'$ can only contain unpaired doubles ($\hat{T}' = \hat{T}'_2$) or both unpaired doubles as well as singles ($\hat{T}' = \hat{T}'_1 + \hat{T}'_2$).

The total response 1-RDM in the fpLCC(S)D framework is then given by

$$\gamma_q^p = {}^{\text{pCCD}}\gamma_q^p + {}^{\text{LCC}}\gamma_q^p. \quad (1.102)$$

whereas ${}^{\text{LCC}}\gamma_q^p$ accounts exclusively for the additional correlation effects introduced

by the LCC correction.

To evaluate the response density matrices, the $\lambda_{ij\dots}^{ab\dots}$ amplitudes are required, which are determined by solving the $\hat{\Lambda}$ -equations associated with the LCC corrections. These are obtained by enforcing stationarity of the LCC Lagrangian with respect to the LCC excitation amplitudes t_ν .

The LCC Lagrangian is defined as

$$\mathcal{L} = E(\mathbf{t}, \boldsymbol{\lambda}) = \langle \Phi_0 | (1 + \hat{\Lambda}') \left\{ e^{-\hat{T}' - \hat{T}_p} \hat{H} e^{\hat{T}_p + \hat{T}'} \right\}_{L'} | \Phi_0 \rangle. \quad (1.103)$$

Stationarity with respect to the LCC amplitudes t_ν requires

$$\frac{\partial \mathcal{L}}{\partial t_\nu} = \frac{\partial}{\partial t_\nu} \langle \Phi_0 | (1 + \hat{\Lambda}') \left\{ e^{-\hat{T}' - \hat{T}_p} \hat{H} e^{\hat{T}_p + \hat{T}'} \right\}_{L'} | \Phi_0 \rangle = 0, \quad (1.104)$$

which can be rewritten as

$$\langle \Phi_0 | (1 + \hat{\Lambda}') \left\{ e^{-\hat{T}' - \hat{T}_p} \left[\hat{H}, \hat{\tau}_\nu \right] e^{\hat{T}_p + \hat{T}'} \right\}_{L'} | \Phi_0 \rangle = 0, \quad (1.105)$$

where $\hat{\tau}_\nu$ denotes the excitation operator associated with the amplitude t_ν .

To summarize, these frameworks provide us with ways of accessing molecular properties such as multipole moments using pCCD and post-pCCD density matrices. Additionally, in the Linear Response (LR) formulation of coupled-cluster (LR-CC) theory, excited-state properties can be computed, including excitation energies, transition dipole moments, and oscillator strengths, etc [98, 99, 100]. The recently implemented Linear Response Pair Coupled Cluster Doubles (LR-pCCD) paves the path for that within the pCCD framework [101].

1.5.2 Expectation value based approach for pCCD and post-pCCD methods

An alternative route to estimate molecular properties is the expectation value approach. The expectation value of an observable corresponding to a general operator $\hat{X}$ in a state described by the wave function Ψ is given by

$$\bar{X} = \frac{\langle \Psi | \hat{X} | \Psi \rangle}{\langle \Psi | \Psi \rangle}. \quad (1.106)$$

Substituting $|\Psi\rangle = e^{\hat{T}}|\Phi_0\rangle$ for the CC wave function into the expectation value

expression yields

$$\bar{X} = \frac{\langle \Phi_0 | e^{\hat{T}^\dagger} \hat{X} e^{\hat{T}} | \Phi_0 \rangle}{\langle \Phi_0 | e^{\hat{T}^\dagger} e^{\hat{T}} | \Phi_0 \rangle}. \quad (1.107)$$

A major obstacle in the direct evaluation of Eq. (1.107) is the appearance of non-terminating BCH expansions arising from the operators $e^{\hat{T}}$ and $e^{\hat{T}^\dagger}$. This makes a straightforward truncation difficult and computationally intractable. To address this issue, an alternate approach was proposed by Jeziorski and Moszyński [102], in which a new excitation operator $\hat{S}$ was introduced. The operator $\hat{S}$ is defined as a sum of n -body excitation components,

$$\hat{S} = \hat{S}_1 + \hat{S}_2 + \hat{S}_3 + \dots,$$

and satisfies the relation

$$e^{\hat{S}} | \Phi_0 \rangle = \frac{e^{\hat{T}^\dagger} e^{\hat{T}} | \Phi_0 \rangle}{\langle \Phi_0 | e^{\hat{T}^\dagger} e^{\hat{T}} | \Phi_0 \rangle}. \quad (1.108)$$

The $\hat{S}$ operator is a connected quantity, as demonstrated in Ref. [102], and satisfies a linear equation involving a finite number of nested commutators of $\hat{T}$ and $\hat{T}^\dagger$. This property ensures that the expectation value formulation remains size-extensive. The connected nature of the $\hat{S}$ operator has been exploited in several works to formulate CC propagators and to compute various single-electron properties. [103, 104, 105] In particular, Korona and co-workers have shown that, for properties such as multipole moments and electrostatic interaction energies, the expectation value approach yields results comparable in accuracy to those obtained from response CCSD. [104, 106, 107]

The crucial advantage of this approach is that the 1-RDM can be constructed using the cluster amplitudes. Unlike the response approach, here we do not require any additional computation. A concise derivation following Korona and Jeziorski [106] of the 1-RDM using the expectation value approach is shown in [Appendix A](#).

1.6 Quantum chemical embedding methods

From the discussion so far, we can agree that pCCD, to a large extent, has the potential to become a standard tool for getting reliable qualitative results for strongly correlated systems, with its formal $\mathcal{O}(N^4)$ scaling. However, achieving chemical accuracy still needs post-pCCD steps, which bring additional costs to the order of $\mathcal{O}(N^6)$. This is prohibitively high for large chemical structures, effectively hindering the wide application of pCCD-based methods. Studying excited states with pCCD also involves steep computational burden.

However, it is also found that in complex systems, the region of our interest, the so-called *active* region, is often rather a small part of a much larger whole structure. The remaining part of the structure plays its role as the *environment*, interacting with the active region through intra- or intermolecular interactions. Quantum embedding exploits this fact by employing two different types of quantum simulations: a high-level, accurate calculation for the active region of interest, and a lower-level, more approximate treatment for the surrounding environment. This fragmentation approach is a way to reduce the overall complexity of the problem while maintaining the desired level of accuracy [108, 109, 110, 111].

A primary classification of such embedding methods can be done based on how the environment is approximated: QM/MM methods [112, 113, 114], which use parametrized force fields obtained from molecular mechanics, and QM/QM methods, which treat both the active and environment regions at the QM level. While QM/MM is computationally much cheaper and can handle extremely large systems, especially in condensed phases, quantum mechanical information about interactions between the environment and active region is lost. QM/QM embedding, on the other hand, explicitly treats the atoms/molecules surrounding the active region at the QM level, thereby improving the description of the environment effect. The quality of the approximations with which we model the inter-fragment interactions, and the definition of the boundaries, determine how close we can go to a full supramolecular high-level QM description of the whole system.

Any embedding approach relies on the partitioning of a full system into an active system part and the rest as its environment. The Frozen Density Embedding (FDE) method, based on early formulations of Cortona [115, 116] and Wesolowski and Warshell [117], provides a DFT-based framework for such partitioning. Here, the total electron density of a complex system is decomposed into distinct, interacting parts as

$$\rho_{\text{tot}} = \rho_I + \rho_{II}. \quad (1.109)$$

In this partitioning, ρ_{tot} represents the density of the entire supramolecular system. The density of the primary region where chemical interest is focused is denoted as ρ_I , while ρ_{II} represents the density of the surrounding, “frozen” environment. This idea of the density-partitioning stems from subsystem DFT [118, 119, 120], where the total energy is minimized with respect to individual subsystem densities. It should be noted that ρ_{II} in Eq. (1.109) does not necessarily represent one single fragment, as the environment often consists of several moieties, i.e., $\rho_{II} = \sum_{i \notin \text{active}} \rho_i$ where i are all fragments in the system other than the active part. Using this decomposition, the total supramolecular energy is calculated as the sum of the energies of each fragment plus an interaction energy term $E_{\text{int}}[\rho_I, \rho_{II}]$,

$$E_{\text{tot}}[\rho_I + \rho_{II}] = E_I[\rho_I] + E_{II}[\rho_{II}] + E_{\text{int}}[\rho_I, \rho_{II}]. \quad (1.110)$$

The energy for an isolated subsystem i is defined by the following expression

$$\begin{aligned} E_i[\rho_i] = & \int \rho_i(\mathbf{r}) v_{\text{nuc}}^i(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \iint \frac{\rho_i(\mathbf{r}) \rho_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \\ & + V_{\text{nuc, nuc}}^i + E_{\text{xc}}[\rho_i] + T_s[\rho_i]. \end{aligned} \quad (1.111)$$

Within this equation, the terms account for the nuclear attraction (v_{nuc}^i), the classical electron-electron Coulomb interaction, the nuclear repulsion energy ($V_{\text{nuc, nuc}}^i$), the exchange–correlation energy (E_{xc}), and the non-interacting kinetic energy (T_s). The coupling between subsystems I and II is captured in the interaction term, which accounts for the cross-interactions between densities and nuclei.

$$\begin{aligned} E_{\text{int}}[\rho_I, \rho_{II}] = & \int \rho_I(\mathbf{r}) v_{\text{nuc}}^{II}(\mathbf{r}) d\mathbf{r} + \int \rho_{II}(\mathbf{r}) v_{\text{nuc}}^I(\mathbf{r}) d\mathbf{r} + \iint \frac{\rho_I(\mathbf{r}) \rho_{II}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \\ & + V_{\text{nuc, nuc}}^{I-II} + E_{\text{xc}}^{\text{nadd}}[\rho_I, \rho_{II}] + T_s^{\text{nadd}}[\rho_I, \rho_{II}]. \end{aligned} \quad (1.112)$$

Finally, the non-additive exchange–correlation and kinetic energy components are defined as the difference between the property of the total system and the sum of the properties of the isolated fragments.

$$E_{\text{xc}}^{\text{nadd}}[\rho_I, \rho_{II}] = E_{\text{xc}}[\rho_I + \rho_{II}] - E_{\text{xc}}[\rho_I] - E_{\text{xc}}[\rho_{II}]. \quad (1.113)$$

$$T_s^{\text{nadd}}[\rho_I, \rho_{II}] = T_s[\rho_I + \rho_{II}] - T_s[\rho_I] - T_s[\rho_{II}] \quad (1.114)$$

While the electrostatic terms in Eq. (1.112) can be calculated very easily, the non-classical contributions from the non-additive terms can only be expressed with density functionals.

It should be noted that the subsystem electron densities change when they interact; hence, densities calculated in isolation do not represent the true density of the subsystem in the actual supramolecule, i.e., the active subsystem gets polarized by the environment. To model this polarization, the environmental effect needs to be included in the quantum-chemical description of the active subsystem. If we follow the DFT formulation, the densities can be obtained from Kohn-Sham (KS) orbitals as $\rho_I = \sum_i |\chi_i^I(\mathbf{r})|^2$. For a given frozen electron density ρ_{II} of subsystem II , the orbitals and the electron density of the active subsystem I can then be determined by minimizing the total energy with respect to ρ_I , while keeping ρ_{II} frozen. Performing this minimization under the constraint that the number of electrons N_I in subsystem I is conserved leads to a set of equations for the orbitals of subsystem I :

$$\left[-\frac{1}{2}\nabla^2 + V_{\text{nuc}}^I(\mathbf{r}) + \int \frac{\rho_I(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r}' + \frac{\delta E_{\text{xc}}[\rho_I]}{\delta \rho_I(\mathbf{r})} + v_{\text{emb}}^I[\rho_I, \rho_{II}](\mathbf{r}) \right] \chi_i^I(\mathbf{r}) = \varepsilon_i^I \chi_i^I(\mathbf{r}). \quad (1.115)$$

The derivative of the coupling energy term defined in Eq. (1.112) with respect to the subsystem density of I gives the embedding potential, which accounts for the presence of the frozen environment II when determining the electron density of the active subsystem.

$$v_{\text{emb}}^I[\rho_I, \rho_{II}](\mathbf{r}) = V_{\text{nuc}}^{II}(\mathbf{r}) + \int \frac{\rho_{II}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r}' + \left. \frac{\delta E_{\text{xc}}^{\text{nadd}}[\rho_I, \rho_{II}]}{\delta \rho(\mathbf{r})} \right|_{\rho_I} + \left. \frac{\delta T_s^{\text{nadd}}[\rho_I, \rho_{II}]}{\delta \rho(\mathbf{r})} \right|_{\rho_I}. \quad (1.116)$$

In an exact embedding approach, the density of the active subsystem should combine with the frozen environment density to reproduce the total density of the full system. This is only possible if the frozen density is physically consistent and should not be larger than the true total density anywhere. However, approximate frozen densities

obtained from isolated subsystems often violate this condition [121]. This problem is tackled by considering the polarization effect of the environment by the active system as well. This gives us a set of coupled equations,

$$\left[-\frac{1}{2}\nabla^2 + V_{\text{nuc}}^I(\mathbf{r}) + \int \frac{\rho_I(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r}' + \frac{\delta E_{\text{xc}}[\rho_I]}{\delta \rho_I(\mathbf{r})} + v_{\text{emb}}^I[\rho_I, \rho_{II}](\mathbf{r}) \right] \chi_i^I(\mathbf{r}) = \varepsilon_i^I \chi_i^I(\mathbf{r}), \quad (1.117)$$

and

$$\left[-\frac{1}{2}\nabla^2 + V_{\text{nuc}}^{II}(\mathbf{r}) + \int \frac{\rho_{II}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r}' + \frac{\delta E_{\text{xc}}[\rho_{II}]}{\delta \rho_{II}(\mathbf{r})} + v_{\text{emb}}^{II}[\rho_I, \rho_{II}](\mathbf{r}) \right] \chi_i^{II}(\mathbf{r}) = \varepsilon_i^{II} \chi_i^{II}(\mathbf{r}). \quad (1.118)$$

As can be seen from these two equations, the embedding potential terms, consequently, the densities of the two subsystems are interdependent. Hence, these equations need to be solved in a self-consistent fashion. This is achieved by employing the *freeze-and-thaw* cycles [122]. First the density of subsystem *I* is obtained in the presence of an approximate frozen density for subsystem *II*. Thereafter, the roles of the two subsystems are swapped, the density calculated for subsystem *I* in the previous step is now kept frozen, whereas subsystem *II* is now thawed, and an updated density is determined for it. This is repeated iteratively til the subsystem densities reach convergence. Each freeze-and-thaw step will involve two microiterations for active and environment subsystems. Alternatively, the coupled equations for all the subsystems can be solved simultaneously. However, because the Kohn–Sham orbitals of different subsystems are not orthogonal, Pauli repulsion is underestimated and not properly corrected by the kinetic energy terms. As a result, in the case of simultaneous solving of the coupled equations, the electron density becomes too diffuse, underestimating the polarization effects, as reported by Iannuzzi et al. [123].

The simple framework of FDE has also been extended to excited-state calculations. A state-specific Δ DFT approach was used in this context, using the embedding potential obtained for the ground state also for the excited states [124, 125, 126]. Excited states can also be accessed by applying TDDFT. In the simplest formulation, the response equations of the active and the environment subsystems are decoupled [127, 128, 129]. An efficient implementation for coupling the response of subsystems was presented by Neugebauer et al. [130, 131].

A particularly attractive feature of FDE is that the embedding potential can also be derived when the active subsystem is described using a wavefunction-based method.

Carter and co-workers developed a density-based embedding framework that extends beyond conventional DFT by incorporating electron correlation explicitly within the embedded region [132]. This formulation was designed to address a broad class of problems, ranging from adsorption on metal surfaces [133] to strongly correlated phenomena such as the Kondo effect [134]. The formal validity of this approach, in the limit of an exact density functional combined with FCI, was demonstrated by Wośowski [135]. While the resulting embedding equations share the same formal structure of the embedding potential as in the earlier formulation, their derivation follows a different route. In contrast to Cortona’s approach [115], which is based on the variational minimization of the total DFT energy with respect to the set of one-electron orbitals χ_i^I spanning the embedded subsystem I , the Carter framework builds on an explicit treatment of correlation effects in the active subsystem. In these so-called WFT-in-DFT approaches, the interaction energy is described as

$$E_{\text{int}} \approx E_{\text{tot}}^{\text{DFT}}[\rho_{\text{tot}}] - E_I^{\text{DFT}}[\rho_I] - E_{II}^{\text{DFT}}[\rho_{II}]. \quad (1.119)$$

Instead of seeking a solution that minimizes the total DFT energy, one considers a more general variational principle with respect to the embedded-subsystem density ρ_I . Assuming that the effect of change in electron density due to inter-subsystem interactions is sufficiently localized within I and that the environment II is sufficiently large, it is reasonable to assume that $\delta E_{II}^{\text{DFT}}[\rho_{II}]/\delta \rho_I = 0$. In such a case, we neglect the polarization of the environment by the active subsystem. This leads to the equation

$$\frac{\delta E_{\text{tot}}[\rho_{\text{tot}}]}{\delta \rho_I(\mathbf{r})} = \frac{\delta E_I[\rho_I]}{\delta \rho_I(\mathbf{r})} + v_{\text{emb}}[\rho_I, \rho_{\text{tot}}](\mathbf{r}) = \mu, \quad (1.120)$$

where μ is a Lagrange multiplier enforcing the constraint $\int d\mathbf{r} \rho_I(\mathbf{r}) = N_I$, and v_{emb} denotes the embedding potential,

$$v_{\text{emb}}[\rho_I, \rho_{\text{tot}}](\mathbf{r}) = \frac{\delta E_{\text{tot}}^{\text{DFT}}[\rho_{\text{tot}}]}{\delta \rho_{\text{tot}}(\mathbf{r})} - \frac{\delta E_I^{\text{DFT}}[\rho_I]}{\delta \rho_I(\mathbf{r})}. \quad (1.121)$$

The density ρ_I satisfying Eq. (1.120) can equivalently be obtained by solving the Schrödinger equation for the embedded subsystem,

$$\left[\hat{H}_I + \sum_j^{N_I} v_{\text{emb}}[\rho_I, \rho_{\text{tot}}](\mathbf{r}_j) \right] \Psi_I = E_I \Psi_I, \quad (1.122)$$

where $\hat{H}_I$ is the Hamiltonian of the active embedded subsystem, and Ψ_I is the

corresponding wave function from which ρ_I can be obtained.

Of the three density terms, only two are independent, so in principle one should minimize the energy with respect to the two chosen independent variables, subject to the constraint that $\rho_{\text{tot}} = \rho_I + \rho_{II}$. But this has not been implemented in practice. Instead, the density of the active subsystem is obtained with CI or CASSCF and is updated iteratively using Eq. (1.122), while ρ_{tot} obtained from a periodic DFT calculation is kept fixed. This strategy was used for the description of ground and excited states of CO molecule adsorbed on metallic (Cu, Pt, Pd) surfaces [133, 136, 137, 138]. The limitations of this strategy were addressed in later work, where the constraint of a fixed total density was relaxed. Instead, the environment density is defined as $\rho_{II} = \rho_{\text{tot}} - \rho_I^{\text{iso}}$, where the total density is obtained from a periodic DFT calculation and ρ_I^{iso} denotes the density of the isolated subsystem I . This environment density is subsequently kept frozen; thus, environmental polarization is accounted for only at the level of the periodic DFT calculation and is not updated self-consistently with the wave function-based subsystem treatment. ρ_I is obtained self-consistently from Eq. (1.122) and ρ_{tot} is calculated as sum of this updated ρ_I and the constant ρ_{II} . This strategy was found to show more numerically stable performance [139, 140, 141].

A simplified WFT-in-DFT embedding scheme has been proposed by Gomes et al. in which the embedding potential is obtained from a preceding DFT-in-DFT embedding calculation, either using a fixed approximate environment density or an environment density polarized through freeze-and-thaw cycles. This approach has been applied to investigate the electronic excitations of NpO_2^{2+} [142] and UO_2^{2+} [143] embedded in ionic crystals with the Intermediate Hamiltonian Fock-Space Coupled Cluster (IHFSCC) framework. Höfener et al. extended the response formalism to WFT-in-DFT [144, 145]. Instead of DFT, densities obtained from CC methods have also been used in the embedding context, i.e., in WFT-in-WFT fashion [146, 147].

To summarize, the QM/QM embedding schemes are best viewed as flexible, easily implementable frameworks that allow for different practical approximations to achieve a suitable balance between accuracy and computational cost. In this context, FDE provides a particularly appealing formulation. Moreover, when combined with response theory, this framework can be extended to the accurate modeling of electronically excited states, since the FDE-based approaches offer the additional advantage of capturing the coupling between excitations in different subsystems, as all regions are treated at the QM level of theory.

Chapter 2

Summary of the publications

2.1 Published articles comprising the thesis

The research presented in this thesis is derived from the articles listed below, which were prepared or published during the course of the doctoral study. Below each article’s data, the article’s main research context and methodology are briefly presented. All the methods discussed here have been implemented in the PyBEST software package [148, 149, 150]. A summary of the results is presented in Section 3.1.

1. Static embedding with pair coupled cluster doubles based methods

Rahul Chakraborty, Katharina Boguslawski, and Paweł Tecmer, *Phys. Chem. Chem. Phys.*, 2023, 25, 25377, DOI: 10.1039/d3cp02502k.

This work presents a new WFT-in-DFT embedding framework based on pCCD-type approaches. We employ the fpLCCSD ansatz in the pCCD optimized Molecular Orbital (MO) basis, i.e., oo-fpLCCSD, for ground states of the active subsystem and extend it to excited states using the EOM formalism. The environment was represented using a static embedding potential obtained from separate DFT-in-DFT calculations, where the subsystem densities were relaxed with freeze-and-thaw cycles [142]. In static embedding theory, the exact response of the subsystem densities, in principle, is achievable with an FCI calculation in a complete basis set. However, a limitation arises in our case, as only the response of the active subsystem is treated explicitly at the oo-fpLCCSD level. Another approximation is the usage of a fixed potential

obtained using the ground state subsystem densities. However, for localized excitations, this approximation is expected to have a minor effect.

The embedding potentials were generated with the Amsterdam Modeling Suite (AMS2022) [151, 152] and then extracted using the PyADF scripting framework. [153] A particularly important question in any DFT FDE scheme is the choice of approximate functionals for the non-additive kinetic term. The GGA-based PW91k functional of Lembarki and Chermette [154] has been recommended for such purpose. [155, 156] In this work, we also used this functional. For the uranyl tetrahalides, we used the second-order scalar relativistic Douglas–Kroll–Hess (DKH) Hamiltonian. [157, 158]

The resulting EOM-fpLCCSD-in-DFT approach was benchmarked for the VEEs of hydrogen-bonded $\text{H}_2\text{O} \cdots \text{NH}_3$ complex, micro-solvated thymine, and uranyl tetrahalides ($\text{UO}_2\text{X}_4^{2-}$, $\text{X} = \text{F}, \text{Cl}, \text{Br}$).

In addition, the embedding potentials for the uranyl tetrahalides were evaluated using orbital entanglement and correlation analysis, as described in 1.5.1. These entanglement measures are orbital-dependent and sensitive to environmental effects. Changes in the surroundings lead to variations in the single-orbital entropies and the orbital-pair mutual information, allowing one to track how orbital correlations evolve. Specifically, these quantities differ between isolated subsystems and full supramolecular systems, while a properly embedded system should reproduce the qualitative features of the latter. Therefore, these measures provide a useful tool for analyzing the quality of the embedding potential in our case.

2. Toward reliable dipole moments without single excitations: The role of orbital rotations and dynamical correlation

Rahul Chakraborty, Matheus Morato F. de Moraes, Katharina Boguslawski, Artur Nowak, Julian Świerczyński, and Paweł Tecmer, *J. Chem. Theory Comput.* 2024, 20, 4689–4702, DOI: 10.1021/acs.jctc.4c00471.

This study investigates the accuracy of dipole moments calculated using pCCD and its fpLCC extensions. While molecular properties have been previously studied with APSG [159, 160], or with the related Piris Natural Orbital Functional 5 (PNOF5) [161, 162, 163], AP1roG/pCCD and post-pCCD methods have not been assessed in this context. To that end, a set of main-group diatomics with a diverse set of bonding patterns was chosen. Specifically,

we compared these values calculated with canonical HF as well as pCCD-optimized MO bases. The fpLCC response density matrices were computed as described in Nowak et al. [97]. Considering the established accuracy of CCSD(T) dipole moment values in comparison to the experimental numbers, we used CCSD(T) as our theoretical reference [164]. Dipole moments of some model complexes were also tested, using our previous embedding approach, where we extracted the embedding potential from a preceding DFT-in-DFT FDE calculation, and then performed pCCD and fpCC calculations. Dipole moment surfaces of HF and CO molecules were also constructed with our methods and compared with theoretical references.

3. Expectation value-pCCD-based methods for single-electron properties

Rahul Chakraborty, Somayeh Ahmadkhani, Julian Świerczyński, Katharina Boguslawski, and Paweł Tecmer, *J. Phys. Chem. A* 2025, 129, 6713–6732, DOI: 10.1021/acs.jpca.5c03859.

In this work, we follow the approach of Jeziorski and Moszyński [102], introducing an alternative excitation operator defined in Eq. (1.108). Previous studies have shown that, for properties such as multipole moments and electrostatic interaction energies, the expectation value approach yields results comparable to response CCSD [106, 107]. In these works, the expansion of $\hat{S}$ was truncated at the same level as standard CCSD. Here, we use this idea for pCCD and post-pCCD methods, assessing the performance of the expectation value CC approach with these ansätze.*

The performance of the expectation value approach with pCCD, fpCC, and fpLCC was assessed against the dipole and quadrupole moments of molecules with diverse chemical characteristics. We benchmarked our results against response CCSD(T) values obtained with canonical HF orbitals, where pCCD and post-pCCD methods were evaluated using both canonical HF and pCCD-optimized MOs.

4. Frozen density embedding with pCCD electron densities

Rahul Chakraborty, and Paweł Tecmer, arXiv:2604.14904, 2026, DOI: 10.48550/arXiv.2604.14904.

*The indices in the expectation value 1-RDM expressions (Eqs. 19–21 in Ref. [165]) are erroneous. Correct expressions are given in Appendix A and also in Ref. [107]

In this work, we present a simple density-embedding scheme based on pCCD electron densities. Its efficiency stems from the low cost of the pCCD $\hat{\Lambda}$ -equations, which enable straightforward evaluation of electron densities. The implemented pCCD-in-pCCD method provides us with a single, simple workflow in PyBEST, which can be used for a wide range of molecules, where DFT can be of limited application.

Freeze-and-thaw cycles are employed, where the subsystem densities obtained from pCCD are updated till convergence of the subsystem energies. Initially, the environment subsystem is treated in isolation at the supramolecular geometry. HF and subsequent pCCD calculations are performed to obtain an initial density, which serves as a starting point for the freeze-and-thaw procedure. The corresponding embedding potential is evaluated on a numerical grid, and its one-electron matrix elements are computed. This is then added to the Fock matrix of the active subsystem A , and the HF equations are solved, allowing its orbitals to relax in the presence of the environment subsystem B . The resulting orbitals are then used in a pCCD calculation in the presence of the static embedding potential. Next, an updated embedding potential is constructed from the pCCD density of the active subsystem A , and the roles of the two subsystems are exchanged. For the pilot test cases, the Slater potential and Thomas-Fermi functional were used for the non-additive terms in the embedding potential. Orbital optimization with pCCD was not performed primarily to avoid the computationally expensive orbital rotations in each freeze-and-thaw macro-iteration. In addition, EOM-fpLCCSD excitation energies show minor differences when using HF orbitals instead of pCCD orbitals [85]. The method is validated by computing dipole moments of $\text{CO}_2 \cdots \text{Rg}$ ($\text{Rg} = \text{He}, \text{Ne}, \text{Ar}, \text{Kr}$) complexes and vertical excitation energies of selected microsolvated systems.

2.2 Other published articles

Additional publications and preprints by the author, which fall outside the scope of this thesis, are listed below. Below each article’s data, the article abstract is presented.

A. Published during doctoral studies

1. Scalable quantum circuit generation for iterative ground state approximation using Majorana Propagation

Rahul Chakraborty, Aaron Miller, Anton Nykänen, Özlem Salehi, Fabio Tarocco, Fabijan Pavošević, Pi AB Haase, Martina Stella, Adam Glos, arXiv:2603.23444 (2026), DOI: 10.48550/arXiv.2603.23444

Abstract: We introduce the Adaptive Derivative-Assembled Pseudo-Trotter ansatz Variational Majorana Propagation Eigensolver (ADAPT-VMPE), a quantum-inspired classical algorithm that exploits Majorana Propagation (MP) to produce circuits for approximating the ground state of molecular Hamiltonians. Equipped with the theoretical guarantees of MP, which provide controllable bounds on the approximation error, ADAPT-VMPE offers an efficient and scalable approach for iterative ansatz construction. A theoretical analysis of the computational complexity demonstrates that it is polynomial in both the number of qubits and the number of iterations. We present an in-depth analysis of circuit construction strategies, analyzing their impact on convergence and providing practical guidance for efficient ansatz generation. Using ADAPT-VMPE, we construct up to 100-qubit ansätze for a strongly correlated photosensitizer currently undergoing human clinical trials for cancer treatment. Our results demonstrate that constant overlap with the ground state across system sizes can be reached in polynomial time with polynomially sized circuits.

2. Simulation of Fermionic circuits using Majorana Propagation

Aaron Miller, Joachim Favre, Zoë Holmes, Özlem Salehi, **Rahul Chakraborty**, Anton Nykänen, Zoltán Zimborás, Adam Glos, Guillermo García-Pérez, arXiv:2503.18939 (2025), DOI:10.48550/arXiv.2503.18939

Abstract: We introduce Majorana Propagation, an algorithmic frame-

work for the classical simulation of Fermionic circuits. Inspired by Pauli Propagation, Majorana Propagation operates by applying successive truncations throughout the Heisenberg evolution of the observable. We identify monomial length as an effective truncation strategy for typical, unstructured circuits by proving that high-length Majorana monomials are exponentially unlikely to contribute to expectation values, and the backflow of high-length monomials to lower-length monomials is quadratically suppressed. We provide performance guarantees by proving analytically that approximation errors decrease exponentially with the truncation threshold and that only polynomial resources are required to compute the expectation value of observables up to a fixed error for an ensemble of circuits relevant to quantum chemistry. Majorana Propagation can be used either independently or in conjunction with quantum hardware to simulate Fermionic systems relevant to quantum chemistry and condensed matter. We exemplify this by using Majorana Propagation to find circuits that approximate ground states for strongly correlated systems of up to 52 Fermionic modes. Our results indicate that Majorana Propagation is orders of magnitude faster and more accurate than state-of-the-art tensor-network-based circuit simulators.

3. PyBEST: Improved functionality and enhanced performance, Katharina Boguslawski, Filip Brzeźek, **Rahul Chakraborty**, Kacper Cieślak, Seyedehdelaram Jahani, Aleksandra Leszczyk, Artur Nowak, Emil Sujkowski, Julian Świerczyński, Somayeh Ahmadvani, Dariusz Kędziera, Maximilian H Kriebel, Piotr Szymon Żuchowski, Paweł Tecmer, *Comput. Phys. Commun.* 297 (2024) 109049, DOI: 10.1016/j.cpc.2023.109049

Abstract: We present a methodological update to the Pythonic Black-box Electronic Structure Tool (PyBEST), a fully fledged modern electronic structure software package. Among other things, these include the extensions to pair Coupled Cluster Doubles (pCCD) open-shell electronic structures via the ionization potential and double ionization potential equation of motion formalism (IP-EOM-pCCD and DIP-EOM-pCCD), tailored Coupled Cluster methods, EOM extensions on top of the linearized Coupled Cluster correction on top of pCCD, the

configuration interaction correction on top of pCCD, and the static pCCD-in-DFT embedding. Furthermore, we have increased the performance of PYBEST by optimizing its underlying tensor contraction engine and providing GPU support via the CuPy library.

B. Published before starting the doctoral studies

1. Evolutionary algorithm based configuration interaction approach
Rahul Chakraborty, Paulami Ghosh, Debashree Ghosh, Int. J. Quantum Chem. 2018;118:e25509. DOI: 10.1002/qua.25509

Abstract: A new evolutionary algorithm for stochastic configuration interaction (CI) method designed as an affordable approximation to full configuration interaction (FCI) has been described here. The key components of the algorithm are initiation, propagation, and termination steps taking inspiration from the genetic algorithm. The propagation step is performed with cloning (retention of a Slater determinant without change), mutation (single excitation/de-excitation), and crossover (exchange of α and β strings between two Slater determinants) and termination is selection of few Slater determinants based on certain fitness function (measure of importance of a determinant in the CI space) and rejection of the rest. We find that the absolute value of the CI coefficients is a suitable fitness function when combined with a fixed selection scheme. We have tested its accuracy in 1D Hubbard problem and ground state potential energy surface (PES) has also been constructed for symmetric bond breaking of water molecule, where the errors are found to be around 10 mEh with low non-parallelity error, when retaining only a small fraction of the total number of Slater determinants in the final population. This shows that this method has the ability to capture both static and dynamic correlation. Performance and convergence properties of the algorithm are also tested for N₂ triple bond breaking problem. The algorithm opens up a promising way for stochastic sampling of the important determinants in the full Hilbert space.

2. Effect of solvation on the ionization of guanine nucleotide: A hybrid QM/EFP study

Rahul Chakraborty, Samik Bose, Debashree Ghosh, J. Comput. Chem. 2017, 38, 2528–2537. DOI: 10.1002/jcc.24913

Abstract: Ionization of nucleobases is affected by their biological environment, which includes both the effect of adjacent nucleotides as well as the presence of water around it. Guanine and its nucleotide have the lowest ionization potentials among the various DNA bases. Therefore, the threshold of ionization is dependent on that of guanine and its characterization is crucial to the prediction of interaction of light with DNA. We investigate the effect of solvation on the vertical ionization energies (VIEs) of guanine and its nucleotide. In this work, we have used hybrid quantum mechanics/molecular mechanics (QM/MM) approach with effective fragment potential as the MM method of choice and equation-of-motion coupled-cluster for ionization potential with singles and doubles (EOM-IP-CCSD) as the QM method. The performance of the hybrid scheme with respect to the full QM method shows an accuracy of ≤ 0.02 – 0.04 eV. The lowest few ionizations of the nucleotide are found to be from different parts of the moiety, that is, the nucleic acid base, phosphate, or sugar, and these ionization energies are very closely spaced giving rise to a very complicated spectrum. Furthermore, microsolvation has large effects on these ionizations and can lead to red or blue shift depending on the position of the water molecule. Even a single water molecule can change the order of ionized states in the nucleotide. The VIEs of the bulk solvated chromophores are predicted and compared to existing experimental spectra. The predominant role of polarization in the solvatochromic shift is noticed.

3. The effect of sequence on the ionization of guanine in DNA

Rahul Chakraborty and Debashree Ghosh, Phys. Chem. Chem. Phys., 2016, 18, 6526–6533, DOI: 10.1039/C5CP07804K

Abstract: The accurate estimation of the ionization energies and understanding the nature of the ionized states of the nucleic acid bases (NABs) are crucial to the understanding of the DNA damage mechanism. The vertical ionization energy (VIE) of guanine is the lowest among the NABs and the ionization energies are strongly affected by the environment, such as solvation and characteristics of nearby NABs. Therefore, we investigate the sequence dependence of the VIEs

of guanine in B-DNA. We use the equation of motion coupled cluster method for the estimation of ionization potential with single and double excitations (EOM-IP-CCSD) and density functional theory with dispersion corrected ω B97x-D for the estimation of VIEs. A significant amount of non-additivity or cooperativity, directly proportional to charge delocalization, is noticed in the change in VIE due to the interaction with the nearby NABs. While the change in VIE due to base pairing originates predominantly from charge–dipole interactions, stacking between base pairs is a more complicated balance of dispersion and charge–dipole interactions as well as stabilization due to the delocalization of the positive charge. The long range interactions are however dominated by $\frac{1}{r^3}$ distance dependence which shows the major role played by charge–dipole interactions. The extent of localization of positive holes on guanine is also estimated for various sequences.

2.3 Author contribution statement

Table 2.1: Author contributions for the article **Static embedding with pair coupled cluster doubles based methods**, *Phys. Chem. Chem. Phys.*, 2023, 25, 25377 (DOI: 10.1039/d3cp02502k)

Module	Work Contribution	Co-participant(s)
Conceptualization	50%	Paweł Tecmer
Literature search	75%	Paweł Tecmer
Implementation of static embedding module in PyBEST	50%	Paweł Tecmer, Katharina Boguslawski
DF-in-DFT calculations	75%	Paweł Tecmer
pCCD-in-DFT calculations	100%	–
Result analysis	50%	Paweł Tecmer, Katharina Boguslawski
Orbital entanglement analysis	50%	Katharina Boguslawski
Article writing, plots	60%	Paweł Tecmer, Katharina Boguslawski

Note: The table summarizes author contributions for the article by Rahul Chakraborty, Katharina Boguslawski, and Paweł Tecmer. The second column describes the contributions of the author of this thesis. **Corresponding authors:** Paweł Tecmer.

Table 2.2: Author contributions for the article **Toward Reliable Dipole Moments without Single Excitations: The Role of Orbital Rotations and Dynamical Correlation**, *J. Chem. Theory Comput.* 2024, 20, 4689–4702, DOI: 10.1021/acs.jctc.4c00471

Module	Work Contribution	Co-participant(s)
Conceptualization	25%	Paweł Tecmer, Matheus Morato F. de Moraes, Katharina Boguslawski
Literature search	50%	Matheus Morato F. de Moraes, Paweł Tecmer
Multipole moment module in PyBEST	50%	Julian Świerczyński, Artur Nowak, katharina Boguslawski
Data generation	100%	–
Result analysis	50%	Matheus Morato F. de Moraes, Paweł Tecmer, Katharina Boguslawski
Article writing, plots	50%	Matheus Morato F. de Moraes, Paweł Tecmer, Katharina Boguslawski

Note: The table summarizes author contributions for the article by Rahul Chakraborty, Matheus Morato F. de Moraes, Katharina Boguslawski, Artur Nowak, Julian Świerczyński, and Paweł Tecmer. The second column describes the contribution of the author of this thesis. **Corresponding authors** – Matheus Morato F. de Moraes, Katharina Boguslawski, Paweł Tecmer.

Table 2.3: Author contributions for the article **Expectation Value-pCCD-Based Methods for Single-Electron Properties**, *J. Phys. Chem. A*, 2025, 129, 6713–6732, DOI: 10.1021/acs.jpca.5c03859.

Module	Work Contribution	Co-participant(s)
Conceptualization	50%	Paweł Tecmer, Somayeh Ahmadkhani
Literature search	75%	Paweł Tecmer
Mathematical verification	50%	Somayeh Ahmadkhani
Code implementation in PyBEST	50%	Somayeh Ahmadkhani, Julian Świerczyński, Katharina Boguslawski
Data generation	100%	–
Result analysis	75%	Paweł Tecmer
Article writing, plots	75%	Paweł Tecmer, Somayeh Ahmadkhani

Note: The table summarizes author contributions for the article by Rahul Chakraborty, Somayeh Ahmadkhani, Julian Świerczyński, Katharina Boguslawski, and Paweł Tecmer. The second column describes the contribution of the author of this thesis. **Corresponding authors** – Somayeh Ahmadkhani, Paweł Tecmer.

Table 2.4: Author contributions for the article **Frozen density embedding with pCCD electron densities**, arXiv:2604.14904, 2026, DOI: 10.48550/arXiv.2604.14904.

Module	Work Contribution	Co-participant(s)
Conceptualization	75%	Paweł Tecmer
Literature search	100%	–
Code implementation in PyBEST	100%	–
Data generation	100%	–
Result analysis	75%	Paweł Tecmer
Article writing, plots	75%	Paweł Tecmer

Note: The table summarizes author contributions for the article by Rahul Chakraborty, and Paweł Tecmer. The second column describes the contribution of the author of this thesis. **Corresponding author** – Rahul Chakraborty.

Chapter 3

Conclusions

3.1 Summary of key findings

Even in this era of high-performance computing, studying large chemical structures with quantum chemical methods still remains a daunting task, both at the theoretical and implementation levels. While this certainly limits the full utilization of quantum chemistry, it also provides opportunities for fundamental theoretical method development. pCCD, a method that gained traction in the last decade, is one example of such development. The discussions so far have shown the computational efficiency and reliable performance of pCCD, especially in its optimized natural orbital basis, in simulating the energetics of strongly correlated systems, where many standard electronic structure methods struggle. Still, achieving chemical accuracy in the energies, a reliable description of excited states, and finer molecular properties requires post-pCCD dynamic correlation corrections, which entail a massive jump in computational expense.

In this thesis, two different strategies have been explored to tackle this problem: a) embedded pCCD, where we use a *divide-and-conquer* approach by bifurcating the system and reserving accurate but costlier methods for a subsystem of our interest, b) efficient expectation-value based single-electron property estimation, where we avoid solving the expensive response $\hat{\Lambda}$ -equations associated with the post-pCCD methods.

In the first work discussed in the thesis [166], a simple WFT-in-DFT method was implemented. A static embedding potential was obtained from a separate DFT-in-DFT FDE calculation, and this was used to approximate the tetrahalide environment in uranyl halides, where the uranyl moiety itself was investigated with pCCD-based methods. The EOM-fpLCCSD-in-DFT calculations produced the correct order of

excited states, with excitation energies in close agreement with the supramolecular EOM-fpLCCSD values, as shown for $\text{UO}_2\text{F}_4^{2-}$ in Fig. 3.1.

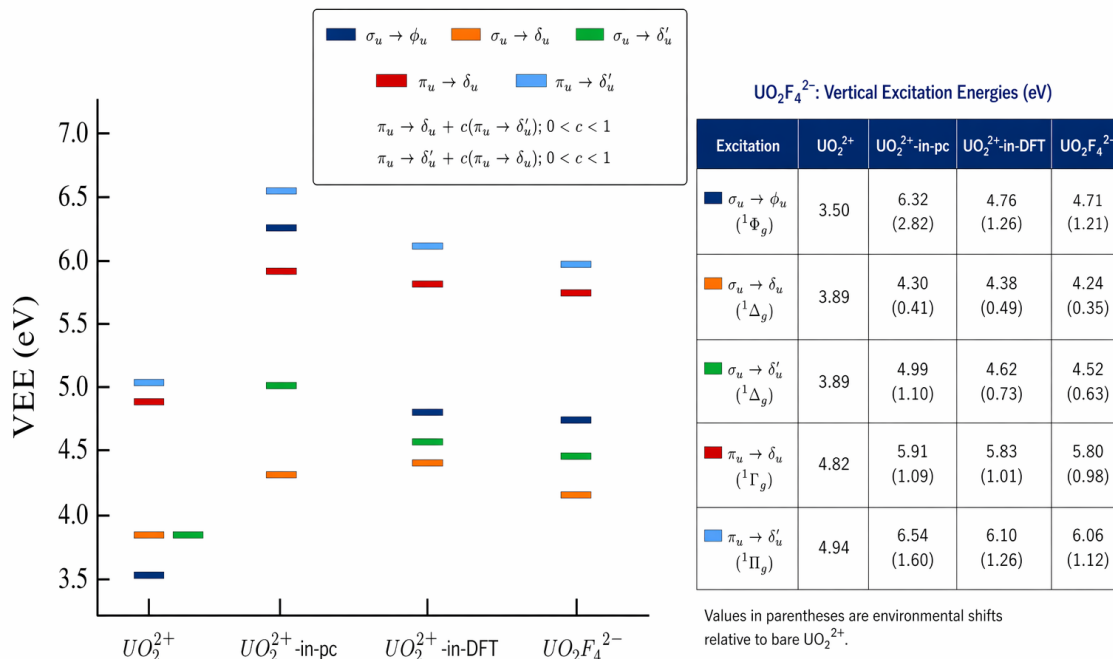


Figure 3.1: Spin-free vertical excitation energies of $\text{UO}_2\text{F}_4^{2-}$ obtained with EOM-fpLCCSD. UO_2^{2+} -in-pc and UO_2^{2+} -in-DFT indicate that the F are replaced by point-charges and embedding potential obtained from DFT-in-DFT, respectively. Schematic diagram on the left side for visual representation.

However, as we move to heavier halides from F to Br, the agreement progressively deteriorates. This can be due to the usage of the embedding potential obtained at the ground state for the modeling of excited states, which may not properly describe the response of excited states of UO_2^{2+} moiety due to the halides, especially for Cl and Br. For heavier halides (Cl and Br), we also observe contributions from ligand-to-metal charge transfer (LMCT) in the $\pi_u \rightarrow \delta_u$ and $\pi_u \rightarrow \delta'_u$ transitions, which can not be described in an embedding model. A further approximated model, where point charges replace the halide atoms, was also investigated in this work, and as expected, it could not reproduce the correct order and reliable excitation energies for all three tetrahalides. It is to be noted that our investigation is limited by the lack of spin-orbit coupling in the implemented method, which prevents an explicit comparison with experimental excitation energies for these systems.

The single-orbital entropies calculated with fpLCCSD indicate that the embedding models reproduce the correlations of the occupied σ_u and π_u orbitals reasonably well, but show significant deviations from the supramolecular reference for the ϕ_u , δ_u , and δ'_u orbitals. The supramolecular calculations reveal enhanced entanglement of the

ϕ_u and δ'_u orbitals, which is not captured by any of the embedding approaches. The orbital-apir mutual information values provide insight into how well the embedding schemes reproduce orbital correlations and allow for a quantitative assessment of environmental effects. Among the studied systems, the agreement is best for $\text{UO}_2\text{Cl}_4^{2-}$, followed by $\text{UO}_2\text{F}_4^{2-}$, while the largest deviations are observed for $\text{UO}_2\text{Br}_4^{2-}$. This suggests that the good performance of EOM-fpLCCSD-in-DFT for the excited states of $\text{UO}_2\text{F}_4^{2-}$ may partly result from a cancellation of errors. Overall, the analysis suggests that orbital entanglement may not always provide a consistent or predictive measure of the quality of the embedding potential.

A crucial aspect for any quantum chemistry method is performance in terms of the estimation of molecular properties. Though traditional CC variants such as CCSD and CCSD(T) are regularly employed to obtain these properties, no such analysis was done previously with pCCD and post-pCCD methods. To that end, we selected ground-state multipole moments as our assessment tool, specifically due to their sensitivity to the quality of density matrices obtained with any particular method. A test set of 20 main group diatomics with varying types of bonding was investigated. We compared pCCD, fpLCCD, and fpLCCSD in two different MO bases, namely canonical HF orbitals and the pCCD-optimized natural orbitals [167]. Based on our numerical results, we arrived at the following conclusions- 1. The oo-fpLCCD method provides the most balanced performance for the majority of the molecules. In particular, oo-fpLCCD achieves CCSD-level accuracy for dipole moments in singly bonded diatomics, while also reproducing dipole moment surfaces comparable to those obtained from multireference methods. 2. In contrast, the linearized singles in the LCC correction on top of the oo-pCCD reference deteriorate the dipole moments, especially for multiply bonded diatomic molecules. The inclusion of singles, though, improves the description of diatomics with van der Waals interaction, such as LiNa, as they play a crucial role in capturing dispersion interactions [168]. Based on the comparison of dipole moments obtained with pCCD, oo-pCCD, and fpLCCSD, the variation among these methods can be used to assess the importance of orbital relaxation and higher-order excitations. When all three methods yield similar values, the dependence on these effects is small, and both fpLCCSD and oo-fpLCCD provide results close to the CCSD(T) reference with average errors of about 4%. If oo-pCCD and fpLCCSD agree, orbital relaxation plays a significant role, making oo-fpLCCSD the preferred choice. In contrast, large discrepancies between oo-pCCD and fpLCCSD indicate the need for higher excitation levels, leading to significantly larger errors. In these cases, excluding carbon-containing molecules, both oo-fpLCCD and oo-fpLCCSD show relative average errors of approximately 30%. Upon inclusion of carbon-containing systems,

the error for oo-fpLCCD decreases to 21%, while that for oo-fpLCCSD increases to as high as 43%. Thus, although orbital optimization improves the overall statistical performance of these methods, a more detailed analysis reveals a strong dependence on the chemical nature of the system under investigation.

The subpar performance of the fpLCCSD model necessitated finding alternate routes of calculating multipole moments. One way to improve the post-pCCD dynamic correlation correction is to retain the nonlinear cluster operator terms in the Hamiltonian, which are excluded in the case of fpLCC, resulting in the fpCC type methods. Already, work of Leszczyk et al. [63] showed that fpCC could solve the numerical divergences shown by the fpLCC in potential energy surfaces, even for multiple bond-breaking cases such as CO. However, solving the response $\hat{\Lambda}$ -equations for fpCC is going to be computationally expensive. An alternate route can be found in the works of Korona et al. [107], building upon the $\hat{S}$ -operator based expectation-value CC formalism by Jeziorski and Moszyński [102]. This method was implemented by us for pCCD, fpCC, and fpLCC. The simple approximations on $\hat{S}$ -operator expansion allow us to construct the density matrix directly with the cluster amplitudes. We did an extensive benchmarking of the performance of this approach, in terms of dipole and quadrupole moments of some small diatomic and polyatomic molecules and ions, as well as some common organic molecules [165]. Again, we used both canonical HF and pCCD natural orbitals for comparison. Our investigation clearly showed the superior performance of XfpCCSD, giving almost identical results to XCCSD in the canonical orbital basis (X denoting ‘expectation value’, to differentiate from the response approach). Use of the natural orbital basis improves the performance of oo-XfpLCCD and oo-XfpLCCSD compared to the canonical basis. Interestingly, for oo-XfpCCSD, we get slightly worse numbers compared to XfpCCSD in the canonical basis. XfpCCD, on the other hand, shows improvement when using the pCCD natural orbital basis, though oo-XfpCCD is still worse than oo-XfpCCSD. This indicates potential overestimation of the relaxation effect in the case of fpCCSD, brought about by both orbital optimization as well as inclusion of single excitations. Still, the overall performance of XfpCCSD is very encouraging for its wider usage. A particular advantage of the expectation value approach, as already pointed out, is the computational savings. A direct comparison is possible for fpLCC, for which the response framework is available in the PyBEST software package. For example, for nitrobenzene with 261 contracted basis functions, 23 active occupied orbitals, and 229 active virtual orbitals—the response $\hat{\Lambda}$ -equations required approximately 103 minutes for fpLCCD and 204 minutes for fpLCCSD (using 8 CPUs on an Intel(R) Xeon(R) Gold 6336Y processor at 2.40 GHz). These numbers are seemingly expected to grow for the corresponding response fpCC calculations, as we retain

many more terms in the fpCC Hamiltonian as shown in Eq. (1.49).

The works discussed so far put us on a strong footing to further extend pCCD and post-pCCD-based molecular property studies to larger systems. Recently, we implemented a pCCD-in-pCCD embedding approach [169]. Here, the effect of other subsystems on the active subsystem of our interest is expressed through an embedding potential obtained using the electron densities obtained with the pCCD density matrix. Obtaining the pCCD density matrix is much cheaper compared to CCSD or other CC methods due to the pair-structure of the excitation and de-excitation operators in pCCD $\hat{A}$ -equations. Also, in cases where the natural orbital basis is used, the pCCD $\hat{A}$ -equations are already solved during the orbital optimization stage. Hence, the oo-pCCD density matrix does not require any extra computational step. The pCCD-in-pCCD approach removes any dependence on DFT for us, providing a single, easily usable framework compared to our previous DFT-in-DFT+pCCD type workflow. The pCCD-in-pCCD framework is depicted in Fig. 3.2 (figure recreated from [169]). After convergence of the pCCD-in-pCCD freeze-and-thaw cycles, we apply the post-pCCD ground and excited state methods only on the subsystem of our interest. Alternatively, we also tested a non-freeze-and-thaw procedure, where the embedding potentials are generated using isolated subsystems at the supramolecular geometry. We applied these to different classes of complexes, from estimation of dipole moments of weakly interacting $\text{CO}_2 \cdots \text{Rg}$ [Rg=He, Ne, Ar, Kr] to local excitations of solutes in comparatively strongly overlapping microsolvation. All calculations were done in the canonical HF orbital basis.

The dipole moments of $\text{CO}_2 \cdots \text{Rg}$ obtained with the expectation-value approach and from the embedding calculations provide excellent agreement with the CCSD(T) reference. This is especially noteworthy as the dipole moments of these complexes are very small in magnitude. In this case, XfpCCSD and XfpLCCSD perform almost similarly. As we go from He to Kr, the dipole moments at any level of theory improve. Other than $\text{CO}_2 \cdots \text{He}$, the freeze-and-thaw cycles worsen the dipole moment to a small extent, especially for $\text{CO}_2 \cdots \text{Ne}$.

The vertical excitation energies of the microsolvated uracil ($\text{C}_4\text{H}_4\text{N}_2\text{O}_2 \cdot (\text{H}_2\text{O})_6$) calculated with the EOM-fpLCCSD-in-pCCD approach, where the solvation sphere consisting of six H_2O molecules is considered the environment, showed good agreement with the supramolecular EOM-fpLCCSD values. Especially the values obtained with the freeze-and-thaw calculation were remarkably accurate. However, another smaller microsolvated system we studied, i.e, $\text{H}_2\text{O} \cdots \text{NH}_3$ showed an opposite trend, where the VEEs become slightly worse when the freeze-and-thaw cycles were used. Interestingly, the same trend was observed in previous CC-in-CC work also [146]. This indicates possible overpolarization of the subsystems during the

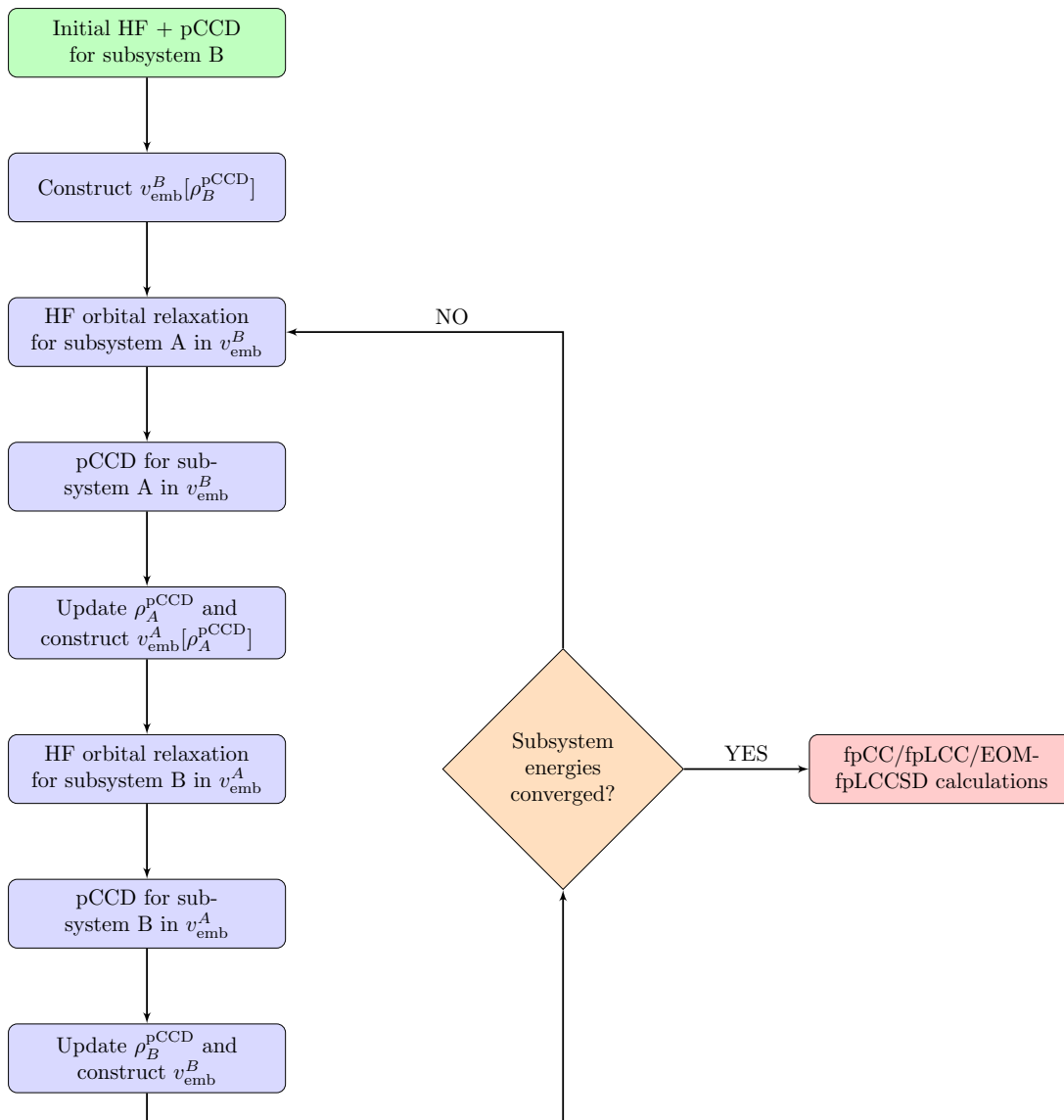


Figure 3.2: Pictorial description of FDE freeze-and-thaw cycles with pCCD densities. The green block represents the initial setup within the isolated subsystem environment. The red block signifies the final stage, where the converged pCCD wave function from the active subsystem is used to perform post-pCCD calculations, including fpCC, fpLCC, and subsequent EOM-fpLCCSD.

freeze-and-thaw steps.

Overall, these results demonstrate that the proposed embedding framework can reliably describe a wide range of intermolecular interactions. Its strong performance on small systems motivates its extension to more challenging cases, particularly actinide complexes embedded in large ligands, as pCCD has proven to be reliable for both ground- and excited-states for actinide moieties [32, 85]. The pCCD-in-pCCD scheme is also expected to be effective for systems in which DFT fails to provide an accurate ground-state description [136, 170].

3.2 Outlook and future Work

The discussion so far has established the idea that the pCCD-based embedding approaches provide a viable pathway for reliably modeling ground and excited states of large molecules, within the constraints of such fragmentation methods. The expectation-value approach for pCCD and the post-pCCD family of methods, on the other hand, provides computationally inexpensive access to some important molecular properties. At this stage, a critical evaluation of the works discussed is needed to gain insights into our limitations and the prospects going forward.

While the expectation value approach is suitable for static properties such as multipole moments, it does not directly provide quantities that depend on wave function response. Properties such as energy gradients, polarizabilities, excitation energies, and transition moments, therefore, require a response-based formulation. For this, the development of a response approach for fpCC represents a natural direction for future development.

A limitation of our implemented pCCD-in-DFT and pCCD-in-pCCD methods is that they do not consider the polarization of the environment due to the excitation in the active subsystem. This is due to our usage of the same static embedding potential obtained from ground state calculations for the excited state. Potentially, this can be improved upon using state-specific embedding potentials [171]. Recently, Kossoski et al. implemented a state-specific pCCD orbital optimization method [79]. Their Δ_{oo} -pCCD method provided excitation energies of doubly excited states in close agreement with EOM-CCSDT. This can potentially be used for our pCCD-in-pCCD embedding; however, orbital optimization with pCCD for multiple states will most likely nullify the computational advantage of the embedding approach, especially because excited-state orbital optimizations are frequently troubled by the presence of multiple local minima and saddle points.

An alternate and more viable option will be to use the framework suggested by Höfener et al. using the linear response approach [144] and implemented by them for the approximate second-order singles and doubles CC2 method. The recent implementation of LR-pCCD+S by Ahmadkhani et al. [101] provides us with the tools to implement this for the pCCD-in-pCCD embedding scheme. The inter-subsystem coupling can then be efficiently established using the transition density matrices of LR-pCCD+S.

To overcome the limitations associated with approximate kinetic energy functionals, Manby, Miller, and co-workers proposed an embedding strategy in which the Fock operator of the active subsystem is augmented by a projection operator [172, 173].

This leads to the modified eigenvalue problem

$$(\mathbf{F} + \mu\mathbf{P})\mathbf{C}_A = \mathbf{S}\mathbf{C}_A\epsilon_A, \quad (3.1)$$

where $\mathbf{F}$ is the Fock matrix of the embedded system, $\mathbf{P}$ is a projection operator that penalizes overlap with subsystem B , and μ is a large positive parameter. Here, subsystem A denotes the active region, while subsystem B represents the environment. For sufficiently large μ , the resulting orbitals become orthogonal to those of subsystem B . In practice, however, the results may depend on the choice of μ , which must be selected carefully.

An improved formulation was later introduced by Kállay and co-workers [174, 175] through the use of an exact embedding potential based on the Huzinaga equation [176]. In this approach, the following matrix equation is solved:

$$(\mathbf{F}^{AA} - \mathbf{S}^{AB}(\mathbf{S}^{BB})^{-1}\mathbf{F}^{BA})\mathbf{C}_A = \mathbf{S}^{AA}\mathbf{C}_A\epsilon_A, \quad (3.2)$$

where $\mathbf{F}^{AB}$ and $\mathbf{S}^{AB}$ denote the Fock and overlap matrix blocks between subsystems A and B . This formulation enforces exact orthogonality between subsystem A and subsystem B orbitals without introducing any empirical parameters.

As a potential extension, the projection-based embedding scheme of Kállay and co-workers can be combined with oo-pCCD by replacing the effective Fock matrix with a Huzinaga-type projected operator,

$$\tilde{\mathbf{F}}^{\text{pCCD}} = \mathbf{F}^{\text{pCCD}} - \mathbf{S}^{AB}(\mathbf{S}^{BB})^{-1}\mathbf{F}^{BA,\text{pCCD}} \quad (3.3)$$

The orbital optimization can then be carried out using the modified gradient

$$g_{pq} = \tilde{F}_{pq}^{\text{pCCD}} - \tilde{F}_{qp}^{\text{pCCD}} \quad (3.4)$$

yielding a parameter-free pCCD-in-pCCD embedding framework.

Orbital entanglement measures derived from the pCCD wavefunction can offer a physically motivated criterion for subsystem partitioning in the projection-based embedding frameworks. Specifically, single-orbital entropies and orbital-pair mutual information obtained from pCCD density matrices enable the identification of strongly correlated orbitals, thereby defining an appropriate active subsystem. In contrast to the Singular Value Decomposition (SVD) based partitioning scheme introduced by Claudino et al. [177], which primarily reflects spatial partitioning,

entanglement-based approaches explicitly account for electronic correlation and are thus better suited to capture nonlocal effects such as charge transfer or delocalized π -systems. Consequently, orbitals from different fragments may be included within a common active subsystem when they exhibit significant mutual correlation. Although this relaxes the strict locality inherent to fragment-based embedding, it remains theoretically consistent within projection-based formulations that operate in orbital space. Notably, orbital entanglement bears a conceptual resemblance to the method of increments, [178] as both aim to identify dominant correlation effects; however, entanglement measures provide this information directly from a single correlated wave function, avoiding the combinatorial scaling associated with incremental expansions.

Finally, we are also working on the development and implementation of a QM/MM-type embedding-based pCCD approach. QM/MM models are widely employed for the description of solvent effects in biomolecular systems, providing an efficient multiscale framework in which a quantum mechanical region is embedded in a classical environment [112, 114]. However, both classical molecular mechanics force fields and continuum solvation models suffer from intrinsic limitations: the former rely on fixed, empirically parameterized point-charge representations that often fail to capture many-body polarization and charge-transfer effects, while the latter provide only a mean-field description of the solvent and cannot account for the discrete molecular structure of the environment.

To overcome these shortcomings, the Effective Fragment Potential (EFP) method has emerged as a highly accurate explicit solvation model that incorporates individual solvent molecules within an ab initio-based framework [179, 180]. Contrary to the conventional classical force fields, EFP provides a substantially more rigorous description of intermolecular interactions by representing electrostatics through distributed multipoles up to the octupole level, rather than fixed point charges, and by including a self-consistent treatment of polarization. This enables a physically sound description of many-body induction effects that are typically only implicitly or approximately captured in standard MM models. Furthermore, its rigorous energy decomposition scheme offers detailed insight into the long-range components of solute-solvent interactions [181]. These features motivate the coupling of pCCD with EFP as a robust framework for modeling actinide chemistry in solution, achieving a balanced description of electronic correlation and explicit solvation at moderate computational cost.

To summarize, the developments discussed in this thesis point toward a unified embedding framework in which pCCD-based correlation methods are combined with advanced subsystem partitioning and environment descriptions. The combination of

projection-based embedding, entanglement-driven orbital space selection, and linear-response formulations provides a promising route toward describing both ground- and excited-state properties in large molecular systems. In parallel, the extension to QM/MM-type pCCD/EFP embedding further broadens the applicability of the approach to chemically relevant condensed-phase environments, offering a balanced treatment of electronic correlation and explicit solvation. These directions collectively highlight the potential of pCCD-based embedding strategies as a versatile and systematically improvable framework for future studies of complex electronic structure problems.

Bibliography

- [1] K. Raghavachari, “Electron correlation techniques in quantum chemistry: Recent advances,” *Annu. Rev. Phys. Chem.*, vol. 42, no. 1, pp. 615–642, 1991.
- [2] K. Raghavachari and J. B. Anderson, “Electron correlation effects in molecules,” *J. Phys. Chem.*, vol. 100, no. 31, pp. 12960–12973, 1996.
- [3] P. E. M. Siegbahn, “The direct CI method,” in *Methods in Computational Molecular Physics* (G. H. F. Diercksen and S. Wilson, eds.), pp. 189–207, Dordrecht: D. Reidel, 1983.
- [4] C. D. Sherrill, “Frontiers in electronic structure theory,” *J. Chem. Phys.*, vol. 132, no. 11, p. 110902, 2010.
- [5] C. D. Sherrill, “An introduction to configuration interaction theory,” *School of Chemistry and Biochemistry, Georgia Institute of Technology*, 1995.
- [6] J. Čížek and J. Paldus, “Coupled cluster approach,” *Phys. Scr.*, vol. 21, no. 3-4, pp. 251–254, 1980.
- [7] T. D. Crawford and H. F. Schaefer III, “An introduction to coupled cluster theory for computational chemists,” *Reviews in computational chemistry*, vol. 14, pp. 33–136, 2007.
- [8] R. J. Bartlett and M. Musiał, “Coupled-cluster theory in quantum chemistry,” *Rev. Mod. Phys.*, vol. 79, no. 1, pp. 291–352, 2007.
- [9] I. Talmi, “Generalized seniority and structure of semi-magic nuclei,” *Nucl. Phys. A*, vol. 172, no. 1, pp. 1–24, 1971.
- [10] S. Shlomo and I. Talmi, “Shell-model hamiltonians with generalized seniority eigenstates,” *Nucl. Phys. A*, vol. 198, no. 1, pp. 81–108, 1972.
- [11] M. Caprio, F. Luo, K. Cai, C. Constantinou, and V. Hellemans, “Generalized seniority with realistic interactions in open-shell nuclei,” *J. Phys. G: Nucl. Part. Phys.*, vol. 39, no. 10, p. 105108, 2012.
- [12] L. Bytautas, T. M. Henderson, C. A. Jiménez-Hoyos, J. K. Ellis, and G. E.

- Scuseria, "Seniority and orbital symmetry as tools for establishing a full configuration interaction hierarchy," *J. Chem. Phys.*, vol. 135, no. 4, p. 044119, 2011.
- [13] L. Bytautas, G. E. Scuseria, and K. Ruedenberg, "Seniority number description of potential energy surfaces: Symmetric dissociation of water, N_2 , C_2 , and Be_2 ," *J. Chem. Phys.*, vol. 143, no. 9, p. 094105, 2015.
- [14] P. R. Surján, "An introduction to the theory of geminals," in *Correlation and localization*, pp. 63–88, Springer, 1999.
- [15] P. R. Surján, *Correlation and localization*, vol. 203. Springer, 2003.
- [16] P. Tecmer and K. Boguslawski, "Geminal-based electronic structure methods in quantum chemistry. Toward a geminal model chemistry," *Phys. Chem. Chem. Phys.*, vol. 24, no. 38, pp. 23026–23048, 2022.
- [17] W. Kutzelnigg, "Direct Determination of Natural Orbitals and Natural Expansion Coefficients of Many-Electron Wavefunctions. I. Natural Orbitals in the Geminal Product Approximation," *J. Chem. Phys.*, vol. 40, no. 12, pp. 3640–3647, 1964.
- [18] D. M. Silver, "Natural orbital expansion of interacting geminals," *J. Chem. Phys.*, vol. 50, no. 12, pp. 5108–5116, 1969.
- [19] D. M. Silver, E. L. Mehler, and K. Ruedenberg, "Electron correlation and separated pair approximation in diatomic molecules. I. Theory," *J. Chem. Phys.*, vol. 52, no. 3, pp. 1174–1180, 1970.
- [20] P. A. Limacher, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck, and P. Bultinck, "A new mean-field method suitable for strongly correlated electrons: Computationally facile antisymmetric products of nonorthogonal geminals," *J. Chem. Theory Comput.*, vol. 9, no. 3, pp. 1394–1401, 2013.
- [21] P. A. Limacher, T. D. Kim, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck, and P. Bultinck, "The influence of orbital rotation on the energy of closed-shell wavefunctions," *Mol. Phys.*, vol. 112, pp. 853–862, Mar. 2014.
- [22] P. R. Surján, Á. Szabados, P. Jeszenszki, and T. Zoboki, "Strongly orthogonal geminals: size-extensive and variational reference states," *J. Math. Chem.*, vol. 50, no. 3, pp. 534–551, 2012.
- [23] T. M. Henderson, I. W. Bulik, T. Stein, and G. E. Scuseria, "Seniority-based coupled cluster theory," *J. Chem. Phys.*, vol. 141, no. 24, p. 244104, 2014.
- [24] K. Boguslawski, P. Tecmer, P. W. Ayers, P. Bultinck, S. De Baerdemacker,

- and D. Van Neck, "Efficient description of strongly correlated electrons with mean-field cost," *Phys. Rev. B*, vol. 89, no. 20, p. 201106, 2014.
- [25] T. Stein, T. M. Henderson, and G. E. Scuseria, "Seniority zero pair coupled cluster doubles theory," *J. Chem. Phys.*, vol. 140, no. 21, p. 214113, 2014.
- [26] K. Boguslawski, P. Tecmer, P. Bultinck, S. De Baerdemacker, D. Van Neck, and P. W. Ayers, "Nonvariational orbital optimization techniques for the AP1roG wave function," *J. Chem. Theory Comput.*, vol. 10, no. 11, pp. 4873–4882, 2014.
- [27] K. Boguslawski, P. Tecmer, P. A. Limacher, P. A. Johnson, P. W. Ayers, P. Bultinck, S. De Baerdemacker, and D. Van Neck, "Projected seniority-two orbital optimization of the antisymmetric product of one-reference orbital geminal," *J. Chem. Phys.*, vol. 140, no. 21, p. 214114, 2014.
- [28] U. Bozkaya, J. M. Turney, Y. Yamaguchi, H. F. Schaefer, and C. D. Sherrill, "Quadratically convergent algorithm for orbital optimization in the orbital-optimized coupled-cluster doubles method and in orbital-optimized second-order Møller-Plesset perturbation theory," *J. Chem. Phys.*, vol. 135, no. 10, p. 104103, 2011.
- [29] U. Bozkaya and C. D. Sherrill, "Analytic energy gradients for the coupled-cluster singles and doubles method with the density-fitting approximation," *J. Chem. Phys.*, vol. 144, no. 17, p. 174103, 2016.
- [30] T. B. Pedersen, S. Lehtola, I. Fdez. Galván, and R. Lindh, "The versatility of the Cholesky decomposition in electronic structure theory," *WIREs Comput Mol Sci.*, vol. 14, no. 1, p. e1692, 2024.
- [31] P. Tecmer, K. Boguslawski, P. A. Johnson, P. A. Limacher, M. Chan, T. Verstraelen, and P. W. Ayers, "Assessing the accuracy of new geminal-based approaches," *J. Phys. Chem. A*, vol. 118, no. 39, pp. 9058–9068, 2014.
- [32] P. Tecmer, K. Boguslawski, and P. W. Ayers, "Singlet ground state actinide chemistry with geminals," *Phys. Chem. Chem. Phys.*, vol. 17, no. 22, pp. 14427–14436, 2015.
- [33] A. J. Garza, I. W. Bulik, T. M. Henderson, and G. E. Scuseria, "Synergy between pair coupled cluster doubles and pair density functional theory," *J. Chem. Phys.*, vol. 142, no. 4, p. 044109, 2015.
- [34] A. J. Garza, I. W. Bulik, T. M. Henderson, and G. E. Scuseria, "Range separated hybrids of pair coupled cluster doubles and density functionals," *Phys. Chem. Chem. Phys.*, vol. 17, no. 34, pp. 22412–22422, 2015.

- [35] A. Nowak and K. Boguslawski, “A configuration interaction correction on top of pair coupled cluster doubles,” *Phys. Chem. Chem. Phys.*, vol. 25, no. 10, pp. 7289–7301, 2023.
- [36] P. A. Limacher, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck, and P. Bultinck, “Simple and inexpensive perturbative correction schemes for antisymmetric products of nonorthogonal geminals,” *Phys. Chem. Chem. Phys.*, vol. 16, no. 11, pp. 5061–5065, 2014.
- [37] K. Boguslawski and P. Tecmer, “Benchmark of dynamic electron correlation models for seniority-zero wave functions and their application to thermochemistry,” *J. Chem. Theory Comput.*, vol. 13, no. 12, pp. 5966–5983, 2017.
- [38] J. Paldus, J. Čížek, and M. Takahashi, “Approximate account of the connected quadruply excited clusters in the coupled-pair many-electron theory,” *Phys. Rev. A*, vol. 30, no. 5, p. 2193, 1984.
- [39] J. Paldus and J. Planelles, “Valence bond corrected single reference coupled cluster approach: I. General formalism,” *Theor. Chem. Acc.*, vol. 89, no. 1, pp. 13–31, 1994.
- [40] J. Planelles, J. Paldus, and X. Li, “Valence bond corrected single reference coupled cluster approach: II. Application to PPP model systems,” *Theor. Chem. Acc.*, vol. 89, no. 1, pp. 33–57, 1994.
- [41] J. Planelles, J. Paldus, and X. Li, “Valence bond corrected single reference coupled cluster approach: III. Simple model of bond breaking or formation,” *Theor. Chem. Acc.*, vol. 89, no. 1, pp. 59–76, 1994.
- [42] J. Paldus, “Externally and internally corrected coupled cluster approaches: an overview,” *J. Math. Chem.*, vol. 55, no. 2, pp. 477–502, 2017.
- [43] L. Z. Stolarczyk, “Complete active space coupled-cluster method. extension of single-reference coupled-cluster method using the CASSCF wavefunction,” *Chem. Phys. Lett.*, vol. 217, no. 1-2, pp. 1–6, 1994.
- [44] G. Peris, J. Planelles, and J. Paldus, “Single-reference CCSD approach employing three-and four-body CASSCF corrections: A preliminary study of a simple model,” *Int. J. Quantum Chem.*, vol. 62, no. 2, pp. 137–151, 1997.
- [45] X. Li, G. Peris, J. Planelles, F. Rajadall, and J. Paldus, “Externally corrected singles and doubles coupled cluster methods for open-shell systems,” *J. Chem. Phys.*, vol. 107, no. 1, pp. 90–98, 1997.
- [46] X. Li and J. Paldus, “Reduced multireference CCSD method: An effective ap-

- proach to quasidegenerate states,” *J. Chem. Phys.*, vol. 107, no. 16, pp. 6257–6269, 1997.
- [47] X. Li and J. Paldus, “Truncated version of the reduced multireference coupled-cluster method with perturbation selection of higher than pair clusters,” *Int. J. Quantum Chem.*, vol. 80, no. 4-5, pp. 743–756, 2000.
- [48] X. Li and J. Paldus, “Full potential energy curve for N_2 by the reduced multireference coupled-cluster method,” *J. Chem. Phys.*, vol. 129, no. 5, p. 054104, 2008.
- [49] G. J. Aroeira, M. M. Davis, J. M. Turney, and H. F. Schaefer III, “Coupled cluster externally corrected by adaptive configuration interaction,” *J. Chem. Theory Comput.*, vol. 17, no. 1, pp. 182–190, 2020.
- [50] K. Jankowski and K. Kowalski, “Approximate coupled cluster methods based on a split-amplitude strategy,” *Chem. Phys. Lett.*, vol. 256, no. 1-2, pp. 141–148, 1996.
- [51] K. Jankowski, I. Grabowski, and K. Kowalski, “Approximate coupled-cluster methods employing split cluster amplitudes: Implementation of an almost-linear coupled-cluster formalism,” *J. Chem. Phys.*, vol. 109, no. 15, pp. 6255–6263, 1998.
- [52] N. Oliphant and L. Adamowicz, “The implementation of the multireference coupled-cluster method based on the single-reference formalism,” *J. Chem. Phys.*, vol. 96, no. 5, pp. 3739–3744, 1992.
- [53] P. Piecuch, N. Oliphant, and L. Adamowicz, “A state-selective multireference coupled-cluster theory employing the single-reference formalism,” *J. Chem. Phys.*, vol. 99, no. 3, pp. 1875–1900, 1993.
- [54] P. Piecuch and L. Adamowicz, “State-selective multireference coupled-cluster theory employing the single-reference formalism: Implementation and application to the H_8 model system,” *J. Chem. Phys.*, vol. 100, no. 8, pp. 5792–5809, 1994.
- [55] T. Kinoshita, O. Hino, and R. J. Bartlett, “Coupled-cluster method tailored by configuration interaction,” *J. Chem. Phys.*, vol. 123, no. 7, p. 074106, 2005.
- [56] O. Hino, T. Kinoshita, G. K. Chan, and R. J. Bartlett, “Tailored coupled cluster singles and doubles method applied to calculations on molecular structure and harmonic vibrational frequencies of ozone,” *J. Chem. Phys.*, vol. 124, no. 11, p. 114311, 2006.

- [57] A. Melnichuk and R. J. Bartlett, “Relaxed active space: Fixing tailored-CC with high order coupled cluster. I,” *J. Chem. Phys.*, vol. 137, no. 21, p. 214103, 2012.
- [58] A. Melnichuk and R. J. Bartlett, “Relaxed active space: Fixing tailored-CC with high order coupled cluster. II,” *J. Chem. Phys.*, vol. 140, no. 6, p. 064113, 2014.
- [59] R. Feldmann, M. Mörchen, J. Lang, M. Lesiuk, and M. Reiher, “Complete active space iterative coupled cluster theory,” *J. Phys. Chem. A*, vol. 128, no. 40, pp. 8615–8627, 2024.
- [60] L. Veis, A. Antalík, J. Brabec, F. Neese, Ö. Legeza, and J. Pittner, “Coupled cluster method with single and double excitations tailored by matrix product state wave functions,” *J. Phys. Chem. Lett.*, vol. 7, no. 20, pp. 4072–4078, 2016.
- [61] F. M. Faulstich, M. Máté, A. Laestadius, M. A. Csirik, L. Veis, A. Antalík, J. Brabec, R. Schneider, J. Pittner, S. Kvaal, *et al.*, “Numerical and theoretical aspects of the dmrg-tcc method exemplified by the nitrogen dimer,” *J. Chem. Theory Comput.*, vol. 15, no. 4, pp. 2206–2220, 2019.
- [62] J. Lang, A. Antalík, L. Veis, J. Brandejs, J. Brabec, Ö. Legeza, and J. Pittner, “Near-linear scaling in DMRG-based tailored coupled clusters: An implementation of DLPNO-TCCSD and DLPNO-TCCSD (T),” *J. Chem. Theory Comput.*, vol. 16, no. 5, pp. 3028–3040, 2020.
- [63] A. Leszczyk, M. Máté, Ö. Legeza, and K. Boguslawski, “Assessing the accuracy of tailored coupled cluster methods corrected by electronic wave functions of polynomial cost,” *J. Chem. Theory Comput.*, vol. 18, no. 1, pp. 96–117, 2022.
- [64] K. Boguslawski and P. W. Ayers, “Linearized Coupled Cluster Correction on the Antisymmetric Product of 1-Reference Orbital Geminals,” *J. Chem. Theory Comput.*, vol. 11, pp. 5252–5261, 2015.
- [65] T. Zoboki, Á. Szabados, and P. R. Surján, “Linearized coupled cluster corrections to antisymmetrized product of strongly orthogonal geminals: Role of dispersive interactions,” *J. Chem. Theory Comput.*, vol. 9, no. 6, pp. 2602–2608, 2013.
- [66] J. Čížek, “On the correlation problem in atomic and molecular systems. calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods,” *Int. J. Quantum Chem.*, vol. 45, no. 11, pp. 4256–4266, 1966.

- [67] K. Jankowski and J. Paldus, “Applicability of coupled-pair theories to quasidegenerate electronic states: A model study,” *Int. J. Quantum Chem.*, vol. 18, no. 5, pp. 1243–1269, 1980.
- [68] A. G. Taube and R. J. Bartlett, “Rethinking linearized coupled-cluster theory,” *J. Chem. Phys.*, vol. 130, no. 14, 2009.
- [69] L. Serrano-Andrés and M. Merchán, “Quantum chemistry of the excited state: 2005 overview,” *J. Mol. Struct. THEOCHEM*, vol. 729, no. 1-2, pp. 99–108, 2005.
- [70] L. González, D. Escudero, and L. Serrano-Andrés, “Progress and challenges in the calculation of electronic excited states,” *Comput. Phys. Commun.*, vol. 13, no. 1, pp. 28–51, 2012.
- [71] U. Kaldor, “The Fock space coupled cluster method: theory and application,” *Theor. Chem. Acc.*, vol. 80, no. 6, pp. 427–439, 1991.
- [72] M. Musiał and L. Meissner, “The Fock-space coupled-cluster method in the calculation of excited state properties,” *Collect. Czech. Chem. Commun.*, vol. 70, no. 6, pp. 811–825, 2005.
- [73] A. Landau, E. Eliav, and U. Kaldor, “Intermediate Hamiltonian Fock-space coupled-cluster method,” *Adv. Quantum Chem.*, vol. 39, pp. 171–188, 2001.
- [74] M. Musiał, L. Meissner, S. A. Kucharski, and R. J. Bartlett, “Molecular applications of the intermediate Hamiltonian Fock-space coupled-cluster method for calculation of excitation energies,” *J. Chem. Phys.*, vol. 122, no. 22, p. 224110, 2005.
- [75] D. I. Lyakh, M. Musiał, V. F. Lotrich, and R. J. Bartlett, “Multireference nature of chemistry: The coupled-cluster view,” *Chem. Rev.*, vol. 112, no. 1, pp. 182–243, 2012.
- [76] D. J. Rowe, “Equations-of-motion method and extended shell model,” *Rev. Mod. Phys.*, vol. 40, pp. 153–166, 1968.
- [77] J. Geersten, M. Rittby, and R. J. Bartlett, “The equation-of-motion coupled cluster method: excitation energies of Be and CO,” *Chem. Phys. Lett.*, vol. 164, pp. 57–62, 1989.
- [78] J. F. Stanton and R. J. Bartlett, “The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties,” *J. Chem. Phys.*, vol. 98, no. 9, pp. 7029–7039, 1993.

- [79] F. Kossoski, A. Marie, A. Scemama, M. Caffarel, and P.-F. Loos, “Excited states from state-specific orbital-optimized pair coupled cluster,” *J. Chem. Theory Comput.*, vol. 17, no. 8, pp. 4756–4768, 2021.
- [80] K. Boguslawski, “Targeting excited states in all-trans polyenes with electron-pair states,” *J. Chem. Phys.*, vol. 145, p. 234105, 2016.
- [81] K. Boguslawski, “Erratum: “Targeting excited states in all-trans polyenes with electron-pair states,”” *J. Chem. Phys.*, vol. 147, no. 13, p. 139901, 2017.
- [82] J. B. Foresman, M. Head-Gordon, J. A. Pople, and M. J. Frisch, “Toward a systematic molecular orbital theory for excited states,” *J. Chem. Phys.*, vol. 96, no. 1, pp. 135–149, 1992.
- [83] K. Boguslawski, “Targeting Doubly Excited States with Equation of Motion Coupled Cluster Theory Restricted to Double Excitations,” *J. Chem. Theory Comput.*, vol. 15, no. 1, pp. 18–24, 2019.
- [84] M. Ravi, A. Perera, Y. C. Park, and R. J. Bartlett, “Excited states with pair coupled cluster doubles tailored coupled cluster theory,” *J. Chem. Phys.*, vol. 159, no. 9, p. 094101, 2023.
- [85] A. Nowak, P. Tecmer, and K. Boguslawski, “Assessing the accuracy of simplified coupled cluster methods for electronic excited states in f0 actinide compounds,” *Phys. Chem. Chem. Phys.*, vol. 21, no. 35, pp. 19039–19053, 2019.
- [86] H. Koch and P. Jørgensen, “Coupled cluster response functions,” *J. Chem. Phys.*, vol. 93, no. 5, pp. 3333–3344, 1990.
- [87] T. B. Pedersen and H. Koch, “Coupled cluster response functions revisited,” *J. Chem. Phys.*, vol. 106, no. 19, pp. 8059–8072, 1997.
- [88] T. Helgaker and P. Jørgensen, “Configuration-interaction energy derivatives in a fully variational formulation,” *Theor. Chem. Acc.*, vol. 75, no. 2, pp. 111–127, 1989.
- [89] P. Jørgensen and T. Helgaker, “Møller–Plesset energy derivatives,” *J. Chem. Phys.*, vol. 89, no. 3, pp. 1560–1570, 1988.
- [90] G. Trucks, E. Salter, J. Noga, and R. Bartlett, “Analytic many-body perturbation theory MBPT(4) response properties,” *Chem. Phys. Lett.*, vol. 150, no. 1-2, pp. 37–44, 1988.
- [91] E. Kraka, J. Gauss, and D. Cremer, “Determination and use of response densities,” *J. Mol. Struct. THEOCHEM*, vol. 234, pp. 95–126, 1991.

- [92] J. Rissler, R. M. Noack, and S. R. White, “Measuring Orbital Interaction Using Quantum Information Theory,” *Chem. Phys.*, vol. 323, pp. 519–531, 2006.
- [93] L. Freitag, S. Knecht, S. F. Keller, M. G. Delcey, F. Aquilante, T. B. Pedersen, R. Lindh, M. Reiher, and L. Gonzalez, “Orbital entanglement and CASSCF analysis of the Ru-NO bond in a Ruthenium nitrosyl complex,” *Phys. Chem. Chem. Phys.*, vol. 17, pp. 13769–13769, 2015.
- [94] K. Boguslawski and P. Tecmer, “Orbital entanglement in quantum chemistry,” *Int. J. Quantum Chem.*, vol. 115, no. 19, pp. 1289–1295, 2015.
- [95] K. Boguslawski and P. Tecmer, “Erratum: Orbital entanglement in quantum chemistry,” *Int. J. Quantum Chem.*, vol. 117, no. 24, p. e25455, 2017.
- [96] K. Boguslawski, P. Tecmer, and Ö. Legeza, “Analysis of two-orbital correlations in wave functions restricted to electron-pair states,” *Phys. Rev. B*, vol. 94, no. 15, p. 155126, 2016.
- [97] A. Nowak, Ö. Legeza, and K. Boguslawski, “Orbital entanglement and correlation from pCCD-tailored coupled cluster wave functions,” *J. Chem. Phys.*, vol. 154, p. 084111, Feb. 2021.
- [98] H. Sekino and R. J. Bartlett, “A linear response, coupled-cluster theory for excitation energy,” *Int. J. Quantum Chem.*, vol. 26, no. S18, pp. 255–265, 1984.
- [99] H. Koch, H. J. A. Jensen, P. Jørgensen, and T. Helgaker, “Excitation energies from the coupled cluster singles and doubles linear response function (CCSDLR). Applications to Be, CH⁺, CO, and H₂O,” *J. Chem. Phys.*, vol. 93, no. 5, pp. 3345–3350, 1990.
- [100] H. Koch, R. Kobayashi, A. Sanchez de Merás, and P. Jørgensen, “Calculation of size-intensive transition moments from the coupled cluster singles and doubles linear response function,” *J. Chem. Phys.*, vol. 100, no. 6, pp. 4393–4400, 1994.
- [101] S. Ahmadkhani, K. Boguslawski, and P. Tecmer, “Linear response pCCD-based methods: LR-pCCD and LR-pCCD+ S approaches for the efficient and reliable modeling of excited state properties,” *J. Chem. Theory Comput.*, vol. 20, no. 23, pp. 10443–10452, 2024.
- [102] B. Jeziorski and R. Moszyński, “Explicitly connected expansion for the average value of an observable in the coupled-cluster theory,” *Int. J. Quantum Chem.*, vol. 48, no. 3, pp. 161–183, 1993.

- [103] R. Moszyński, P. S. Żuchowski, and B. Jeziorski, “Time-independent coupled-cluster theory of the polarization propagator,” *Collect. Czechoslov. Chem. Commun.*, vol. 70, no. 8, pp. 1109–1132, 2005.
- [104] T. Korona, M. Przybytek, and B. Jeziorski, “Time-independent coupled cluster theory of the polarization propagator. Implementation and application of the singles and doubles model to dynamic polarizabilities and van der Waals constants,” *Mol. Phys.*, vol. 104, no. 13-14, pp. 2303–2316, 2006.
- [105] T. Korona and B. Jeziorski, “Dispersion energy from density-fitted density susceptibilities of singles and doubles coupled cluster theory,” *J. Chem. Phys.*, vol. 128, no. 14, p. 144107, 2008.
- [106] T. Korona and B. Jeziorski, “One-electron properties and electrostatic interaction energies from the expectation value expression and wave function of singles and doubles coupled cluster theory,” *J. Chem. Phys.*, vol. 125, p. 184109, Nov. 2006.
- [107] T. Korona, “The effect of local approximations on first-order properties from expectation-value coupled cluster theory,” *Theor. Chem. Acc.*, vol. 129, pp. 15–30, 2011.
- [108] M. S. Gordon, D. G. Fedorov, S. R. Pruitt, and L. V. Slipchenko, “Fragmentation methods: A route to accurate calculations on large systems,” *Chem. Rev.*, vol. 112, no. 1, pp. 632–672, 2012.
- [109] A. Severo Pereira Gomes and C. R. Jacob, “Quantum-chemical embedding methods for treating local electronic excitations in complex chemical systems,” *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.*, vol. 108, p. 222, 2012.
- [110] Q. Sun and G. K.-L. Chan, “Quantum embedding theories,” *Acc. Chem. Res.*, vol. 49, no. 12, pp. 2705–2712, 2016.
- [111] L. O. Jones, M. A. Mosquera, G. C. Schatz, and M. A. Ratner, “Embedding methods for quantum chemistry: applications from materials to life sciences,” *J. Am. Chem. Soc.*, vol. 142, no. 7, pp. 3281–3295, 2020.
- [112] A. Warshel and M. Levitt, “Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme,” *J. Mol. Biol.*, vol. 103, no. 2, pp. 227–249, 1976.
- [113] M. J. Field, P. A. Bash, and M. Karplus, “A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations,” *J. Comput. Chem.*, vol. 11, no. 6, pp. 700–733, 1990.

- [114] H. M. Senn and W. Thiel, “QM/MM methods for biomolecular systems,” *Angew. Chem. Int. Ed.*, vol. 48, no. 7, pp. 1198–1229, 2009.
- [115] P. Cortona, “Self-consistently determined properties of solids without band-structure calculations,” *Phys. Rev. B*, vol. 44, no. 16, p. 8454, 1991.
- [116] P. Cortona, “Direct determination of self-consistent total energies and charge densities of solids: A study of the cohesive properties of the alkali halides,” *Phys. Rev. B*, vol. 46, no. 4, p. 2008, 1992.
- [117] T. A. Wesolowski and A. Warshel, “Frozen density functional approach for ab initio calculations of solvated molecules,” *J. Phys. Chem.*, vol. 97, no. 30, pp. 8050–8053, 1993.
- [118] G. Senatore and K. Subbaswamy, “Density dependence of the dielectric constant of rare-gas crystals,” *Phys. Rev. B*, vol. 34, no. 8, p. 5754, 1986.
- [119] M. Johnson, K. Subbaswamy, and G. Senatore, “Hyperpolarizabilities of alkali halide crystals using the local-density approximation,” *Phys. Rev. B*, vol. 36, no. 17, p. 9202, 1987.
- [120] C. R. Jacob and J. Neugebauer, “Subsystem density-functional theory,” *WIREs Comput. Mol. Sci.*, vol. 4, no. 4, pp. 325–362, 2014.
- [121] N. Ricardi, C. E. González-Espinoza, and T. A. Wesolowski, “N-representability of the target density in frozen-density embedding theory based methods: Numerical significance and its relation to electronic polarization,” *J. Chem. Phys.*, vol. 157, no. 6, p. 064108, 2022.
- [122] T. A. Wesolowski and J. Weber, “Kohn-Sham equations with constrained electron density: an iterative evaluation of the ground-state electron density of interacting molecules,” *Chem. Phys. Lett.*, vol. 248, no. 1-2, pp. 71–76, 1996.
- [123] M. Iannuzzi, B. Kirchner, and J. Hutter, “Density functional embedding for molecular systems,” *Chem. Phys. Lett.*, vol. 421, no. 1-3, pp. 16–20, 2006.
- [124] M. Zbiri, M. Atanasov, C. Daul, J. M. Garcia-Lastra, and T. A. Wesolowski, “Application of the density functional theory derived orbital-free embedding potential to calculate the splitting energies of lanthanide cations in chloroelpasolite crystals,” *Chem. Phys. Lett.*, vol. 397, no. 4-6, pp. 441–446, 2004.
- [125] J. García-Lastra, T. Wesolowski, M. Barriuso, J. Aramburu, and M. Moreno, “Optical and vibrational properties of MnF_6^{4-} complexes in cubic fluoroperovskites: insight through embedding calculations using Kohn–Sham equations

- with constrained electron density,” *J. Phys.: Condens. Matter*, vol. 18, no. 5, pp. 1519–1534, 2006.
- [126] M. Zbiri, C. A. Daul, and T. A. Wesolowski, “Effect of the f-orbital delocalization on the ligand-field splitting energies in lanthanide-containing elpasolites,” *J. Chem. Theory Comput.*, vol. 2, no. 4, pp. 1106–1111, 2006.
- [127] T. A. Wesolowski, “Hydrogen-bonding-induced shifts of the excitation energies in nucleic acid bases: an interplay between electrostatic and electron density overlap effects,” *J. Am. Chem. Soc.*, vol. 126, no. 37, pp. 11444–11445, 2004.
- [128] J. Neugebauer, C. R. Jacob, T. A. Wesolowski, and E. J. Baerends, “An explicit quantum chemical method for modeling large solvation shells applied to aminocoumarin C151,” *J. Phys. Chem. A*, vol. 109, no. 34, pp. 7805–7814, 2005.
- [129] J. Neugebauer, M. J. Louwerse, E. J. Baerends, and T. A. Wesolowski, “The merits of the frozen-density embedding scheme to model solvatochromic shifts,” *J. Chem. Phys.*, vol. 122, no. 9, p. 094115, 2005.
- [130] J. Neugebauer, “Couplings between electronic transitions in a subsystem formulation of time-dependent density functional theory,” *J. Chem. Phys.*, vol. 126, no. 13, p. 134116, 2007.
- [131] J. Neugebauer, “On the calculation of general response properties in subsystem density functional theory,” *J. Chem. Phys.*, vol. 131, no. 8, p. 084104, 2009.
- [132] P. Huang and E. A. Carter, “Advances in correlated electronic structure methods for solids, surfaces, and nanostructures,” *Annu. Rev. Phys. Chem.*, vol. 59, no. 1, pp. 261–290, 2008.
- [133] N. Govind, Y. Wang, A. Da Silva, and E. Carter, “Accurate ab initio energetics of extended systems via explicit correlation embedded in a density functional environment,” *Chem. Phys. Lett.*, vol. 295, no. 1-2, pp. 129–134, 1998.
- [134] P. Huang and E. A. Carter, “Ab initio explanation of tunneling line shapes for the Kondo impurity state,” *Nano Lett.*, vol. 8, no. 4, pp. 1265–1269, 2008.
- [135] T. A. Wesolowski, “Embedding a multideterminantal wave function in an orbital-free environment,” *Phys. Rev. A*, vol. 77, no. 1, p. 012504, 2008.
- [136] N. Govind, Y. A. Wang, and E. A. Carter, “Electronic-structure calculations by first-principles density-based embedding of explicitly correlated systems,” *J. Chem. Phys.*, vol. 110, no. 16, pp. 7677–7688, 1999.
- [137] T. Klüner, N. Govind, Y. A. Wang, and E. A. Carter, “Prediction of elec-

- tronic excited states of adsorbates on metal surfaces from first principles,” *Phys. Rev. Lett.*, vol. 86, no. 26, p. 5954, 2001.
- [138] T. Klüner, N. Govind, Y. A. Wang, and E. A. Carter, “Periodic density functional embedding theory for complete active space self-consistent field and configuration interaction calculations: Ground and excited states,” *J. Chem. Phys.*, vol. 116, no. 1, pp. 42–54, 2002.
- [139] P. Huang and E. A. Carter, “Local electronic structure around a single Kondo impurity,” *Nano Lett.*, vol. 6, no. 6, pp. 1146–1150, 2006.
- [140] P. Huang and E. A. Carter, “Self-consistent embedding theory for locally correlated configuration interaction wave functions in condensed matter,” *J. Chem. Phys.*, vol. 125, no. 8, p. 084102, 2006.
- [141] D. Lahav and T. Klüner, “A self-consistent density based embedding scheme applied to the adsorption of CO on Pd(111),” *J. Phys.: Condens. Matter*, vol. 19, no. 22, p. 226001, 2007.
- [142] A. S. P. Gomes, C. R. Jacob, and L. Visscher, “Calculation of local excitations in large systems by embedding wave-function theory in density-functional theory,” *Phys. Chem. Chem. Phys.*, vol. 10, no. 35, pp. 5353–5362, 2008.
- [143] A. S. P. Gomes, C. R. Jacob, F. Réal, L. Visscher, and V. Vallet, “Towards systematically improvable models for actinides in condensed phase: the electronic spectrum of uranyl in $\text{Cs}_2\text{UO}_2\text{Cl}_4$ as a test case,” *Phys. Chem. Chem. Phys.*, vol. 15, no. 36, p. 15153, 2013.
- [144] S. Höfener, A. Severo Pereira Gomes, and L. Visscher, “Molecular properties via a subsystem density functional theory formulation: A common framework for electronic embedding,” *J. Chem. Phys.*, vol. 136, no. 4, p. 044104, 2012.
- [145] S. Höfener, A. S. P. Gomes, and L. Visscher, “Solvatochromic shifts from coupled-cluster theory embedded in density functional theory,” *J. Chem. Phys.*, vol. 139, no. 10, p. 104106, 2013.
- [146] S. Höfener and L. Visscher, “Calculation of electronic excitations using wave-function in wave-function frozen-density embedding,” *J. Chem. Phys.*, vol. 137, no. 20, p. 204120, 2012.
- [147] S. Höfener and L. Visscher, “Wave Function Frozen-Density Embedding: Coupled Excitations,” *J. Chem. Theory Comput.*, vol. 12, pp. 549–557, Feb. 2016.
- [148] K. Boguslawski, A. Leszczyk, A. Nowak, F. Brzęk, P. S. Żuchowski, D. Kędziera, and P. Tecmer, “Pythonic black-box electronic structure tool

- (PyBEST). an open-source python platform for electronic structure calculations at the interface between chemistry and physics,” *Comput. Phys. Commun.*, vol. 264, p. 107933, 2021.
- [149] K. Boguslawski, F. Brzęk, R. Chakraborty, K. Cieślak, S. Jahani, A. Leszczyk, A. Nowak, E. Sujkowski, J. Świerczyński, S. Ahmadkhani, D. Kędziera, M. H. Kriebel, P. S. Żuchowski, and P. Tecmer, “Pybest: improved functionality and enhanced performance,” *Comput. Phys. Commun.*, vol. 297, p. 109049, 2024.
- [150] L. Szczuczko and K. Boguslawski, “A cross-platform graphical user interface using web technologies: Simplifying the setup for pybest calculations,” *Int. J. Quantum Chem.*, vol. 125, no. 17, p. e70095, 2025.
- [151] G. t. Te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. van Gisbergen, J. G. Snijders, and T. Ziegler, “Chemistry with ADF,” *J. Comput. Chem.*, vol. 22, no. 9, pp. 931–967, 2001.
- [152] 2022. AMS2022, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, <http://www.scm.com>.
- [153] C. R. Jacob, S. M. Beyhan, R. E. Bulo, A. S. P. Gomes, A. W. Götz, K. Kiewisch, J. Sikkema, and L. Visscher, “Software news and updates PyADF—a scripting framework for multiscale quantum chemistry,” *J. Comput. Chem.*, vol. 32, no. 10, p. 2328, 2011.
- [154] A. Lembarki and H. Chermette, “Obtaining a gradient-corrected kinetic-energy functional from the Perdew-Wang exchange functional,” *Phys. Rev. A*, vol. 50, p. 5328, 1994.
- [155] T. A. Wesółowski, “Density functional theory with approximate kinetic energy functionals applied to hydrogen bonds,” *J. Chem. Phys.*, vol. 106, no. 20, pp. 8516–8526, 1997.
- [156] T. Wesółowski, Y. Ellinger, and J. Weber, “Density functional theory with an approximate kinetic energy functional applied to study structure and stability of weak van der Waals complexes,” *J. Chem. Phys.*, vol. 108, no. 15, pp. 6078–6083, 1998.
- [157] N. Douglas and N. M. Kroll, “Quantum electrodynamical corrections to fine-structure of helium,” *Ann. Phys.*, vol. 82, pp. 89–155, 1974.
- [158] B. A. Hess, “Relativistic electronic-structure calculations employing a 2-component no-pair formalism with external-fields projection operators,” *Phys. Rev. A*, vol. 33, pp. 3742–3748, 1986.

- [159] G. Peinel, “Calculation of dipole moments with a CNDO-APSG procedure,” *Mol. Phys.*, vol. 29, no. 2, pp. 641–643, 1975.
- [160] J. Ogilvie, W. Rodwell, and R. Tipping, “Dipole moment functions of the hydrogen halides,” *J. Chem. Phys.*, vol. 73, no. 10, pp. 5221–5229, 1980.
- [161] M. Piris, X. Lopez, F. Ruipérez, J. Matxain, and J. Ugalde, “A natural orbital functional for multiconfigurational states,” *J. Chem. Phys.*, vol. 134, no. 16, 2011.
- [162] K. Pernal, “The equivalence of the Piris Natural Orbital Functional 5 (PNOF5) and the antisymmetrized product of strongly orthogonal geminal theory,” *Comput. Theor. Chem.*, vol. 1003, pp. 127–129, 2013.
- [163] X. Lopez, M. Piris, M. Nakano, and B. Champagne, “Natural orbital functional calculations of molecular polarizabilities and second hyperpolarizabilities. the hydrogen molecule as a test case,” *J. Phys. B: At. Mol. Opt. Phys.*, vol. 47, no. 1, p. 015101, 2014.
- [164] X. Liu, L. McKemmish, and J. Pérez-Ríos, “The performance of CCSD(T) for the calculation of dipole moments in diatomics,” *Phys. Chem. Chem. Phys.*, vol. 25, no. 5, pp. 4093–4104, 2023.
- [165] R. Chakraborty, S. Ahmadkhani, J. Świerczyński, K. Boguslawski, and P. Tecmer, “Expectation Value-pCCD-Based Methods for Single-Electron Properties,” *J. Phys. Chem. A*, vol. 129, no. 29, pp. 6713–6732, 2025.
- [166] R. Chakraborty, K. Boguslawski, and P. Tecmer, “Static embedding with pair coupled cluster doubles based methods,” *Phys. Chem. Chem. Phys.*, vol. 25, no. 37, pp. 25377–25388, 2023.
- [167] R. Chakraborty, M. M. F. de Moraes, K. Boguslawski, A. Nowak, J. Świerczyński, and P. Tecmer, “Toward Reliable Dipole Moments without Single Excitations: The Role of Orbital Rotations and Dynamical Correlation,” *J. Chem. Theory Comput.*, vol. 20, pp. 4689–4702, 2024.
- [168] A. Hesselmann and O. R. Meitei, “Intermolecular dispersion energies from coupled exact-exchange Kohn-Sham excitation energies and vectors,” *Comput. Theor. Chem.*, vol. 1129, pp. 57–69, 2018.
- [169] R. Chakraborty and P. Tecmer, “Frozen density embedding with pccd electron densities,” *arXiv preprint arXiv:2604.14904*, 2026.
- [170] Z.-L. Cai, K. Sendt, and J. R. Reimers, “Failure of density-functional theory

- and time-dependent density-functional theory for large extended π systems,” *J. Chem. Phys.*, vol. 117, no. 12, pp. 5543–5549, 2002.
- [171] C. Daday, C. Kønig, O. Valsson, J. Neugebauer, and C. Filippi, “State-specific embedding potentials for excitation-energy calculations,” *J. Chem. Theory Comput.*, vol. 9, no. 5, pp. 2355–2367, 2013.
- [172] F. R. Manby, M. Stella, J. D. Goodpaster, and T. F. Miller III, “A simple, exact density-functional-theory embedding scheme,” *J. Chem. Theory Comput.*, vol. 8, no. 8, pp. 2564–2568, 2012.
- [173] J. D. Goodpaster, T. A. Barnes, F. R. Manby, and T. F. Miller, “Accurate and systematically improvable density functional theory embedding for correlated wavefunctions,” *J. Chem. Phys.*, vol. 140, no. 18, p. 18A507, 2014.
- [174] B. Hégely, P. R. Nagy, G. G. Ferenczy, and M. Kállay, “Exact density functional and wave function embedding schemes based on orbital localization,” *J. Chem. Phys.*, vol. 145, no. 6, p. 064107, 2016.
- [175] J. Csóka, B. Hégely, P. R. Nagy, and M. Kállay, “Development of analytic gradients for the Huzinaga quantum embedding method and its applications to large-scale hybrid and double hybrid DFT forces,” *J. Chem. Phys.*, vol. 160, no. 12, p. 124113, 2024.
- [176] S. Huzinaga and A. A. Cantu, “Theory of separability of many-electron systems,” *J. Chem. Phys.*, vol. 55, no. 12, pp. 5543–5549, 1971.
- [177] D. Claudino and N. J. Mayhall, “Automatic partition of orbital spaces based on singular value decomposition in the context of embedding theories,” *J. Chem. Theory Comput.*, vol. 15, no. 2, pp. 1053–1064, 2019.
- [178] B. Paulus, “The method of increments—a wavefunction-based ab initio correlation method for solids,” *Physics reports*, vol. 428, no. 1, pp. 1–52, 2006.
- [179] M. S. Gordon, M. A. Freitag, P. Bandyopadhyay, J. H. Jensen, V. Kairys, and W. J. Stevens, “The Effective Fragment Potential Method: A QM-Based MM Approach to Modeling Environmental Effects in Chemistry,” *J. Phys. Chem. A*, vol. 105, pp. 293–307, Jan. 2001.
- [180] L. V. Slipchenko and P. K. Gurunathan, “Effective fragment potential method: Past, present, and future,” *Fragmentation: Toward Accurate Calculations on Complex Molecular Systems*, pp. 183–208, 2017.
- [181] L. V. Slipchenko, “Detangling solvatochromic effects by the effective fragment potential method,” *J. Phys. Chem. A*, vol. 128, no. 3, pp. 656–669, 2024.

Appendix A

Derivation of XCCSD density matrix

A.1 CC expectation value

The coupled-cluster expectation value of a one-electron operator $\hat{X}$ is written as

$$\bar{X} = \frac{\langle \Phi_0 | e^{\hat{T}^\dagger} \hat{X} e^{\hat{T}} | \Phi_0 \rangle}{\langle \Phi_0 | e^{\hat{T}^\dagger} e^{\hat{T}} | \Phi_0 \rangle}, \quad \hat{T} = \hat{T}_1 + \hat{T}_2. \quad (\text{A.1})$$

We define the normalization factor

$$\mathcal{N} = \langle \Phi_0 | e^{\hat{T}^\dagger} e^{\hat{T}} | \Phi_0 \rangle. \quad (\text{A.2})$$

Thus,

$$\bar{X} = \mathcal{N}^{-1} \langle \Phi_0 | e^{\hat{T}^\dagger} \hat{X} e^{\hat{T}} | \Phi_0 \rangle. \quad (\text{A.3})$$

Inserting the identity operator $\mathbf{1} = e^{\hat{T}} e^{-\hat{T}}$ between $e^{\hat{T}^\dagger}$ and $\hat{X}$, we obtain

$$\bar{X} = \mathcal{N}^{-1} \langle \Phi_0 | e^{\hat{T}^\dagger} e^{\hat{T}} e^{-\hat{T}} \hat{X} e^{\hat{T}} | \Phi_0 \rangle. \quad (\text{A.4})$$

The similarity-transformed operator is

$$\hat{X}_T = e^{-\hat{T}} \hat{X} e^{\hat{T}}. \quad (\text{A.5})$$

Therefore,

$$\bar{X} = \mathcal{N}^{-1} \langle \Phi_0 | e^{\hat{T}^\dagger} e^{\hat{T}} \hat{X}_T | \Phi_0 \rangle. \quad (\text{A.6})$$

Following Jeziorski et al. [102], the operator $\hat{S}$ is defined by

$$e^{\hat{S}} | \Phi_0 \rangle = \frac{e^{\hat{T}^\dagger} e^{\hat{T}} | \Phi_0 \rangle}{\langle \Phi_0 | e^{\hat{T}^\dagger} e^{\hat{T}} | \Phi_0 \rangle}. \quad (\text{A.7})$$

Equivalently,

$$e^{\hat{S}}|\Phi_0\rangle = \mathcal{N}^{-1}e^{\hat{T}^\dagger}e^{\hat{T}}|\Phi_0\rangle. \quad (\text{A.8})$$

Taking the Hermitian conjugate gives

$$\langle\Phi_0|e^{\hat{S}^\dagger} = \mathcal{N}^{-1}\langle\Phi_0|e^{\hat{T}^\dagger}e^{\hat{T}}. \quad (\text{A.9})$$

Substitution into Eq. (A.6) gives

$$\bar{X} = \langle\Phi_0|e^{\hat{S}^\dagger}\hat{X}_T|\Phi_0\rangle. \quad (\text{A.10})$$

Since

$$\hat{X}_T = e^{-\hat{T}}\hat{X}e^{\hat{T}}, \quad (\text{A.11})$$

we have

$$\bar{X} = \langle\Phi_0|e^{\hat{S}^\dagger}e^{-\hat{T}}\hat{X}e^{\hat{T}}|\Phi_0\rangle. \quad (\text{A.12})$$

Because $\hat{S}^\dagger$ is a de-excitation operator,

$$\hat{S}^\dagger|\Phi_0\rangle = 0, \quad e^{-\hat{S}^\dagger}|\Phi_0\rangle = |\Phi_0\rangle. \quad (\text{A.13})$$

Thus the expectation value may be written as

$$\bar{X} = \langle\Phi_0|e^{\hat{S}^\dagger}e^{-\hat{T}}\hat{X}e^{\hat{T}}e^{-\hat{S}^\dagger}|\Phi_0\rangle. \quad (\text{A.14})$$

The exponentials in Eq. (A.14) can be expanded as a product.

The first similarity transformation is

$$e^{-\hat{T}}\hat{X}e^{\hat{T}} = \hat{X} + [\hat{X}, \hat{T}] + \frac{1}{2}[[\hat{X}, \hat{T}], \hat{T}] + \dots. \quad (\text{A.15})$$

The second transformation is

$$e^{\hat{S}^\dagger}\hat{O}e^{-\hat{S}^\dagger} = \hat{O} + [\hat{S}^\dagger, \hat{O}] + \frac{1}{2}[\hat{S}^\dagger, [\hat{S}^\dagger, \hat{O}]] + \dots. \quad (\text{A.16})$$

Applying this to $\hat{O} = e^{-\hat{T}}\hat{X}e^{\hat{T}}$, and using the approximations in Ref. [106] for a

one-particle density matrix, gives

$$\begin{aligned}\bar{X} \approx & \langle \hat{X} \rangle + \langle [\hat{X}, \hat{T}_1] \rangle + \langle [\hat{S}_1^{[1]\dagger}, \hat{X}] \rangle + \langle [\hat{S}_1^{[2]\dagger}, \hat{X}] \rangle \\ & + \langle [\hat{S}_1^{[1]\dagger}, [\hat{X}, \hat{T}_2]] \rangle + \langle [\hat{S}_2^{[1]\dagger}, [\hat{X}, \hat{T}_2]] \rangle.\end{aligned}\quad (\text{A.17})$$

The operator $\hat{S}$ is expanded as

$$\hat{S} = \hat{S}^{[1]} + \hat{S}^{[2]} + \dots \quad (\text{A.18})$$

At first order,

$$\hat{S}_1^{[1]} = \hat{T}_1, \quad (\text{A.19})$$

$$\hat{S}_2^{[1]} = \hat{T}_2. \quad (\text{A.20})$$

The second-order single-excitation contribution is obtained from the connected single-excitation part of

$$[\hat{T}_1^\dagger, \hat{T}_2]. \quad (\text{A.21})$$

Thus,

$$\hat{S}_1^{[2]} = \hat{P}_1 \left([\hat{T}_1^\dagger, \hat{T}_2] \right), \quad (\text{A.22})$$

where

$$\hat{P}_n = \sum_{\substack{i_1 < \dots < i_n \\ a_1 < \dots < a_n}} |\Phi_{i_1 \dots i_n}^{a_1 \dots a_n}\rangle \langle \Phi_{i_1 \dots i_n}^{a_1 \dots a_n}|. \quad (\text{A.23})$$

In particular,

$$\hat{P}_1 = \sum_{ia} |\Phi_i^a\rangle \langle \Phi_i^a|. \quad (\text{A.24})$$

The orders of the terms are

$$\hat{T}_1, \hat{T}_2, \hat{S}_1^{[1]}, \hat{S}_2^{[1]} \sim \mathcal{O}(\hat{T}), \quad \hat{S}_1^{[2]} \sim \mathcal{O}(\hat{T}^2). \quad (\text{A.25})$$

A.2 One-particle density matrix operator

We choose

$$\hat{X} = \hat{E}_p^q = a_q^\dagger a_p. \quad (\text{A.26})$$

The one-particle density matrix is defined as

$$\gamma_p^q = \bar{E}_p^q. \quad (\text{A.27})$$

Therefore, using Eq. (A.17),

$$\begin{aligned}\gamma_p^q = & \langle \hat{E}_p^q \rangle + \langle [\hat{E}_p^q, \hat{T}_1] \rangle + \langle [\hat{S}_1^{[1]\dagger}, \hat{E}_p^q] \rangle + \langle [\hat{S}_1^{[2]\dagger}, \hat{E}_p^q] \rangle \\ & + \langle [\hat{S}_1^{[1]\dagger}, [\hat{E}_p^q, \hat{T}_2]] \rangle + \langle [\hat{S}_2^{[1]\dagger}, [\hat{E}_p^q, \hat{T}_2]] \rangle.\end{aligned}\quad (\text{A.28})$$

A.3 Derivation of $\hat{S}_1^{[2]}$

Define

$$\mathcal{Z} = [\hat{T}_1^\dagger, \hat{T}_2]. \quad (\text{A.29})$$

With

$$\hat{T}_1^\dagger = \sum_{jb} t_j^b a_j^\dagger a_b, \quad (\text{A.30})$$

$$\hat{T}_2 = \frac{1}{4} \sum_{klcd} t_{kl}^{cd} a_c^\dagger a_d^\dagger a_l a_k, \quad (\text{A.31})$$

we have

$$\mathcal{Z} = \frac{1}{4} \sum_{jbklcd} t_j^b t_{kl}^{cd} [a_j^\dagger a_b, a_c^\dagger a_d^\dagger a_l a_k]. \quad (\text{A.32})$$

Using canonical fermionic anticommutation relations, one obtains

$$\begin{aligned}[a_j^\dagger a_b, a_c^\dagger a_d^\dagger a_l a_k] = & \delta_{bc} a_j^\dagger a_d^\dagger a_l a_k - \delta_{bd} a_j^\dagger a_c^\dagger a_l a_k \\ & - \delta_{jk} a_c^\dagger a_d^\dagger a_l a_b + \delta_{jl} a_c^\dagger a_d^\dagger a_k a_b.\end{aligned}\quad (\text{A.33})$$

The single-excitation part is obtained by projection,

$$\hat{S}_1^{[2]} = \hat{P}_1 \mathcal{Z}. \quad (\text{A.34})$$

Equivalently,

$$s_i^a = \langle \Phi_i^a | \mathcal{Z} | \Phi_0 \rangle. \quad (\text{A.35})$$

Using

$$|\Phi_i^a\rangle = a_a^\dagger a_i |\Phi_0\rangle, \quad \langle \Phi_i^a | = \langle \Phi_0 | a_i^\dagger a_a, \quad (\text{A.36})$$

one finds, after connected Wick contractions,

$$s_i^a = \sum_{jb} t_j^b (2t_{ij}^{ab} - t_{ij}^{ba}). \quad (\text{A.37})$$

Hence,

$$\hat{S}_1^{[2]} = \sum_{ia} s_i^a a_a^\dagger a_i, \quad s_i^a = \sum_{jb} t_j^b (2t_{ij}^{ab} - t_{ij}^{ba}). \quad (\text{A.38})$$

A.4 Basic commutators with one-particle operators

Let [Note: not to be confused with the singlet excitation operator defined in Eq. (1.52)]

$$\hat{E}_p^q = a_q^\dagger a_p. \quad (\text{A.39})$$

For a single-excitation operator

$$\hat{E}_i^a = a_a^\dagger a_i, \quad (\text{A.40})$$

the commutator is

$$[a_q^\dagger a_p, a_a^\dagger a_i] = \delta_{pa} a_q^\dagger a_i - \delta_{iq} a_a^\dagger a_p. \quad (\text{A.41})$$

Thus,

$$\begin{aligned} [\hat{E}_p^q, \hat{T}_1] &= \sum_{ia} t_i^a [a_q^\dagger a_p, a_a^\dagger a_i] \\ &= \sum_{ia} t_i^a (\delta_{pa} a_q^\dagger a_i - \delta_{iq} a_a^\dagger a_p). \end{aligned} \quad (\text{A.42})$$

For the occupied–virtual operator

$$\hat{E}_i^a = a_a^\dagger a_i, \quad (\text{A.43})$$

one obtains

$$\langle [\hat{E}_i^a, \hat{T}_1] \rangle = 0. \quad (\text{A.44})$$

The corresponding de-excitation contribution is nonzero. Since

$$\hat{S}_1^{[1]\dagger} = \hat{T}_1^\dagger = \sum_{jb} t_j^b a_j^\dagger a_b, \quad (\text{A.45})$$

we have

$$\langle [\hat{S}_1^{[1]\dagger}, \hat{E}_i^a] \rangle = t_i^a. \quad (\text{A.46})$$

A.5 Double commutator involving $\hat{T}_2$

The double-excitation cluster operator is

$$\hat{T}_2 = \frac{1}{4} \sum_{klcd} t_{kl}^{cd} a_c^\dagger a_d^\dagger a_l a_k. \quad (\text{A.47})$$

The commutator of $\hat{E}_p^q = a_q^\dagger a_p$ with $\hat{T}_2$ is

$$\begin{aligned} [\hat{E}_p^q, \hat{T}_2] &= \frac{1}{4} \sum_{klcd} t_{kl}^{cd} [a_q^\dagger a_p, a_c^\dagger a_d^\dagger a_l a_k] \\ &= \frac{1}{4} \sum_{klcd} t_{kl}^{cd} \left(\delta_{pc} a_q^\dagger a_d^\dagger a_l a_k - \delta_{pd} a_q^\dagger a_c^\dagger a_l a_k \right. \\ &\quad \left. - \delta_{ql} a_c^\dagger a_d^\dagger a_p a_k + \delta_{qk} a_c^\dagger a_d^\dagger a_l a_p \right). \end{aligned} \quad (\text{A.48})$$

The retained double-commutator contributions in XCCSD are

$$\langle [\hat{S}_1^{[1]\dagger}, [\hat{E}_p^q, \hat{T}_2]] \rangle \quad (\text{A.49})$$

and

$$\langle [\hat{S}_2^{[1]\dagger}, [\hat{E}_p^q, \hat{T}_2]] \rangle. \quad (\text{A.50})$$

The first term contributes to the occupied–virtual block.

$$\langle [\hat{S}_1^{[1]\dagger}, [\hat{E}_i^a, \hat{T}_2]] \rangle = \sum_{jb} t_j^b (2t_{ij}^{ab} - t_{ij}^{ba}). \quad (\text{A.51})$$

A.6 Occupied–occupied block

For the occupied–occupied block, set

$$\hat{X} = \hat{E}_i^j = a_j^\dagger a_i, \quad i, j \in \text{occ.} \quad (\text{A.52})$$

The reference contribution is

$$\langle \hat{E}_i^j \rangle = \delta_{ij}. \quad (\text{A.53})$$

$$\langle [\hat{S}_2^{[1]\dagger}, [\hat{E}_i^j, \hat{T}_2]] \rangle. \quad (\text{A.54})$$

Using

$$\hat{S}_2^{[1]\dagger} = \hat{T}_2^\dagger = \frac{1}{4} \sum_{mnab} t_{mn}^{ab} a_m^\dagger a_n^\dagger a_b a_a, \quad (\text{A.55})$$

we obtain

$$\gamma_i^j = \delta_{ij} + \langle [\hat{S}_2^{[1]\dagger}, [\hat{E}_i^j, \hat{T}_2]] \rangle. \quad (\text{A.56})$$

Evaluation of the connected Wick contractions gives

$$\gamma_i^j = \delta_{ij} - \sum_{kab} t_{ik}^{ab} (2t_{jk}^{ab} - t_{jk}^{ba}). \quad (\text{A.57})$$

Thus,

$$\boxed{\gamma_i^j = \delta_{ij} - \sum_{kab} t_{ik}^{ab} (2t_{jk}^{ab} - t_{jk}^{ba})}. \quad (\text{A.58})$$

A.7 Virtual–virtual block

For the virtual–virtual block, we set

$$\hat{X} = \hat{E}_a^b = a_b^\dagger a_a, \quad a, b \in \text{virt.} \quad (\text{A.59})$$

The reference contribution vanishes.

$$\langle \hat{E}_a^b \rangle = 0. \quad (\text{A.60})$$

The leading connected contribution comes from

$$\langle [\hat{S}_2^{[1]\dagger}, [\hat{E}_a^b, \hat{T}_2]] \rangle. \quad (\text{A.61})$$

Thus,

$$\gamma_a^b = \langle [\hat{S}_2^{[1]\dagger}, [\hat{E}_a^b, \hat{T}_2]] \rangle. \quad (\text{A.62})$$

Evaluation of the connected Wick contractions yields

$$\gamma_a^b = \sum_{ijc} t_{ij}^{ac} (2t_{ij}^{bc} - t_{ij}^{cb}). \quad (\text{A.63})$$

Thus,

$$\boxed{\gamma_a^b = \sum_{ijc} t_{ij}^{ac} (2t_{ij}^{bc} - t_{ij}^{cb})}. \quad (\text{A.64})$$

A.8 Occupied–virtual

For the occupied–virtual block, we set

$$\hat{X} = \hat{E}_i^a = a_a^\dagger a_i, \quad i \in \text{occ}, \quad a \in \text{virt.} \quad (\text{A.65})$$

The reference contribution vanishes.

$$\langle \hat{E}_i^a \rangle = 0. \quad (\text{A.66})$$

The direct $\hat{T}_1$ commutator gives no contribution.

$$\langle [\hat{E}_i^a, \hat{T}_1] \rangle = 0. \quad (\text{A.67})$$

The leading contribution comes from the de-excitation term.

$$\langle [\hat{S}_1^{[1]\dagger}, \hat{E}_i^a] \rangle = t_i^a. \quad (\text{A.68})$$

The second-order contribution is

$$\langle [\hat{S}_1^{[2]\dagger}, \hat{E}_i^a] \rangle = \sum_{jb} t_j^b (2t_{ij}^{ab} - t_{ij}^{ba}). \quad (\text{A.69})$$

Thus,

$$\boxed{\gamma_i^a = t_i^a + \sum_{jb} t_j^b (2t_{ij}^{ab} - t_{ij}^{ba})}. \quad (\text{A.70})$$

A.9 Final XCCSD one-particle density matrix

Define

$$\tilde{t}_{ij}^{ab} = 2t_{ij}^{ab} - t_{ij}^{ba}. \quad (\text{A.71})$$

The occupied–occupied block is

$$\boxed{\gamma_i^j = \delta_{ij} - \sum_{kab} t_{ik}^{ab} \tilde{t}_{jk}^{ab}}. \quad (\text{A.72})$$

The virtual–virtual block is

$$\boxed{\gamma_a^b = \sum_{ijc} t_{ij}^{ac} \tilde{t}_{ij}^{bc}}. \quad (\text{A.73})$$

The occupied–virtual block is

$$\boxed{\gamma_i^a = t_i^a + \sum_{jb} t_j^b \tilde{t}_{ij}^{ab}}. \quad (\text{A.74})$$

Appendix B

Publications



Cite this: *Phys. Chem. Chem. Phys.*,
2023, 25, 25377

Received 30th May 2023,
Accepted 18th August 2023

DOI: 10.1039/d3cp02502k

rsc.li/pccp

Static embedding with pair coupled cluster doubles based methods†

Rahul Chakraborty,[†] Katharina Boguslawski[†] and Paweł Tecmer[†]*

Quantum embedding methods have recently been significantly developed to model large molecular structures. This work proposes a novel wave function theory in a density functional theory (WTF-in-DFT) embedding scheme based on pair-coupled cluster doubles (pCCD)-type methods. While pCCD can reliably describe strongly-correlated systems with mean-field-like computational cost, the large extent of the dynamic correlation can be accounted for by (linearized) coupled-cluster corrections on top of the pCCD wave function. Here we focus on the linearized coupled-cluster singles and doubles (LCCSD) ansatz for electronic ground states and its extension to excited states within the equation of motion (EOM) formalism. We test our EOM-pCCD-LCCSD-in-DFT approach for the vertical excitation energies of the hydrogen-bonded water–ammonia complex, micro-solvated thymine, and uranyl tetrahalides ($\text{UO}_2\text{X}_4^{2-}$, X = F, Cl, Br). Furthermore, we assess the quality of the embedding potential using an orbital entanglement and correlation analysis. The approximate embedding models successfully capture changes in the excitation energies going from bare fragments to supramolecular structures and represent a promising computational method for excited states in large molecular systems.

1. Introduction

The applicability of traditional highly accurate wave function theory (WFT)-based quantum chemistry methods like configuration interaction,¹ coupled cluster,² and their various truncation schemes are restricted due to their steep computational scaling. An alternate approach to circumvent this problem is embedding methods.^{3–5} Guided by the fact that the electron correlation is ‘local’ in nature,^{6,7} more accurate, albeit computationally expensive, higher-level methods are used to treat only a small subset of atoms in the whole molecular structure, also called the system. The remaining larger part of the molecular structure, often called the environment, is approximately represented by lower-level methods.^{8–12} Hence, any computational approach based on the fragmentation^{13–16} of the supramolecular structure and mix-and-match of various methods for different fragments can be regarded as embedding. The overall

goal is to reduce the computational cost of treating the supramolecule with expensive methods and focus on a small region of interest where the physical or chemical processes are centered, like electronic transitions.

The different flavors of quantum embedding arise from combining various computational methods. The quantum mechanics/molecular mechanics (QM/MM) variant is a popular class of embedding schemes.^{17,18} Since the quantum mechanical description of the environment fragment must explicitly include the purely quantum mechanical interaction terms such as Pauli repulsion, QM/QM embedding methods were introduced. In this class of embedding approaches, the system part is described with orbitals. At the same time, the environment is mimicked only through its nuclear electrostatic potential and its electron density ρ_{env} , calculated in the presence of the electron density of the system ρ_{sys} .

One of the most prominent examples of the QM/QM embedding approach is the DFT-in-DFT embedding.^{19–24} The frozen density embedding (FDE) method was developed in this context.^{25,26} In this approach, the density of the environment is kept frozen, and the system’s ground state density and energy are calculated using the frozen environment density as an external field. An iterative ‘freeze-and-thaw’ cycle was also employed to improve the fragment (system and environment) densities in response to changes in one another.²⁷ FDE has also been extended to excited states using time-dependent density functional theory (TD-DFT).^{28–30} Still, the DFT-in-DFT approach can only give an approximate picture of the interactions

*Institute of Physics, Faculty of Physics, Astronomy, and Informatics, Nicolaus Copernicus University in Toruń, Grudziadzka 5, 87-100 Toruń, Poland.
E-mail: ptecmer@fizyka.umk.pl*

† Electronic supplementary information (ESI) available: Optimized geometry of $\text{H}_2\text{O} \cdots \text{NH}_3$, EOM-CCSD excitation energies of $\text{H}_2\text{O} \cdots \text{NH}_3$ and thymine in the H_2O cluster, the individual monomers, environmental shifts calculated with EOM-pCCD-LCCSD and EOM-CCSD, EOM-CCSD vertical excitation energies, single-orbital entropies, pCCD optimized orbitals, and mutual information of uranyl and its tetrahalides, the influence of uranyl spin-free orbital energies by the presence of spin-orbit and spin-other-orbit effects, and EOM-CCSD vertical excitation energies of thymine and micro-solvated thymine. See DOI: <https://doi.org/10.1039/d3cp02502k>

between the structure fragments mainly because of the non-additive kinetic potential term present in the interaction potential and intrinsic limitations of DFT.³¹ Developing more accurate embedding schemes is still an active field of research. Examples are the studies with inverted potentials^{32,33} and with orthogonal orbitals from different fragments.^{34,35}

Parallel to the popularity of DFT-in-DFT embedding, WFTs have also been augmented with DFT in the so-called WFT-in-DFT embedding.^{8,36–43} In the basic WFT-in-DFT embedding formalism, the energy functional is expressed as^{36–38,44}

$$E[\Psi_{\text{sys}}^{\text{WFT}}, \rho_{\text{env}}^{\text{DFT}}] = E_{\text{sys}}[\Psi_{\text{sys}}^{\text{WFT}}] + E_{\text{env}}[\rho_{\text{env}}^{\text{DFT}}] + E_{\text{int}}[\rho_{\text{sys}}^{\text{WFT}}, \rho_{\text{env}}^{\text{DFT}}], \quad (1)$$

where E_{sys} is the energy of the embedded system calculated with WFT through its characteristic wave function $\Psi_{\text{sys}}^{\text{WFT}}$, whereas E_{env} is the energy of the environment obtained with DFT using its electron density $\rho_{\text{env}}^{\text{DFT}}$. The third term in eqn (1) is described as

$$\begin{aligned} E_{\text{int}} = & E_{\text{NN}} + \int \rho_{\text{sys}}(r) \nu_{\text{env}}^{\text{nuc}} dr + \int \rho_{\text{env}}(r) \nu_{\text{sys}}^{\text{nuc}} dr \\ & + \frac{1}{2} \int \frac{\rho_{\text{sys}}(r) \rho_{\text{env}}(r')}{|r - r'|} dr dr' \\ & + E_{\text{xc}}^{\text{nadd}}[\rho_{\text{sys}}, \rho_{\text{env}}] + T_{\text{s}}^{\text{nadd}}[\rho_{\text{sys}}, \rho_{\text{env}}]. \end{aligned} \quad (2)$$

In the above equation, E_{NN} is the nuclear–nuclear repulsion energy, $\rho_{\text{sys}}^{\text{WFT}}$ is the electron density of the system fragment obtained using the WFT method of choice, $\nu_{\text{sys}}^{\text{nuc}}$ and $\nu_{\text{env}}^{\text{nuc}}$ are nuclear electrostatic potential terms from the system and environment fragments respectively. $E_{\text{xc}}^{\text{nadd}}$ and $T_{\text{s}}^{\text{nadd}}$ correspond to the non-additive parts of the exchange–correlation (xc) and kinetic energy functionals, respectively. The main attractive feature of WFT-in-DFT is that any WFT ansatz can be used along with DFT, and the existing quantum chemistry codes can be modified to account for the DFT environment.^{8,45} The approach has drawn attention mainly due to the shortcomings of DFT in treating bond-breaking, strongly-correlated systems, and electronically excited states. The method has seen extensive applications in surface chemistry and materials design through the pioneering works of Carter and coworkers.^{36,46,47} Manby and coworkers did extensive work coupling DFT with correlated wave function methods like coupled cluster and other projection-based wave functions.^{34,48–50} The WFT-in-DFT embedding has also been used to study reaction pathways^{50,51} and the excited state,^{3,52–55} including four-component Hamiltonians.⁸

Given the ever-expanding horizon of electronic structure theory, there is still room for new WFT-in-DFT embedding schemes in terms of applicability and computational cost. In this work, we propose the recently developed geminal-motivated⁵⁶ pair coupled cluster doubles (pCCD) ansatz^{57–59} as the WFT component in the WFT-in-DFT scheme. Unlike any other WFT method used within the WFT-in-DFT framework, pCCD can work within its optimized orbital basis without defining active spaces. The orbital-optimized variant of pCCD has been shown to model strongly-correlated systems, albeit

with mean-field-like computational costs effectively.^{58–62} The size-extensive and size-consistent nature of orbital-optimized pCCD has led to a wide range of applications for covalent molecules,^{60,62–68} non-covalent systems,^{69,70} and excited states.^{71–75} The successes of pCCD-based methods in modeling complex electronic structures motivate us to use them within the WFT-in-DFT framework. An additional advantage of pCCD-type methods is the ability to compute orbital entanglement and correlation in terms of the single-orbital entropy and mutual information.^{76–80} Such measures can be used to partition the supramolecule into a system and environment or assess the quality of the embedding potential in the WFT-in-DFT approach.

This work is organized as follows. Section 2 summarizes the pCCD-based methods applied in this work, the implementation details of the embedding schemes used, and orbital correlations. The computation details are provided in Section 3. Section 4 includes the analysis of electronic excitation energies in the selected test systems obtained from supramolecular and embedding calculations. Furthermore, we analyze the single-orbital entropies in uranyl complexes to validate the quality of the embedding potential. We conclude in Section 5.

2. Theory

2.1 pCCD-based methods

One of the newest simplified members in the group of coupled cluster methods is the pCCD ansatz^{56,57,59} and reads as

$$|\Psi_{\text{pCCD}}\rangle = \exp\left(\sum_i^{\text{occ}} \sum_a^{\text{virt}} c_i^a \hat{a}_a^\dagger \hat{a}_i\right) |\Phi_0\rangle = \exp(\hat{T}_{\text{p}}) |\Phi_0\rangle, \quad (3)$$

where only the electron-pair excitations operator $\hat{T}_{\text{p}}$ enters the cluster operator, which substitutes an $(\alpha\beta)$ electron-pair from an occupied orbital to a virtual orbital ($i\bar{i} \rightarrow a\bar{a}$). In the above equation, $\hat{a}^\dagger(\hat{a})$ are the electron creation (annihilation) operators running over all virtual (occupied) orbitals a (i). The Φ_0 is some independent particle model, usually the Hartree–Fock wave function. While the method is size-extensive, size-consistency is achieved through variational orbital optimization.^{58,60,62} Numerous computational studies demonstrate⁵⁹ that the pCCD wave function captures static electron correlation with a mean field-like cost, with accurately reproducing the electron correlation spectrum compared to more expensive wave-function-based methods.⁷⁹ As pCCD includes only electron-pair excitations, it, however, misses a large fraction of the weak/dynamic electron correlation effects. To compensate for that, various *a posteriori* corrections have been proposed.⁵⁶ One example is a linearized coupled cluster (LCC) ansatz on top of the pCCD reference state.⁸¹ In the LCC correction, we approximate the exponential coupled cluster ansatz with a pCCD reference wave function as

$$\Psi_{\text{pCCD-LCC}} \approx (1 + \hat{T}') |\Psi_{\text{pCCD}}\rangle, \quad (4)$$

where $\hat{T}'$ is the cluster operator containing various levels of excitations, excluding electron-pair excitations as they are already accounted for by the pCCD reference state. When we

include both single as well as non-pair double electron excitation operators,

$$\hat{T}' = \hat{T}_1 + \hat{T}_2' = \sum_i^{\text{occ}} \sum_a^{\text{virt}} t_i^a \hat{E}_{ai} + \frac{1}{2} \sum_{i,j}^{\text{occ}} \sum_{a,b}^{\text{virt}} t_{ij}^{ab} \hat{E}_{ai} \hat{E}_{bj}, \quad (5)$$

where

$$\hat{E}_{ai} = a_a^\dagger a_i + a_a^\dagger a_i \quad (6)$$

is the singlet excitation operator, we obtain the pCCD-LCCSD (pCCD-LCC singles and doubles) approach. The prime in the summation of eqn (5) indicates that we exclude electron-pair excitation, that is, terms where simultaneously $i = j$ and $a = b$. This method has shown quite robust performances in various theoretical studies,^{67,69,73,82} especially in cases involving the interplay of static and dynamic electron correlations.

2.2 Targeting excited states: EOM-pCCD

Excited states can be modeled within the pCCD framework using the equation of motion (EOM) formalism,^{83–85} giving various EOM-pCCD flavours depending on the choice of the cluster operator or EOM ansatz.^{71,72,86,87} Currently, there are three variants of EOM-pCCD available. The first one, pCCD+S involves a posterior inclusion of single excitations in the EOM ansatz. Thus, the configuration-interaction-type excitation operator in the EOM treatment includes both the single and pair-doubles excitation operators. Since the EOM-pCCD+S model treats the CC reference state differently (only electron-pair excitations) compared to the EOM formalism (both electron-pair and single excitations), it breaks size-intensivity. Although being numerically marginal, the error can be cured by considering an *a posteriori* CCS calculation on top of the pCCD reference function, followed by an EOM treatment of electronic excited states. This approach was labeled as EOM-pCCD-CCS to highlight that a CCS correction on top of pCCD is applied. To target general (non-pair) doubly excited states, the double excitation operator $\hat{T}_2$ is included in the ansatz for $\hat{T}'$ in eqn (4). In that case, the pCCD reference wave function becomes pCCD-LCCSD, and the approach is known as EOM-pCCD-LCCSD.⁷² The EOM-pCCD-LCCSD has shown promising results in modeling electronic excitations in polyenes and actinide systems.^{73,74}

2.3 Implementation of static embedding and point charges

In the WFT-in-DFT scheme, the effect of the environment modeled through DFT embedding is accounted for as a static external potential. This external potential term includes the Coulomb potential of the nuclei and the electron density of the environment, along with the non-additive parts of the exchange–correlation and kinetic energy terms,³⁶

$$\begin{aligned} \nu^{\text{emb}}[\rho_{\text{sys}}^{\text{WFT}}, \rho_{\text{env}}^{\text{DFT}}](r) = & \nu_{\text{env}}^{\text{nuc}}(r) + \int \mathrm{d}r' \frac{\rho_{\text{env}}^{\text{DFT}}(r')}{|r - r'|} \\ & + \frac{\delta E_{\text{XC}}^{\text{nadd}}[\rho_{\text{sys}}^{\text{WFT}}, \rho_{\text{env}}^{\text{DFT}}]}{\delta \rho_{\text{WFT}}} \\ & + \frac{\delta T_{\text{s}}^{\text{nadd}}[\rho_{\text{sys}}^{\text{WFT}}, \rho_{\text{env}}^{\text{DFT}}]}{\delta \rho_{\text{WFT}}}. \end{aligned} \quad (7)$$

This form of potential has seen extensive usage from solid state structures^{36,37} to small hydrogen-bonded clusters.⁸⁸

In this work, we follow the standard procedure of the static embedding approach and further simplify the WFT-in-DFT embedding model by arguing that for molecules where DFT can describe the ground state electronic structure reasonably well, ρ_{sys} can also be obtained through DFT calculations. Thus, the $\rho_{\text{sys}}^{\text{WFT}}$ term can be replaced by $\rho_{\text{sys}}^{\text{DFT}}$ in eqn (7). This assumption helps to avoid multiple computationally expensive and non-trivial WFT calculations in the freeze-and-thaw cycles just to obtain the particular density term. Hence, we use a DFT-in-DFT embedding technique to calculate ρ_{sys} and ρ_{env} . Once the embedding potential is obtained, the matrix elements of $\nu^{\text{emb}}(r)$ represented in the atomic orbital basis set are calculated as

$$\nu_{ij} = \langle \phi_i | \nu^{\text{emb}}(r) | \phi_j \rangle \approx \sum_k w_k \nu^{\text{emb}}(r_k) \phi_i(r_k) \phi_j(r_k), \quad (8)$$

where i, j labels atomic orbitals, and the embedding potential is approximated on a grid with grid points $\{r_k\}$ and associated weights w_k . Thus, the embedding potential expressed in the atomic orbital basis can be straightforwardly determined by evaluating the (product of two) atomic orbitals for each grid point weighted by the embedding potential at this point. The resulting one-particle term is then included in the Hamiltonian of the WFT calculation. We refer the reader to a recent paper by Sharma and Sierka⁸⁹ for more details on efficient embedding implementations.

In the point-charge model, to account for external point charges, we compute an additional set of integrals (in the atomic orbital basis) due to the electron–external-point-charges attraction/repulsion and the potential energy term due to the Coulomb interaction between the external point charges. To properly combine the point-charge model with the Douglas–Kroll–Hess scalar relativistic Hamiltonian (DKH), the DKH transformation also includes the external potential/integrals due to the point-charges to avoid large picture-change errors.^{90,91}

2.4 Orbital entanglement analysis

A quantitative measure of interactions between two orbitals can be achieved by calculating the single-orbital $s(1)_i$ and two-orbital $s(2)_{ij}$ entropies. The first term $s(1)_i$ quantifies the interaction of the i -th orbital with all other orbitals and reads

$$s(1)_i = - \sum_{\alpha=1}^4 \omega_{\alpha,i} \ln \omega_{\alpha,i}, \quad (9)$$

where $\omega_{\alpha,i}$ are the eigenvalues of the one-orbital reduced density matrix (1-ORDM) for the i -th orbital. The 1-ORDM is determined from an N -particle RDM $\hat{\rho} = |\Psi\rangle\langle\Psi|$, tracing out all degrees of freedom other than those of the i -th orbital. Hence, the dimension of the 1-ORDM is equal to that of the one-orbital Fock space, which is 4 for spatial orbitals (unoccupied, doubly occupied, occupied by a single α or β electron).

Similarly, $s(2)_{ij}$ measures the interaction of a pair of orbitals (i,j) with other orbitals, calculated from the eigenvalues $\omega_{\alpha,ij}$ of the two-orbital RDM (2-ORDM),

$$s(2)_{ij} = - \sum_{\alpha=1}^{16} \omega_{\alpha,ij} \ln \omega_{\alpha,ij}, \quad (10)$$

where we have 16 possible occupations for a pair of spatial orbitals. The total correlation between any pair of orbitals can be determined through the orbital-pair mutual information I_{ij} , which is calculated as

$$I_{ij} = s(1)_i + s(1)_j - s(2)_{ij}, \quad (11)$$

with the constraint $i \neq j$ to exclude self-correlations.

These orbital-based entanglement and correlation measures have recently been applied to quantify orbital interactions^{76–78,92–94} in various chemical processes, like bond formation. Since these measures are orbital-dependent and a change in the environment will translate to a variation in single-orbital entropies and the orbital-pair mutual information, they allow us to probe the change in orbital interactions/correlations with respect to a change in the environment. Specifically, $s(1)_i$, $s(2)_{ij}$, and I_{ij} (evaluated for similar orbitals) will differ for the bare subsystem and the supramolecular system. In contrast, the entanglement and correlation spectrum of the embedded system should qualitatively agree with the corresponding spectrum of the latter due to the perturbations experienced by the orbitals in the presence of the environment electron density. Furthermore, the ‘better’ the description of the static embedding, the closer the agreement in the exact and approximated entanglement and correlation spectra (quantitative agreement). Such entanglement and correlation measures can thus be used as a reliable tool for selecting molecular subspaces.

3. Computational details

3.1 Structures

The geometry of the $\text{H}_2\text{O} \cdots \text{NH}_3$ is optimized with the coupled cluster singles doubles and perturbative triples (CCSD(T)) method and the augmented Dunning-type correlation consistent basis sets of quadruple- ζ quality (aug-cc-pVQZ).⁹⁵ Coordinates of the optimized structures are provided in Table S1 of the ESI.† The coordinates of the thymine in a small cluster of water with 6 H_2O molecules around it are taken from Etinski and Marian.⁹⁶ The structural parameters of the uranyl tetrahalides are taken from Ruipérez and Wahlgren.⁹⁷ Point group symmetry is not used in any DFT-in-DFT or WFT-in-DFT calculations.

3.2 Generation of the embedding potential and description of the point charge model

The embedding potentials are generated within the Amsterdam Modeling Suite (AMS2022)^{98–100} and then extracted with the help of the PyADF¹⁰¹ scripting framework. For the uranyl tetrahalides, the scalar ZORA Hamiltonian is used.¹⁰² In all

DFT-in-DFT calculations, we used the triple- ζ double polarization (TZ2P) basis set,¹⁰³ the PW91^{104,105} xc functional, and the PW91k¹⁰⁶ kinetic energy functional. This particular combination of density functionals has been widely used in embedding studies. A point to note here is that the non-additive exchange–correlation term has a small influence on the subsystem density calculations.¹⁰⁷ A detailed analysis of different kinetic energy functionals for FDE was carried out by Visscher and coworkers.¹⁰⁸ The generalized gradient approximation (GGA) functionals were found to show better performance than local density approximation (LDA) functionals for various non-covalently bound molecular fragments as well as coordination compounds. Among GGA functionals, PBE is prone to convergence issues in FDE iterations. The PW91k kinetic energy functional, on the other hand, has been used in a range of embedding studies,^{109,110} including actinides^{8,111,112} and proved to be reliable. In our embedding DFT-in-DFT calculations, a supramolecular DFT calculation is done, followed by DFT calculations on the unrelaxed system and environment fragments. Thus, we did not explore the effect of the choice of the kinetic energy functional on our DFT-in-DFT potential. Then, a freeze-and-thaw cycle is employed to interchange the definition of the system and environment part. Subsystem DFT calculations are carried out on each of them, allowing each to relax in the presence of the electron density of the other for a fixed number of iterations. After 20 freeze-and-thaw cycles, the DFT-in-DFT embedding potential is extracted to perform further quantum chemical calculations in PyBEST.⁶⁶

Due to the partial ionic nature of the bond between the uranium atom and the halides, a point-charge embedding has also been tested as an approximate model for the halide atoms.^{112,113} We included four negative point charges of $-0.5e$ each, centered at the halide coordinates, to replace the halide atoms in a given supramolecular geometry.

3.3 Wave function methods

The pCCD-based electronic structure calculations were carried out in a developer version of the PyBEST software package,⁶⁶ while the EOM-CCSD calculations were done in the Molpro2020.2.1 software package.^{114–116} The static embedding module in PyBEST is written in the Python programming language with some operations handled in C++, interfaced with Pybind11.¹¹⁷ All PyBEST calculations are run with 6 CPUs. Parallelization is used *via* external MKL/BLAS libraries and Python packages like Scipy¹¹⁸ and NumPy.¹¹⁹ Cholesky decomposition was used for the $\text{H}_2\text{O} \cdots \text{NH}_3$ and thymine in H_2O cluster calculations in PyBEST, with a tight threshold of 10^{-8} for Cholesky vectors.¹²⁰ We should stress that the WFT part's computational cost largely determines the WFT-in-DFT approach's computational scaling. Specifically, for pCCD-based methods, the computational scaling ranges from formally $\mathcal{O}(o^2v^2)$, through $\mathcal{O}(o^2v^3)$, to at most $\mathcal{O}(o^2v^4)$ (similar to the CCSD method).

First, we performed orbital optimized pCCD^{58,62} calculations for the ground state and corrected for the missing dynamical correlation energy using an LCCSD correction

(pCCD-LCCSD).⁸¹ Pipek-Mezey orbital localization was performed¹²¹ before the orbital optimization process to accelerate the convergence. We should note that the pCCD orbital optimization results in localized/symmetry-broken orbitals. Still, comparing the atomic contributions and characteristic features of canonical orbitals, the optimized pCCD orbitals can be identified as their traditional canonical counterparts.^{122,123} We follow these notations for the orbitals and excited states only for ease of labeling purpose and not in the strict spectroscopic meaning. The excitation energies were obtained within the EOM formalism^{124,125} on top of pCCD-LCCSD.⁷² In each system, we calculated the 15 lowest-lying roots.

In our EOM-pCCD-LCCSD and EOM-CCSD calculations on the hydrogen-bonded complexes, we applied the aug-cc-pVDZ basis set.⁹⁵ For the uranyl tetrahalides, we used the ANO-RCC-VDZP basis set^{126,127} and the DKH Hamiltonian of second order (DKH2).^{128–131} In correlated calculations for uranyl and its tetrahalides, we introduced a frozen core, where the 1s–4d orbitals are kept frozen for U, along with the 1s of O and F, 1s–2p of Cl, and 1s–3d orbitals of the Br atoms.

4. Results and discussion

In the following, we assess the performance of our proposed WFT-in-DFT methods for electronic excitation energies in the model water-ammonia complex, thymine in a small solvation shell with 6 water molecules, and uranyl tetrahalide complexes ($\text{UO}_2\text{X}_4^{2-}$, X = F, Cl, Br). The latter is further compared to the more approximate point charge models and the supramolecular results.

4.1 Water-ammonia complex

The water-ammonia complex ($\text{H}_2\text{O} \cdots \text{NH}_3$), shown schematically in Fig. 1(a), is a commonly used moiety for testing quantum embedding models.^{88,132} The structure can be easily partitioned into fragments that interact weakly. The ground state of the supramolecular structure is well-represented by single-reference methods, making it a suitable test case for any embedding study.

Table 1 lists the four lowest-lying vertical excitation energies computed with the EOM-pCCD-LCCSD method. The fragment and supramolecular EOM-pCCD-LCCSD calculations align with the EOM-CCSD results in Table S2 of the ESI.† In addition, the shifts from fragment to supramolecular structure calculated with EOM-pCCD-LCCSD are in very good agreement with EOM-CCSD (*cf.* Table S3 of the ESI†). This observation builds upon an earlier study by some of us⁷⁴ showing that EOM-pCCD-LCCSD is, on average, closer to EOM-CCSD(T) than EOM-CCSD. Hence, the supramolecular EOM-pCCD-LCCSD results work as a suitable reference to test our embedding approach.

The embedding calculation is set up by taking the fragment, which is the source of the excitation in the supramolecular case, as the system and the other fragment as the environment. These are denoted as NH_3 and H_2O fragments, respectively, in Table 1. The NH_3 -in- H_2O embedding correctly mimics the supramolecular excitation energies, and the errors are no larger than 0.15 eV. The H_2O -in- NH_3 embedding excitation energy is

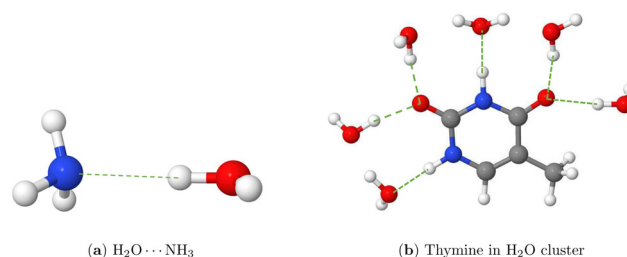


Fig. 1 Structures of the H-bonded species studied. The hydrogen bonds between the system and environment fragments are shown as green dashed lines.

Table 1 EOM-pCCD-LCCSD vertical excitation energies (in eV) for the $\text{H}_2\text{O} \cdots \text{NH}_3$ model system. Comparison between the bare fragments, EOM-pCCD-LCCSD-in-DFT embedding, and supramolecular EOM-pCCD-LCCSD results. The environmental shifts, from fragment to complex, calculated with WFT-in-DFT and supramolecular approaches are given in parenthesis

Fragment	WFT-in-DFT	Supramolecule
NH_3	NH_3 -in- H_2O	$\text{H}_2\text{O} \cdots \text{NH}_3$
6.49	7.05 (0.56)	6.92 (0.43)
8.03	8.63 (0.60)	8.49 (0.46)
8.03	8.67 (0.64)	8.54 (0.51)
H_2O	H_2O -in- NH_3	$\text{H}_2\text{O} \cdots \text{NH}_3$
7.66	7.38 (−0.28)	7.56 (−0.10)

underestimated by 0.18 eV compared to the supramolecular structure. We should note that similar errors were observed in the WFT-in-DFT study by Höfener *et al.*¹³² What is important is that both the H_2O -in- NH_3 and NH_3 -in- H_2O embedding models predict the right signs of the supramolecular shifts with respect to their fragments.

4.2 Thymine in a water cluster

In our next test case, we study the electronic excitations in nucleic acid base thymine in micro-solvation with six water molecules around it. The water molecules have been treated as the environment with a DFT-based embedding potential. The results are shown in Table 2 for the lowest three transitions originating from the thymine fragment. The WFT-in-DFT method correctly predicts the qualitative blue shift in the excitation energies from bare thymine to micro-solvation, as

Table 2 EOM-pCCD-LCCSD vertical excitation energies (in eV) for the thymine microsolvated structure. A comparison between the excitation energies of bare thymine and that of thymine with micro-solvation calculated with EOM-pCCD-LCCSD-in-DFT embedding is shown here. The environmental shifts, from fragment to micro-solvation, calculated with WFT-in-DFT, are given in parenthesis

Thymine	Thymine in micro-solvation (WFT-in-DFT)	ΔE_{env}
5.62	5.75	+0.13
5.80	6.23	+0.43
5.97	6.37	+0.40

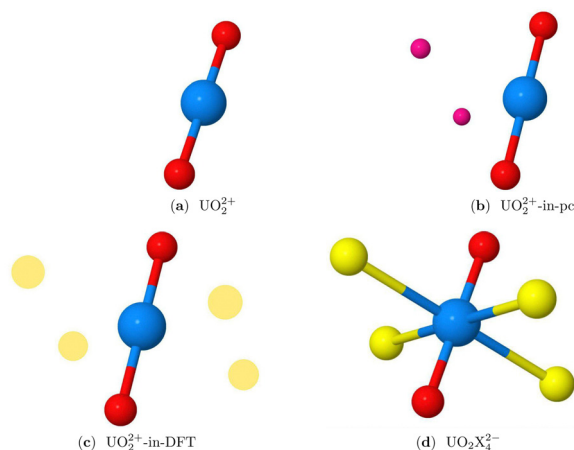


Fig. 2 Schematic diagram of uranyl tetrahalide models studied.

predicted in previous studies^{96,133} and the EOM-CCSD data reported in Table S4 of the ESI.†

4.3 Uranyl tetrahalides

4.3.1 Analysis of electronic spectra. The reliable modeling of the electronic structures and excitation energies of uranium-containing compounds remains a challenge for present-day quantum chemistry. The main reasons behind it are the presence of energetically very close-lying excitations and the involvement of the *f* orbitals. pCCD-based methods represent a promising alternative to standard electronic structure methods in modeling actinide chemistry.^{62,64,65,68,74} We should emphasize that the uranyl pCCD-orbitals are symmetry-broken, and for the σ_u and π_u set of orbitals, we obtain degenerate orbital pairs for each canonical counterpart⁶⁴ as shown in Fig. 3–5. Uranyl tetrahalides are the prototypes of many important actinyl complexes. Their electronic structures and photophysics have been the subject of many theoretical studies.^{8,107,134–142}

Their electronic structure properties and structural parameters are often compared to the bare uranyl cation. Fig. 2 schematically shows the structures of the uranyl and its tetrahalides, including the two model systems where we approximate the tetrahalide environment with point charges (Fig. 2b) and a static embedding potential (Fig. 2c). The presence of ligand-to-metal charge transfer transitions in the upper part of the spectrum poses an additional challenge.^{97,107} In this work, we track the changes caused by the introduction of the halide atoms in three types of the lowest-lying transitions occurring in bare uranyl and investigate whether the trends can be predicted with embedding models. While focusing on spin-free electronic excitations, we should stress that spin-orbit coupling is mandatory to reproduce experimental spectra quantitatively. The influence of spin-orbit coupling on the orbital picture of uranyl and its bonding is discussed, for example, by De Jong *et al.*¹⁴³ and by Dyall.¹⁴⁴ The effect of spin-orbit and spin-other-orbit couplings on the orbital energy levels of uranyl is illustrated in Fig. S1 of the ESI.†

Table 3 shows the spin-free vertical excitation energies of the bare uranyl, uranyl with point charges (WFT-in-pc), uranyl with DFT embedding (WFT-in-DFT), and the supramolecular $\text{UO}_2\text{X}_4^{2-}$ systems.

In Table 3, we focus on the lowest-lying transition energies located on the uranyl fragment. Specifically, we analyze the excitation energies, including an electron transfer from the occupied σ_u and π_u orbitals to the virtual ϕ_u and δ_u orbitals. In the bare uranyl, the $5f_\phi$ and $5f_\delta$ of U do not participate in the U–O bonding and hence remain purely atomic.^{64,134,145} The lowest excitations are $\sigma_u \rightarrow \phi_u$ across the series. The following transitions are of $\sigma_u \rightarrow \delta_u$ nature. The ϕ_u and δ_u orbitals remain degenerate in the absence of ligands, but the δ_u orbitals are split in the presence of external fields. These are denoted as δ_u and δ'_u , respectively. The third transition is of the $\pi_u \rightarrow \delta_u/\pi_u \rightarrow \delta'_u$ character.

Table 3 Spin-free vertical excitation energies (in eV) of uranyl and uranyl tetrahalides calculated with EOM-pCCD-LCCSD. The WFT-in-pc and WFT-in-DFT denote the UO_2^{2+} embedded in point charges and static DFT embedding, respectively. The orbitals (except δ'_u) and molecular term symbols are labeled according to their characters in the bare uranyl. Environmental shifts from bare uranyl are shown in parenthesis

System	Character (Term)	UO_2^{2+}	WFT-in-pc	WFT-in-DFT	Supramolecule
$\text{UO}_2\text{F}_4^{2-}$	$\sigma_u \rightarrow \phi_u (^1\Phi_g)$	3.50	6.32 (2.82)	4.76 (1.26)	4.71 (1.21)
	$\sigma_u \rightarrow \delta_u (^1\Delta_g)$	3.89	4.30 (0.41)	4.38 (0.49)	4.24 (0.35)
	$\sigma_u \rightarrow \delta'_u (^1\Delta_g)$	3.89	4.99 (1.10)	4.62 (0.73)	4.52 (0.63)
	$\pi_u \rightarrow \delta_u (^1\Gamma_g)$	4.82	5.91 (1.09)	5.83 (1.01)	5.80 (0.98)
	$\pi_u \rightarrow \delta'_u (^1\Pi_g)$	4.94	6.54 (1.60)	6.10 (1.26)	6.06 (1.12)
$\text{UO}_2\text{Cl}_4^{2-}$	$\sigma_u \rightarrow \phi_u (^1\Phi_g)$	3.68	5.14 (1.46)	4.53 (0.85)	4.48 (0.80)
	$\sigma_u \rightarrow \delta_u (^1\Delta_g)$	4.05	4.43 (0.38)	4.46 (0.41)	4.33 (0.28)
	$\sigma_u \rightarrow \delta'_u (^1\Delta_g)$	4.05	4.82 (0.77)	4.57 (0.52)	4.46 (0.41)
	$\pi_u \rightarrow \delta_u (^1\Gamma_g)$	5.02	5.99 (0.97)	5.81 (0.79)	5.48 (0.46)
	$\pi_u \rightarrow \delta'_u (^1\Pi_g)$	5.13	6.39 (1.26)	5.97 (0.84)	5.60 (0.47)
$\text{UO}_2\text{Br}_4^{2-}$	$\sigma_u \rightarrow \phi_u (^1\Phi_g)$	3.69	5.04 (1.35)	4.55 (0.86)	4.24 (0.55)
	$\sigma_u \rightarrow \delta_u (^1\Delta_g)$	4.07	4.43 (0.36)	4.50 (0.43)	4.11 (0.04)
	$\sigma_u \rightarrow \delta'_u (^1\Delta_g)$	4.07	4.76 (0.69)	4.59 (0.52)	4.24 (0.17)
	$\pi_u \rightarrow \delta_u (^1\Gamma_g)$	5.05	5.97 (0.92)	5.85 (0.80)	4.76 (−0.29)
	$\pi_u \rightarrow \delta'_u (^1\Pi_g)$	5.16	6.32 (1.16)	6.00 (0.84)	4.84 (−0.32)

The U–O bond lengths in each structure are similar. Thus, the electronic excitation energies for uranyl shown in Table 3 are also alike. The presence of the equatorial fluorine, chlorine, and bromine ligands significantly perturbs the electronic spectra of uranyl. The most significant perturbation is observed for the $\text{UO}_2\text{F}_4^{2-}$ molecule, where the effect is as substantial as 1.2 eV. Moving down the halogen group, the impact of the environment decreases from $\text{UO}_2\text{F}_4^{2-}$ to $\text{UO}_2\text{Cl}_4^{2-}$ to $\text{UO}_2\text{Br}_4^{2-}$. A similar trend is observed in the EOM-CCSD excitation energies listed in Table S4 of the ESI†. The perturbation effect due to the presence of equatorial ligands is reasonably well predicted by both the point-charge and static embedding models. However, the WFT-in-DFT scheme provides more consistent energies with the reference supramolecular spectra across the $\text{UO}_2\text{X}_4^{2-}$ series and overall smaller errors (compare the numbers in parentheses in Table 3). The simplistic point-charge model shows a larger overestimation of the excitation energies compared to WFT-in-DFT for most of the transitions. The model is derived by replacing the halide atoms with charges, excluding any relaxation of the environment. Such a simple electrostatic picture is not suitable for modeling the ligand field. The best agreement between the WFT-in-DFT and supramolecular spectra is obtained for $\text{UO}_2\text{F}_4^{2-}$. The performance of WFT-in-DFT embedding drops down significantly for the $\pi_u \rightarrow \delta_u$ and $\pi_u \rightarrow \delta'_u$ electronic transitions in $\text{UO}_2\text{Cl}_4^{2-}$ and $\text{UO}_2\text{Br}_4^{2-}$. For heavier halides, these electronic transitions include some admixture of ligand-to-metal charge transfer character, which is challenging to mimic within the static WFT-in-DFT approach.

4.3.2 Orbital entanglement and correlation. Fig. 3–5 display the single orbital entropies and the decay of the orbital-pair mutual information for the orbitals involved in the lowest-lying excitation energies in the uranyl tetrahalide complexes and their simplified models. The individual values of the single-orbital entropies are collected in Table S4 of the ESI†

and depicted with the corresponding molecular orbital in Fig. S1–S3 of the ESI†. The single-orbital entropies in the left panels of Fig. 3–5 demonstrate that the approximate models perform well for the occupied σ_u and π_u orbitals, but deteriorate from the supramolecular data for the ϕ_u , δ_u , and δ'_u molecular orbitals. An exception is the $\text{UO}_2\text{F}_4^{2-}$ system, for which the point charge model performs erroneously. This behavior substantiates the overall bad performance of this model in predicting excitation energies as summarized in Table 3. Interestingly, all supramolecular uranyl tetrahalide structures predict an increased entanglement of ϕ_u and δ'_u orbitals, which are not captured by any of the approximate models. The strong entanglement of the δ'_u orbital in the supramolecular complex is to be expected as non-negligible interactions of that orbital with the environment have been noticed by other authors.^{97,142} At the same time, the corresponding δ_u orbital is significantly less entangled.

The right panel of Fig. 3–5 depicts the mutual information between selected (ground-state) orbitals involved in the lower-lying electronic transitions of the investigated excited states. The analysis of the I_{ij} values allows us to monitor how the embedded models can reproduce orbital correlations. Comparing the mutual information for various embedding schemes with the supramolecular spectrum further allows us to quantify the environmental effects on the system: the larger the deviations in the correlation spectrum between the bare and the supramolecular compounds, the stronger the environmental effects on the orbitals in question. Fig. 3–5 highlight that the WFT-in-DFT embedding schemes correct the orbital pair correlations compared to the bare uranyl and bring them closer to the supramolecular results. The selected I_{ij} spectrum suggests that WFT-in-DFT performs best for the $\text{UO}_2\text{Cl}_4^{2-}$ complex, followed by the $\text{UO}_2\text{F}_4^{2-}$ compound, and finally $\text{UO}_2\text{Br}_4^{2-}$. Our simple WFT-in-pc model is too restrictive to even qualitatively

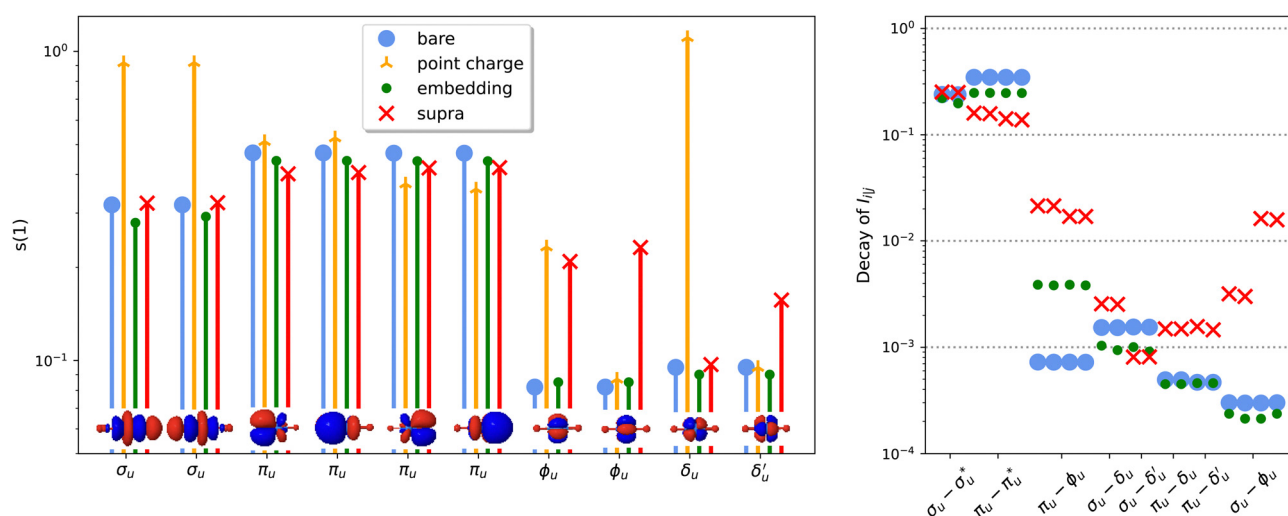


Fig. 3 Single orbital entropies (left panel) and decay of the mutual information (right panel) of selected orbitals in $\text{UO}_2\text{F}_4^{2-}$ and approximate models. The orbitals correspond to the bare uranyl and serve as a guide for the eye. Fig. S4 of the ESI† lists the orbitals involved in the excitation energies from different models.

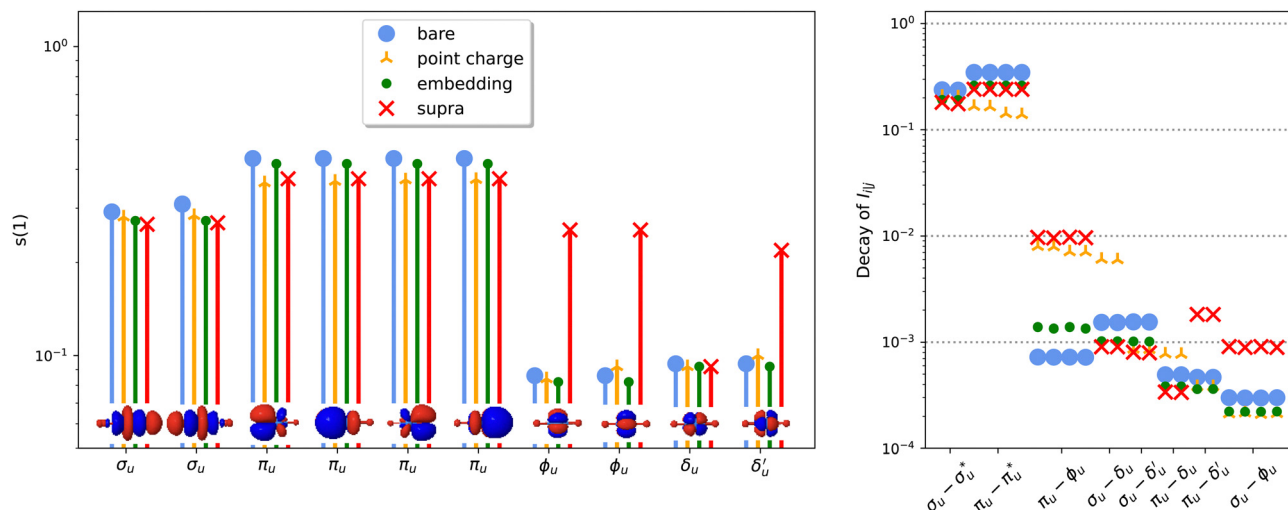


Fig. 4 Single orbital entropies (left panel) and decay of the mutual information (right panel) of selected orbitals in $\text{UO}_2\text{Cl}_4^{2-}$ and approximate models. The orbitals correspond to the bare uranyl and serve as a guide for the eye. Fig. S5 of the ESI† lists the orbitals involved in the excitation energies from different models.

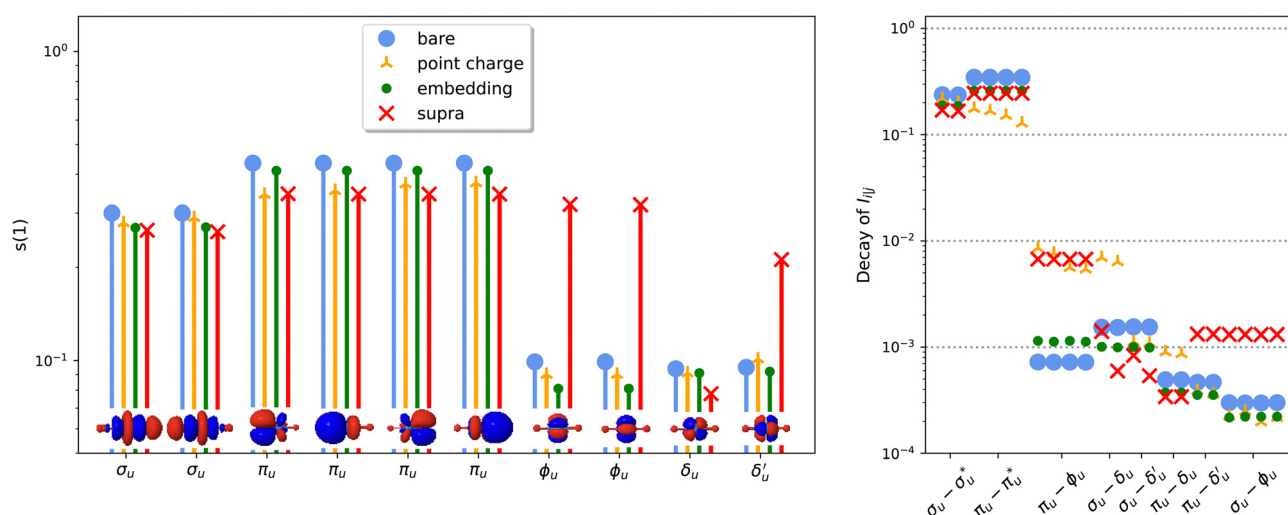


Fig. 5 Single orbital entropies (left panel) and decay of the mutual information (right panel) of selected orbitals in $\text{UO}_2\text{Br}_4^{2-}$ and approximate models. The orbitals correspond to the bare uranyl and serve as a guide for the eye. Fig. S6 of the ESI† lists the orbitals involved in the excitation energies from different models.

reproduce the (selected) orbital correlation spectra, which points to the limitations of the point-charge approach. Especially problematic for embedding models to correct are orbital correlations involving δ'_u and ϕ_u orbitals as their supramolecular counterparts feature correlations with environmental orbitals. The orbital entropy values and mutual information for the WFT-in-DFT method for all three species do not differ much, as we observe from Fig. 3–5. This indicates that the nature of the halide from which the static environment potential is obtained has little effect on the orbital interactions. That is another limitation of the orbital correlation analysis with our static embedding scheme, especially when we compare it with the supramolecular picture. This suggests the good performance of WFT-in-DFT in modeling excited states of $\text{UO}_2\text{F}_4^{2-}$ could be

just be a coincidence or originate from the cancellation of errors in the corresponding excited states. Note that our implementation allows us to analyze orbital correlations for electronic ground states. Finally, we should mention that the orbital pair correlations for the UO_2^{2+} -in-pc model of $\text{UO}_2\text{F}_4^{2-}$ are not shown in Fig. 3 due to numerical problems.

5. Conclusions

In this work, we have implemented a static WFT-in-DFT embedding scheme and a WFT-in-point-charges model, where pCCD-based methods approximate the WFT part. Such an embedding allows us to mimic the effect of the environment with the pCCD family of methods. We focused on the

EOM-pCCD-LCCSD-in-DFT approach to model electronic excitation energies in the water-ammonia complex and uranyl tetrahalides series of compounds. Our EOM-pCCD-LCCSD-in-DFT results are comparable to other WFT-in-DFT studies available in the literature but have the advantages of pCCD-based methods to model strong correlation⁵⁸ and provide an orbital correlation and entanglement analysis.^{78,80,146} The latter allows us to assess the quality of embedding models (point charges and static DFT embedding) with respect to the supramolecular reference. Our analysis of single-orbital entropies and the orbital-pair mutual information provides valuable information on orbital entanglement and correlation in approximate models compared to supramolecular structures. Specifically, we showed that the UO_2^{2+} -in-pc model does not provide a reliable description of orbital entanglement and correlation effects. This model is particularly erratic for the fluoride environment, reflecting the poor description of excited states. The UO_2^{2+} -in-DFT is less erratic but still has some problems with the accurate description of the uranyl ϕ_u and δ_u orbitals in terms of single-orbital entropies and mutual information. For heavier halides (Cl and Br), we also observe some ligand-to-metal-charge-transfer contributions to the $\pi_u \rightarrow \delta_u$ and $\pi_u \rightarrow \delta'_u$ transitions, which pose an additional challenge to embedding models.

To conclude, with the new embedding methods, we can extend the applicability of pCCD-based methods to larger systems, essential, for instance, in organic electronics, and perform a qualitative and quantitative analysis of embedding models in terms of orbital entanglement and correlations. These orbital entanglement and correlation measures can also be used to define appropriate fragments of supramolecular structures that minimize the quantum entanglement and correlation of the system and environment. To extend the applicability of our WFT-in-DFT models to larger systems, efficient implementations of the WFT-in-DFT approach are mandatory. This can be achieved, for instance, by accelerating and optimizing the WFT part. PyBEST's modular design allows for hosting external software libraries and enables us to improve bottleneck operations conveniently. For example, the Intel distribution for Python offers a thin layered interface for Intel MKL, providing out-of-the-box parallelization and acceleration of NumPy or SciPy routines. Another example is the efficient use of the CuPy library to enable graphics processing unit (GPU)-accelerated computing, where the bottleneck contractions in CC calculations are conveniently performed on a GPU.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

R. C. and P. T. acknowledge financial support from the OPUS research grant from the National Science Centre, Poland (Grant No. 2019/33/B/ST4/02114). P. T. acknowledges the scholarship

for outstanding young scientists from the Ministry of Science and Higher Education. K. B. acknowledges financial support from a SONATA BIS grant of the National Science Centre, Poland (no. 2015/18/E/ST4/00584). The authors thank Christoph Jacob for assistance with the PyADF software package. We have had many helpful discussions with Artur Nowak and Aleksandra Leszczyk.

Notes and references

- 1 P. G. Szalay, T. Müller, G. Gidofalvi, H. Lischka and R. Shepard, *Chem. Rev.*, 2012, **112**, 108–181.
- 2 R. J. Bartlett and M. Musiał, *Rev. Mod. Phys.*, 2007, **79**, 291–350.
- 3 A. S. P. Gomes and C. R. Jacob, *Annu. Rep. Prog. Chem., Sect. C*, 2012, **108**, 222.
- 4 G. Knizia and G. K.-L. Chan, *Phys. Rev. Lett.*, 2012, **109**, 186404.
- 5 K. Pernal, *Phys. Chem. Chem. Phys.*, 2016, **18**, 21111–21121.
- 6 P. Pulay, *Chem. Phys. Lett.*, 1983, **100**, 151–154.
- 7 C. Hampel and H.-J. Werner, *J. Chem. Phys.*, 1996, **104**, 6286–6297.
- 8 A. S. P. Gomes, C. R. Jacob and L. Visscher, *Phys. Chem. Chem. Phys.*, 2008, **10**, 5353–5362.
- 9 J. Neugebauer, *Chem. Phys. Chem.*, 2009, **10**, 3148–3173.
- 10 J. Neugebauer, *Phys. Rep.*, 2010, **489**, 1–87.
- 11 A. Solovyeva, M. Pavanello and J. Neugebauer, *J. Chem. Phys.*, 2014, **140**, 164103.
- 12 C. R. Jacob and J. Neugebauer, *WIREs Comput. Mol. Sci.*, 2014, **4**, 325–362.
- 13 W. Li, S. Li and Y. Jiang, *J. Phys. Chem. A*, 2007, **111**, 2193–2199.
- 14 E. E. Dahlke and D. G. Truhlar, *J. Chem. Theory Comput.*, 2007, **3**, 46–53.
- 15 S. R. Pruitt, M. A. Addicoat, M. A. Collins and M. S. Gordon, *Phys. Chem. Chem. Phys.*, 2012, **14**, 7752–7764.
- 16 E. Masoumifeshani and T. Korona, *Comput. Theory Chem.*, 2021, **1202**, 113303.
- 17 H. Lin and D. G. Truhlar, *Theor. Chem. Acc.*, 2007, **117**, 185–199.
- 18 H. M. Senn and W. Thiel, *Angew. Chem., Int. Ed.*, 2009, **48**, 1198–1229.
- 19 P. Cortona, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1991, **44**, 8454.
- 20 P. Cortona, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1992, **46**, 2008.
- 21 P. Cortona and A. V. Monteleone, *Int. J. Quantum Chem.*, 1994, **52**, 987–992.
- 22 J.-M. Gillet and P. Cortona, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1999, **60**, 8569.
- 23 G. Senatore and K. R. Subbaswamy, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1986, **34**, 5754.
- 24 M. Johnson, K. Subbaswamy and G. Senatore, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1987, **36**, 9202.

- 25 T. A. Wesolowski and A. Warshel, *J. Phys. Chem.*, 1993, **97**, 8050.
- 26 T. Wesolowski and A. Warshel, *J. Phys. Chem.*, 1994, **98**, 5183–5187.
- 27 T. A. Wesolowski and J. Weber, *Chem. Phys. Lett.*, 1996, **248**, 71.
- 28 M. E. Casida and T. A. Wesolowski, *Int. J. Quantum Chem.*, 2004, **96**, 577.
- 29 J. Neugebauer, *J. Chem. Phys.*, 2007, **126**, 134116.
- 30 J. Neugebauer, C. Curutchet, A. Munoz-Losa and B. Mennucci, *J. Chem. Theory Comput.*, 2010, **6**, 1843–1851.
- 31 T. A. Wesolowski, *Comput. Chem.*, 2006, **10**, 1–82.
- 32 M. Banafsheh and T. Adam Wesolowski, *Int. J. Quantum Chem.*, 2018, **118**, e25410.
- 33 T. A. Wesolowski, *J. Chem. Phys.*, 2022, **157**, 081102.
- 34 F. R. Manby, M. Stella, J. D. Goodpaster and T. F. Miller III, *J. Chem. Theory Comput.*, 2012, **8**, 2564–2568.
- 35 T. Culpitt, K. R. Brorsen and S. Hammes-Schiffer, *J. Chem. Phys.*, 2017, **146**, 211101.
- 36 N. Govind, Y. A. Wang, A. J. R. da Silva and E. Carter, *Chem. Phys. Lett.*, 1998, **295**, 129.
- 37 N. Govind, Y. A. Wang and E. A. Carter, *J. Chem. Phys.*, 1999, **110**, 7677.
- 38 P. Huang and E. A. Carter, *J. Chem. Phys.*, 2006, **125**, 084102.
- 39 T. A. Wesolowski, *Phys. Rev. A*, 2008, **77**, 012504.
- 40 C. Daday, C. König, J. Neugebauer and C. Filippi, *Comput. Phys. Commun.*, 2014, **15**, 3205–3217.
- 41 B. Hégyel, P. R. Nagy, G. G. Ferenczy and M. Kállay, *J. Chem. Phys.*, 2016, **145**, 064107.
- 42 B. Hégyel, P. R. Nagy and M. Kállay, *J. Chem. Theory Comput.*, 2018, **14**, 4600–4615.
- 43 S. Ghosh, P. Verma, C. J. Cramer, L. Gagliardi and D. G. Truhlar, *Chem. Rev.*, 2018, **118**, 7249–7292.
- 44 T. Klüner, N. Govind, Y. A. Wang and E. A. Carter, *Phys. Rev. Lett.*, 2001, **86**, 5954.
- 45 Y. Bouchafra, A. Shee, F. Réal, V. Vallet and A. S. P. Gomes, *Phys. Rev. Lett.*, 2018, **121**, 266001.
- 46 C. Huang and E. A. Carter, *J. Chem. Phys.*, 2011, **135**, 194104.
- 47 K. Yu, F. Libisch and E. A. Carter, *J. Chem. Phys.*, 2015, **143**, 102806.
- 48 J. D. Goodpaster, T. A. Barnes, F. R. Manby and T. F. Miller III, *J. Chem. Phys.*, 2012, **137**, 224113.
- 49 J. D. Goodpaster, T. A. Barnes, F. R. Manby and T. F. Miller III, *J. Chem. Phys.*, 2014, **140**, 18A507.
- 50 S. J. Bennie, M. W. van der Kamp, R. C. Pennifold, M. Stella, F. R. Manby and A. J. Mulholland, *J. Chem. Theory Comput.*, 2016, **12**, 2689–2697.
- 51 X. Zhang, S. J. Bennie, M. W. van der Kamp, D. R. Glowacki, F. R. Manby and A. J. Mulholland, *R. Soc. Open Sci.*, 2018, **5**, 171390.
- 52 Y. G. Khait and M. R. Hoffmann, *J. Chem. Phys.*, 2010, **133**, 044107.
- 53 S. J. Bennie, B. F. Curchod, F. R. Manby and D. R. Glowacki, *J. Phys. Chem. Lett.*, 2017, **8**, 5559–5565.
- 54 S. Ghosh, P. Verma, C. J. Cramer, L. Gagliardi and D. G. Truhlar, *Chem. Rev.*, 2018, **118**, 7249–7292.
- 55 R. S. Treß, J. Liu, C. Hättig and S. Höfener, *J. Chem. Phys.*, 2022, **157**, 204101.
- 56 P. Tecmer and K. Boguslawski, *Phys. Chem. Chem. Phys.*, 2022, **24**, 23026–23048.
- 57 P. A. Limacher, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck and P. Bultinck, *J. Chem. Theory Comput.*, 2013, **9**, 1394–1401.
- 58 K. Boguslawski, P. Tecmer, P. W. Ayers, P. Bultinck, S. De Baerdemacker and D. Van Neck, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2014, **89**, 201106(R).
- 59 T. Stein, T. M. Henderson and G. E. Scuseria, *J. Chem. Phys.*, 2014, **140**, 214113.
- 60 K. Boguslawski, P. Tecmer, P. A. Limacher, P. A. Johnson, P. W. Ayers, P. Bultinck, S. De Baerdemacker and D. Van Neck, *J. Chem. Phys.*, 2014, **140**, 214114.
- 61 P. A. Limacher, T. D. Kim, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck and P. Bultinck, *Mol. Phys.*, 2014, **112**, 853–862.
- 62 K. Boguslawski, P. Tecmer, P. W. Ayers, P. Bultinck, S. De Baerdemacker and D. Van Neck, *J. Chem. Theory Comput.*, 2014, **10**, 4873–4882.
- 63 P. Tecmer, K. Boguslawski, P. A. Limacher, P. A. Johnson, M. Chan, T. Verstraelen and P. W. Ayers, *J. Phys. Chem. A*, 2014, **118**, 9058–9068.
- 64 P. Tecmer, K. Boguslawski and P. W. Ayers, *Phys. Chem. Chem. Phys.*, 2015, **17**, 14427–14436.
- 65 A. J. Garza, A. G. S. Alencar and G. E. Scuseria, *J. Chem. Phys.*, 2015, **143**, 244106.
- 66 K. Boguslawski, A. Leszczyk, A. Nowak, F. Brzek, P. S. Żuchowski, D. Kędziera and P. Tecmer, *Comput. Phys. Commun.*, 2021, **264**, 107933.
- 67 A. Leszczyk, M. Máté, Ö. Legeza and K. Boguslawski, *J. Chem. Theory Comput.*, 2022, **18**, 96–117.
- 68 A. Leszczyk, T. Dome, P. Tecmer, D. Kędziera and K. Boguslawski, *Phys. Chem. Chem. Phys.*, 2022, **24**, 21296–21307.
- 69 F. Brzek, K. Boguslawski, P. Tecmer and P. S. Żuchowski, *J. Chem. Theory Comput.*, 2019, **15**, 4021–4035.
- 70 A. J. Garza, I. W. Bulik, T. M. Henderson and G. E. Scuseria, *Phys. Chem. Chem. Phys.*, 2015, **17**, 22412–22422.
- 71 K. Boguslawski, *J. Chem. Phys.*, 2016, **145**, 234105.
- 72 K. Boguslawski, *J. Chem. Theory Comput.*, 2019, **15**, 18–24.
- 73 P. Tecmer, K. Boguslawski, M. Borkowski, P. S. Żuchowski and D. Kędziera, *Int. J. Quantum Chem.*, 2019, **119**, e25983.
- 74 A. Nowak, P. Tecmer and K. Boguslawski, *Phys. Chem. Chem. Phys.*, 2019, **21**, 19039–19053.
- 75 S. Jahani, K. Boguslawski and P. Tecmer, arXiv, 2023, preprint, arXiv:2304.14547.
- 76 J. Rissler, R. M. Noack and S. R. White, *Chem. Phys.*, 2006, **323**, 519–531.
- 77 K. Boguslawski, P. Tecmer, O. Legeza and M. Reiher, *J. Phys. Chem. Lett.*, 2012, **3**, 3129–3135.
- 78 K. Boguslawski and P. Tecmer, *Int. J. Quantum Chem.*, 2017, **117**, e25455.

- 79 K. Boguslawski, P. Tecmer and Ö. Legeza, *Phys. Rev. B*, 2016, **94**, 155126.
- 80 A. Nowak, O. Legeza and K. Boguslawski, *J. Chem. Phys.*, 2021, **154**, 084111.
- 81 K. Boguslawski and P. W. Ayers, *J. Chem. Theory Comput.*, 2015, **11**, 5252–5261.
- 82 K. Boguslawski and P. Tecmer, *J. Chem. Theory Comput.*, 2017, **13**, 5966–5983.
- 83 D. J. Rowe, *Rev. Mod. Phys.*, 1968, **40**, 153–166.
- 84 H. Sekino and R. J. Bartlett, *Int. J. Quantum Chem.*, 1984, **18**, 255–265.
- 85 J. N. Byrd, V. Rishi, A. Perera and R. J. Bartlett, *J. Chem. Phys.*, 2015, **143**, 164103.
- 86 K. Boguslawski, *J. Chem. Phys.*, 2017, **147**, 139901.
- 87 K. Boguslawski, *Chem. Commun.*, 2021, **57**, 12277–12280.
- 88 M. De Santis, L. Belpassi, C. R. Jacob, A. Severo Pereira Gomes, F. Tarantelli, L. Visscher and L. Storchi, *J. Chem. Theory Comput.*, 2020, **16**, 5695–5711.
- 89 M. Sharma and M. Sierka, *J. Chem. Theory Comput.*, 2022, **18**, 6892–6904.
- 90 V. Kellö and A. J. Sadlej, *Int. J. Quantum Chem.*, 1998, **68**, 159–174.
- 91 K. G. Dyall and J. K. Fægri, *Introduction to Relativistic Quantum Chemistry*, Oxford, 2007.
- 92 K. Boguslawski, P. Tecmer, G. Barcza, O. Legeza and M. Reiher, *J. Chem. Theory Comput.*, 2013, **9**, 2959–2973.
- 93 C. Duperrouzel, P. Tecmer, K. Boguslawski, G. Barcza, O. Legeza and P. W. Ayers, *Chem. Phys. Lett.*, 2015, **621**, 160–164.
- 94 C. J. Stein and M. Reiher, *J. Chem. Theory Comput.*, 2016, **12**, 1760–1771.
- 95 T. Dunning Jr., *J. Chem. Phys.*, 1989, **90**, 1007–1023.
- 96 M. Etinski and C. M. Marian, *Phys. Chem. Chem. Phys.*, 2010, **12**, 4915–4923.
- 97 F. Ruipérez and U. Wahlgren, *J. Phys. Chem. A*, 2010, **114**, 3615.
- 98 G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. F. Guerra, S. J. A. van Gisbergen, J. G. Snijders and T. Ziegler, *J. Comput. Chem.*, 2001, **22**, 931–967.
- 99 2022, AMS2022, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, <https://www.scm.com>.
- 100 C. F. Guerra, J. G. Snijders, G. te Velde and E. J. Baerends, *Theor. Chem. Acc.*, 1998, **99**, 391–403.
- 101 C. R. Jacob, S. M. Beyhan, R. E. Bulo, A. S. P. Gomes, A. W. Götz, K. Kiewisch, J. Sikkema and L. Visscher, *J. Comput. Chem.*, 2011, **32**, 2328.
- 102 E. van Lenthe, E. J. Baerends and J. G. Snijders, *J. Chem. Phys.*, 1993, **99**, 4597–4610.
- 103 E. van Lenthe and E. J. Baerends, *J. Comput. Chem.*, 2003, **24**, 1142–1156.
- 104 K. Burke, J. P. Perdew and Y. Wang, *Electronic Density Functional Theory: recent progress and new directions*, 1998, pp. 81–111.
- 105 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.
- 106 A. Lembarki and H. Chermette, *Phys. Rev. A*, 1994, **50**, 5328.
- 107 P. Tecmer, R. Bast, K. Ruud and L. Visscher, *J. Phys. Chem. A*, 2012, **116**, 7397–7404.
- 108 A. W. Götz, S. M. Beyhan and L. Visscher, *J. Chem. Theory Comput.*, 2009, **5**, 3161–3174.
- 109 T. A. Wesolowski, *J. Chem. Phys.*, 1997, **106**, 8516–8526.
- 110 T. Wesolowski, Y. Ellinger and J. Weber, *J. Chem. Phys.*, 1998, **108**, 6078–6083.
- 111 P. Tecmer, H. van Lingen, A. S. P. Gomes and L. Visscher, *J. Chem. Phys.*, 2012, **137**, 084308.
- 112 A. S. P. Gomes, C. R. Jacob, F. Réal, L. Visscher and V. Vallet, *Phys. Chem. Chem. Phys.*, 2013, **15**, 15153–15162.
- 113 G. K. Liu, *J. Phys. Chem. A*, 2011, **115**, 12419.
- 114 H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby and M. Schütz, *WIREs Comput. Mol. Sci.*, 2012, **2**, 242–253.
- 115 H.-J. Werner, P. J. Knowles, F. R. Manby, J. A. Black, K. Doll, A. Heßelmann, D. Kats, A. Köhn, T. Korona and D. A. Kreplin, *et al.*, *J. Chem. Phys.*, 2020, **152**, 144107.
- 116 H.-J. Werner, P. J. Knowles *et al.*, MOLPRO, version 2020.2.1, a package of ab initio programs, 2020, see <https://www.molpro.net>.
- 117 2020, See <https://pybind11.readthedocs.io/en/stable/for-more-information-about-the-pybind11-project> (accessed June 23, 2023).
- 118 P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser and J. Bright, *et al.*, *Nat. Methods*, 2020, **17**, 261–272.
- 119 C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke and T. E. Oliphant, *Nature*, 2020, **585**, 357–362.
- 120 F. Aquilante, L. Boman, J. Boström, H. Koch, R. Lindh, A. S. de Merás and T. B. Pedersen, *Linear-Scaling Techniques in Computational Chemistry and Physics: Methods and Applications*, 2011, pp. 301–343.
- 121 J. Pipek and P. G. Mezey, *J. Chem. Phys.*, 1989, **90**, 4916–4926.
- 122 A. Leszczyk, T. Dome, P. Tecmer, D. Kedziera and K. Boguslawski, *Phys. Chem. Chem. Phys.*, 2022, **24**, 21296–21307.
- 123 A. Nowak, P. Tecmer and K. Boguslawski, *Phys. Chem. Chem. Phys.*, 2019, **21**, 19039–19053.
- 124 D. J. Rowe, *Rev. Mod. Phys.*, 1968, **40**, 153–166.
- 125 J. F. Stanton and R. J. Bartlett, *J. Chem. Phys.*, 1993, **98**, 7029–7039.
- 126 B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov and P.-O. Widmark, *J. Phys. Chem. A*, 2004, **108**, 2851–2858.
- 127 B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov and P.-O. Widmark, *Chem. Phys. Lett.*, 2005, **409**, 295–299.
- 128 N. Douglas and N. M. Kroll, *Ann. Phys.*, 1974, **82**, 89–155.
- 129 B. A. Hess, *Phys. Rev. A*, 1986, **33**, 3742–3748.
- 130 M. Reiher, *Theor. Chem. Acc.*, 2006, **116**, 241–252.
- 131 P. Tecmer, K. Boguslawski and D. Kedziera, *Handbook of computational chemistry*, 2017, vol. 2, pp. 885–926.

- 132 S. Höfener and L. Visscher, *J. Chem. Phys.*, 2012, **137**, 204120.
- 133 M. Busker, M. Nispel, T. Häber, K. Kleinermanns, M. Etinski and T. Fleig, *Chem. Phys. Chem.*, 2008, **9**, 1570–1577.
- 134 M. Pepper and B. E. Bursten, *Chem. Rev.*, 1991, **91**, 719–741.
- 135 G. Schreckenbach, P. J. Hay and R. L. Martin, *J. Comput. Chem.*, 1999, **20**, 70–90.
- 136 G. A. Shamov, G. Schreckenbach and T. N. Vo, *Chem. – Eur. J.*, 2007, **13**, 4932.
- 137 G. Schreckenbach and G. A. Shamov, *Acc. Chem. Res.*, 2010, **43**, 19–29.
- 138 M. Matsika and M. Pitzer, *J. Phys. Chem. A*, 2001, **105**, 637.
- 139 J. Shen and P. Piecuch, *J. Chem. Phys.*, 2013, **138**, 194102.
- 140 G. S. Groenewold, M. J. van Stipdonk, J. Oomens, W. A. de Jong, G. L. Gresham and M. E. McIlwain, *Int. J. Mass Spectrom.*, 2010, **297**, 67–75.
- 141 P. D. Dau, J. Su, H.-T. Liu, D.-L. Huang, J. Li and L.-S. Wang, *J. Chem. Phys.*, 2012, **137**, 064315.
- 142 G. Shen, M. Guo, Z. Wang and F. Wang, *Int. J. Quantum Chem.*, 2022, **122**, e26981.
- 143 W. A. de Jong, L. Visscher and W. C. Nieuwpoort, *J. Mol. Struct. THEOCHEM*, 1998, **458**, 41–52.
- 144 K. Dyall, *Mol. Phys.*, 1999, **96**, 511.
- 145 R. Denning, *J. Phys. Chem. A*, 2007, **111**, 4125–4143.
- 146 K. Boguslawski and P. Tecmer, *Int. J. Quantum Chem.*, 2015, **115**, 1289–1295.

Static Embedding with Pair Coupled Cluster Doubles Based Methods

Rahul Chakraborty, Katharina Boguslawski, and Paweł Tecmer

*Institute of Physics, Faculty of Physics, Astronomy and Informatics,
Nicolaus Copernicus University in Torun,
Grudziadzka 5, 87-100 Torun, Poland*

Supplementary Information

S1 XYZ Coordinates of $\text{H}_2\text{O} \cdots \text{NH}_3$ Complex

Table S1: XYZ Coordinates of the optimized $\text{H}_2\text{O} \cdots \text{NH}_3$ structure in Å.

N	-0.3247549	-1.0163481	0.2284577
H	0.3853716	-1.6985302	0.3923495
H	-0.9158405	-0.9874769	1.0326156
H	-0.8749286	-1.3247542	-0.5458969
O	0.3171256	1.9052994	-0.5156991
H	1.2106186	2.1227180	-0.3263054
H	0.2024080	0.9990920	-0.2655214

S2 EOM-CCSD vertical excitation energies of the H-bonded structures

TableS2: Vertical excitation energies (in eV) of $\text{H}_2\text{O} \cdots \text{NH}_3$ calculated with EOM-CCSD and aug-cc-pVDZ basis set. For each particular excitation, we show the VEE of the isolated source of excitation at the unrelaxed geometry of the complex.

	Fragment	supramolecule
	H_2O	NH_3
		6.42
		6.86
7.55		7.47
		7.96
		8.43
		7.96
		8.48

TableS3: The environmental shifts (ΔE_{env}) in vertical excitation energies (in eV) for the $\text{H}_2\text{O} \cdots \text{NH}_3$ complex calculated with EOM-pCCD-LCCSD and compared with EOM-CCSD values. ΔE_{env} is calculated as $(VEE_{\text{supra}} - VEE_{\text{frag}})$, where VEE_{supra} is the VEE of the supramolecule and VEE_{frag} is the same for the bare source of excitation at the unrelaxed geometry of the complex.

Source of excitation	EOM-pCCD-LCCSD	EOM-CCSD
NH_3	+0.43	+0.44
H_2O	-0.10	-0.08
NH_3	+0.46	+0.47
NH_3	+0.51	+0.52

Table S4: Vertical excitation energies (in eV) of thymine in microsolvation calculated with EOM-CCSD and aug-cc-pVDZ basis set. ΔE_{env} is calculated as $(VEE_{\text{supra}} - VEE_{\text{frag}})$, where VEE_{supra} is the VEE of the supramolecule and VEE_{frag} is the same for the bare thymine at the unrelaxed geometry of the complex.

Thymine	Thymine in microsolvation	ΔE_{env}
5.34	5.86	+0.52
5.87	6.30	+0.43
6.77	7.15	+0.38

S3 Spin-Orbit Coupling

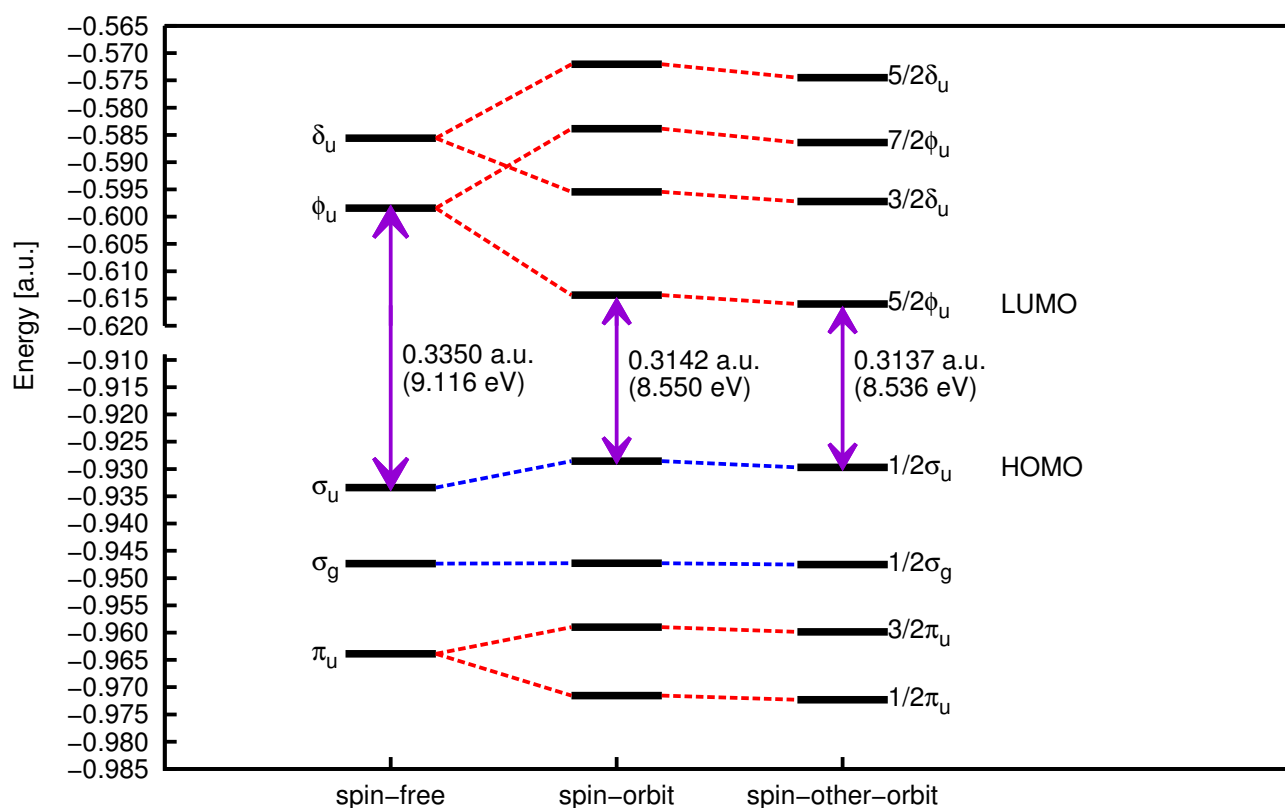


Figure S1: Comparison of Dirac-Hartree-Fock orbital energies of uranyl computed with different Hamiltonians: Dyall spin-free, (spin-orbit) Dirac-Coulomb, and Dirac-Coulomb-Breit (spin-other-orbit). The computations were performed with the Dirac2012 software package, aug-cc-pVTZ (O), and dyall.v3z (U) basis sets for U-O bond length of 1.708 Å.

S4 EOM-CCSD vertical excitation energies of uranyl tetrahalides

Table S5 shows the vertical excitation energies of the uranyl tetrahalides and the corresponding bare uranyl moieties. For the bare uranyl moieties, it is to be noted that the two δ_u orbitals are degenerate. We label them as δ_u and δ'_u to keep parity with the tetrahalides. We observe two transitions with both $\pi_u \rightarrow \delta_u$ and $\pi_u \rightarrow \delta'_u$ characters, in case of bare uranyls. For $\text{UO}_2\text{Br}_4^{2-}$, the $\pi_u \rightarrow \delta'_u$ transition is not observed with the 15 roots targeted in our EOM-CCSD calculation.

Table S5: EOM-CCSD/ANO-RCC-VDZP vertical excitation energies (in eV) of uranyl and its tetrahalides. The transitions are labelled according to their transition characters and not in the order of energies. The bare UO_2^{2+} calculations are done at each uranyl tetrahalide geometry, removing the halide atoms. The values in boldface in the 3rd column are the corresponding EOM-pCCD-LCCSD excitation energies of the bare uranyl, given here for comparison. Environmental shifts in EOM-CCSD excitation energies of uranyl due to the tetrahalide environment are given in the parenthesis in the 4th column.

System	Character (Term)	UO_2^{2+}	$\text{UO}_2\text{X}_4^{2-}$
$\text{UO}_2\text{F}_4^{2-}$	$\sigma_u \rightarrow \phi_u$ ($^1\Phi_g$)	3.02 (3.50)	4.22 (1.20)
	$\sigma_u \rightarrow \delta_u$ ($^1\Delta_g$)	3.39 (3.89)	3.89 (0.50)
	$\sigma_u \rightarrow \delta'_u$ ($^1\Delta_g$)	3.39 (3.89)	4.14 (0.75)
	$\pi_u \rightarrow \delta_u$ ($^1\Gamma_g$)	3.83 (4.82)	4.91 (1.08)
	$\pi_u \rightarrow \delta'_u$ ($^1\Pi_g$)	3.92 (4.94)	5.24 (1.32)
$\text{UO}_2\text{Cl}_4^{2-}$	$\sigma_u \rightarrow \phi_u$ ($^1\Phi_g$)	3.26 (3.68)	4.04 (0.78)
	$\sigma_u \rightarrow \delta_u$ ($^1\Delta_g$)	3.62 (4.05)	3.98 (0.36)
	$\sigma_u \rightarrow \delta'_u$ ($^1\Delta_g$)	3.62 (4.05)	4.08 (0.46)
	$\pi_u \rightarrow \delta_u$ ($^1\Gamma_g$)	4.16 (5.02)	5.01 (0.85)
	$\pi_u \rightarrow \delta'_u$ ($^1\Pi_g$)	4.25 (5.13)	5.08 (0.83)
$\text{UO}_2\text{Br}_4^{2-}$	$\sigma_u \rightarrow \phi_u$ ($^1\Phi_g$)	3.29 (3.69)	3.83 (0.54)
	$\sigma_u \rightarrow \delta_u$ ($^1\Delta_g$)	3.66 (4.07)	3.79 (0.13)
	$\sigma_u \rightarrow \delta'_u$ ($^1\Delta_g$)	3.66 (4.07)	3.87 (0.21)
	$\pi_u \rightarrow \delta_u$ ($^1\Gamma_g$)	4.21 (5.05)	4.44 (0.23)
	$\pi_u \rightarrow \delta'_u$ ($^1\Pi_g$)	4.30 (5.16)	—

S5 Orbital entropies and mutual information of uranyl tetrahalides

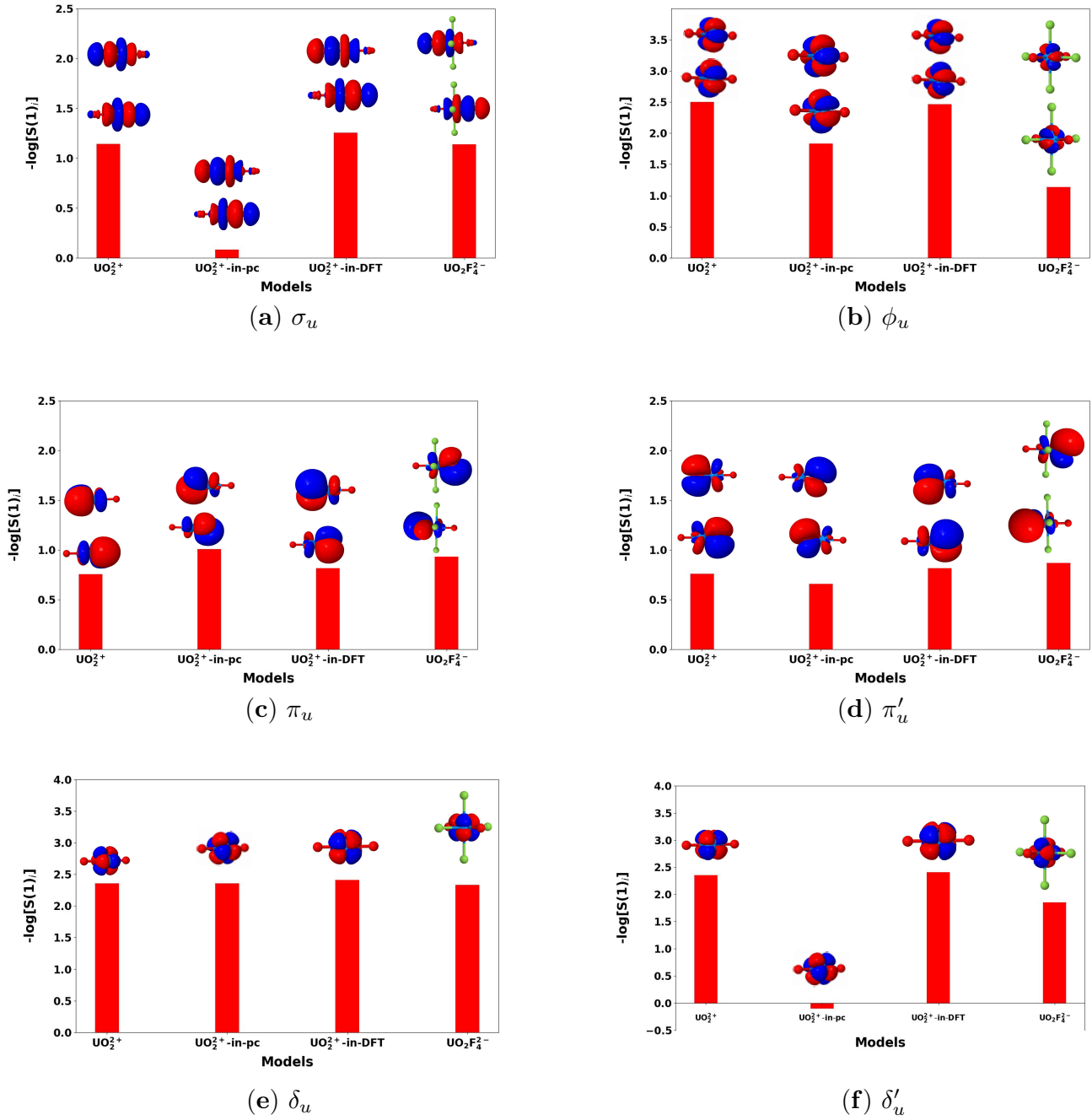
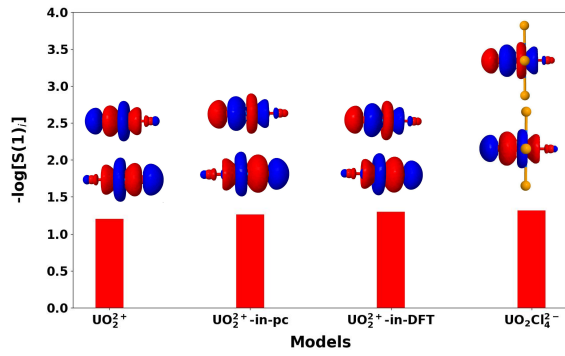
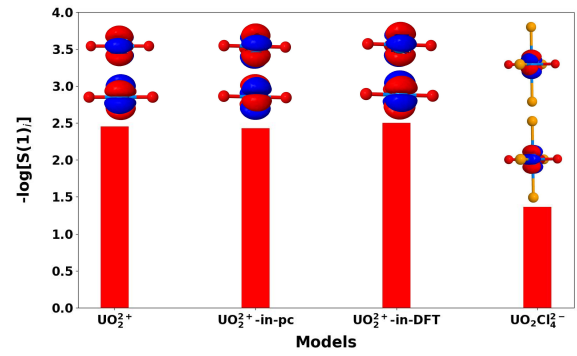
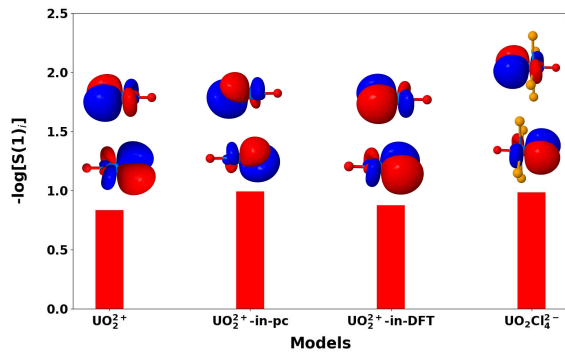
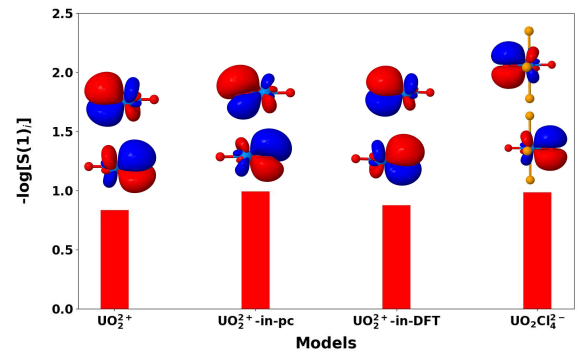
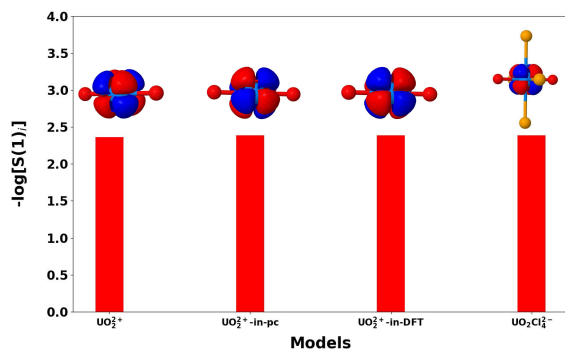
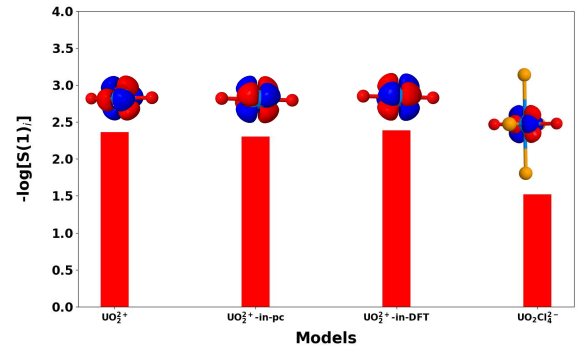


Figure S2: Single-orbital entropies of $\text{UO}_2\text{F}_4^{2-}$ and related models.

(a) σ_u (b) ϕ_u (c) π_u (d) π'_u (e) δ_u (f) δ'_u Figure S3: Single-orbital entropies of $\text{UO}_2\text{Cl}_4^{2-}$ and related models.

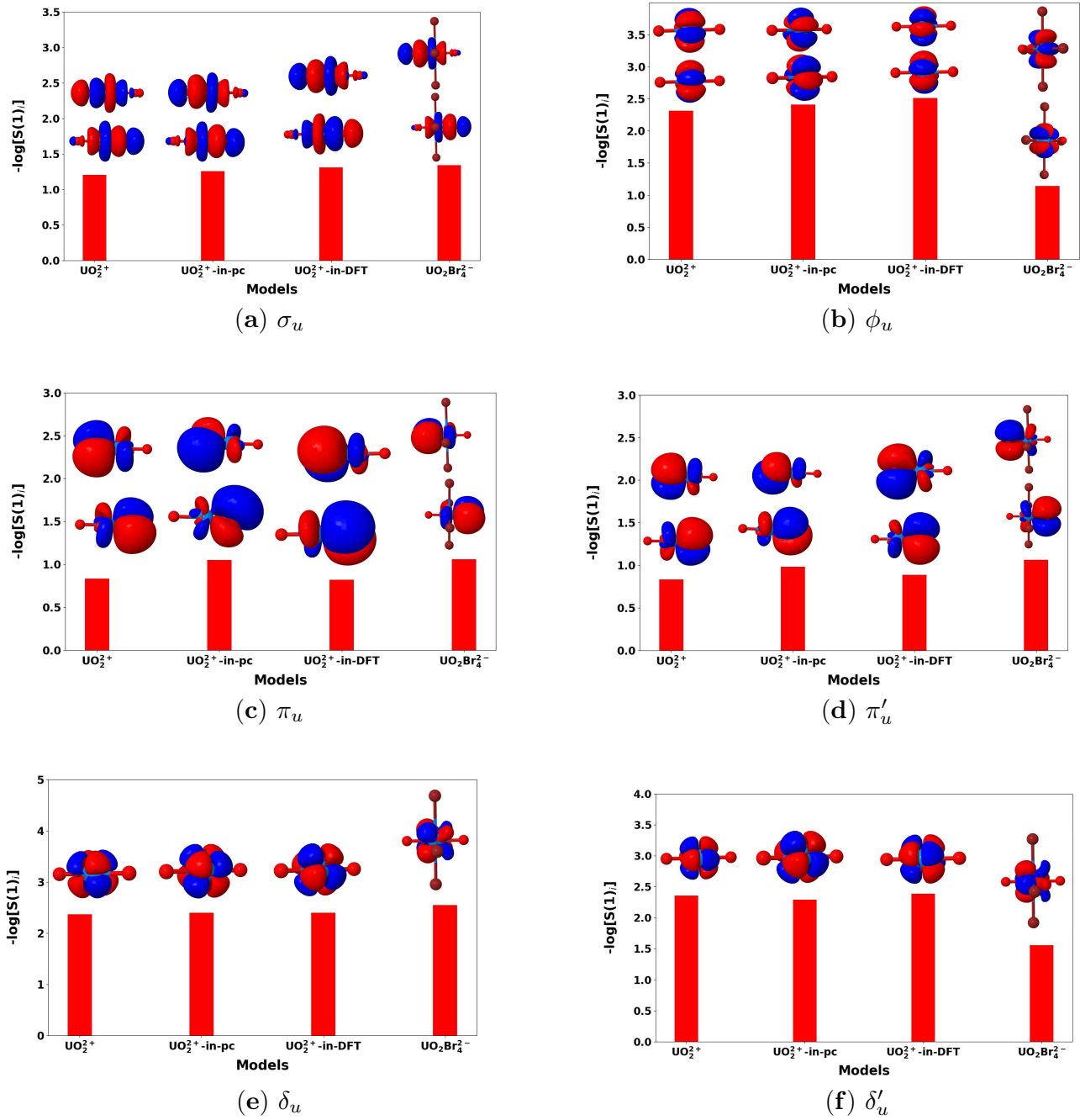


Figure S4: Single-orbital entropies of $\text{UO}_2\text{Br}_4^{2-}$ and related models.

Table S6: Single-orbital entropies ($s(1)_i$) of the uranyl orbitals involved in the electronic spectra of the different moieties. The $s(1)_i$ values are calculated with the pCCD-LCCSD method.

System	Orbital	UO ₂ ²⁺	WFT-in-pc	WFT-in-DFT	Supramolecule
UO ₂ F ₄ ²⁻	σ_u	0.319	0.920	0.279	0.323
		0.319	0.921	0.292	0.324
	π_u	0.470	0.373	0.443	0.371
		0.470	0.359	0.443	0.402
		0.469	0.512	0.442	0.420
		0.469	0.526	0.442	0.421
	ϕ_u	0.082	0.087	0.085	0.209
		0.082	0.233	0.085	0.232
	δ_u δ'_u	0.095	0.095	0.090	0.097
		0.095	1.11	0.090	0.157
UO ₂ Cl ₄ ²⁻	σ_u	0.292	0.281	0.273	0.266
		0.309	0.284	0.273	0.269
	π_u	0.434	0.362	0.417	0.373
		0.434	0.366	0.417	0.373
		0.434	0.370	0.417	0.373
		0.434	0.371	0.417	0.373
	ϕ_u	0.086	0.084	0.082	0.255
		0.086	0.092	0.082	0.255
	δ_u δ'_u	0.094	0.092	0.092	0.092
		0.094	0.100	0.092	0.219
UO ₂ Br ₄ ²⁻	σ_u	0.300	0.279	0.269	0.260
		0.300	0.290	0.270	0.264
	π_u	0.435	0.344	0.411	0.346
		0.435	0.354	0.411	0.346
		0.435	0.372	0.411	0.345
		0.435	0.375	0.411	0.345
	ϕ_u	0.099	0.090	0.081	0.319
		0.099	0.090	0.081	0.320
	δ_u δ'_u	0.094	0.091	0.091	0.078
		0.095	0.101	0.092	0.211

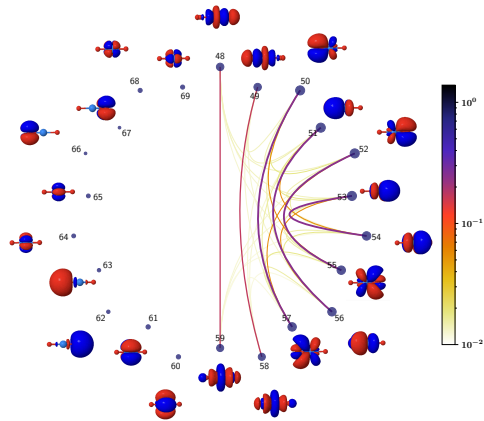
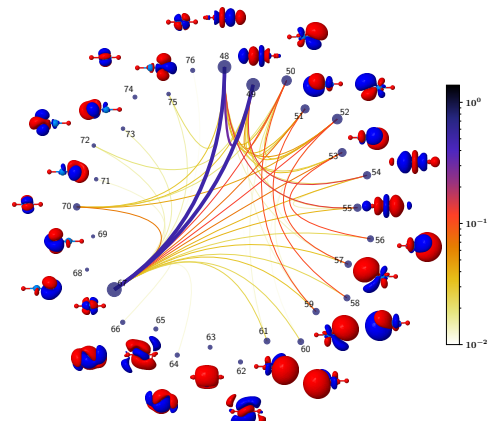
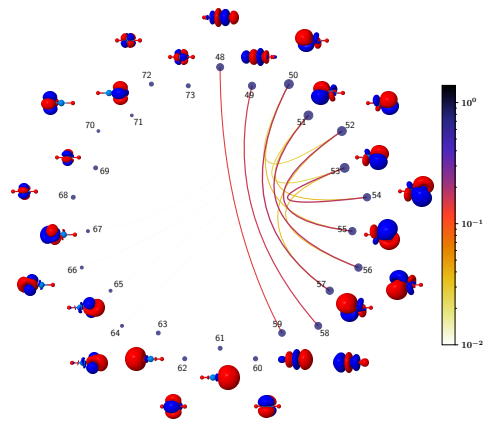
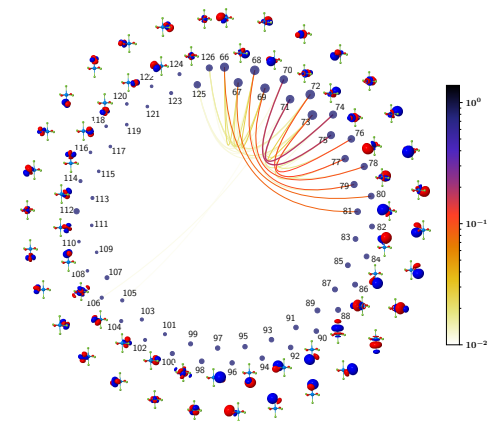
(a) UO_2^{2+} (b) $\text{UO}_2^{2+}\text{-in-pc}$ (c) $\text{UO}_2^{2+}\text{-in-DFT}$ (d) $\text{UO}_2\text{F}_4^{2-}$

Figure S5: Mutual information (I_{ij}) of $\text{UO}_2\text{F}_4^{2-}$ and related models studied in this work. The orbitals are numbered as obtained after pCCD optimization. The I_{ij} values are calculated based on single-orbital entropies calculated with the pCCD-LCCSD method.

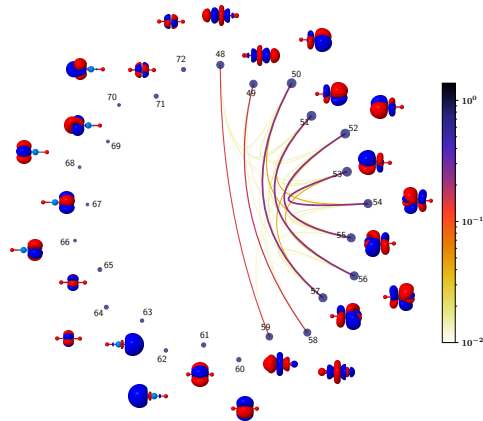
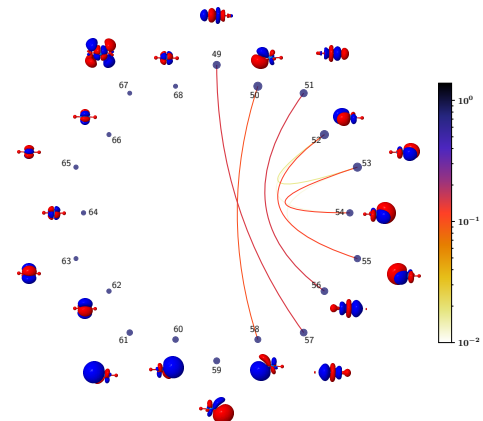
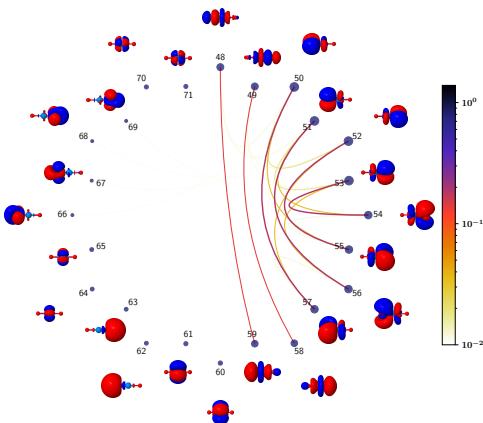
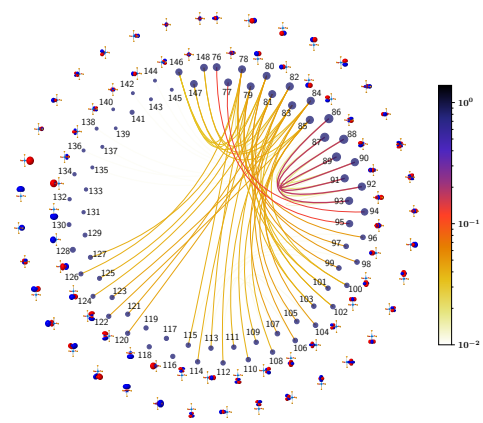
(a) UO_2^{2+} (b) UO_2^{2+} -in-pc(c) UO_2^{2+} -in-DFT(d) $\text{UO}_2\text{Cl}_4^{2-}$

Figure S6: Mutual information ($I_{i|j}$) of $\text{UO}_2\text{Cl}_4^{2-}$ and related models studied in this work. The orbitals are numbered as obtained after pCCD optimization. The $I_{i|j}$ values are calculated based on single-orbital entropies calculated with the pCCD-LCCSD method.

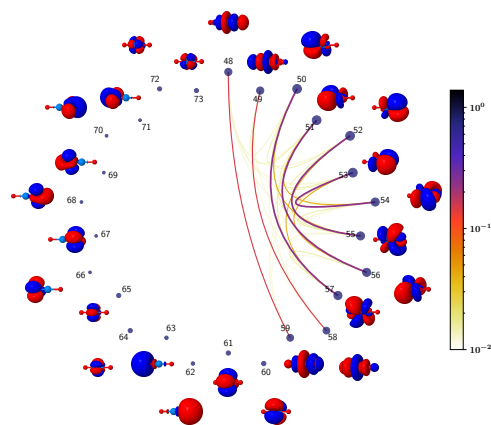
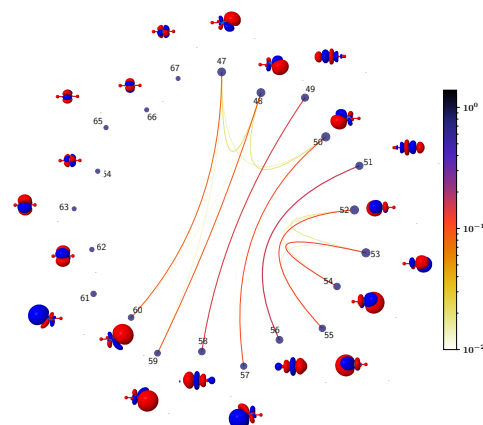
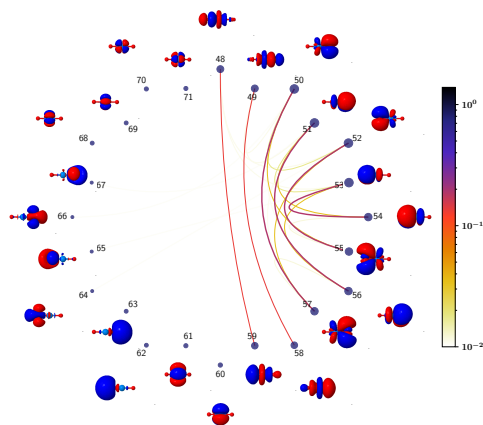
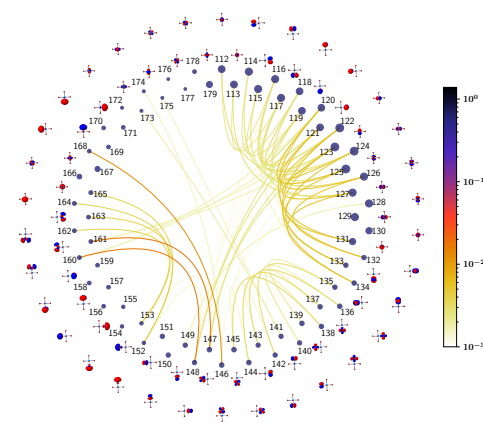
(a) UO_2^{2+} (b) $\text{UO}_2^{2+}\text{-in-pc}$ (c) $\text{UO}_2^{2+}\text{-in-DFT}$ (d) $\text{UO}_2\text{Br}_4^{2-}$

Figure S7: Mutual information (I_{ij}) of $\text{UO}_2\text{Br}_4^{2-}$ and related models studied in this work. The orbitals are numbered as obtained after pCCD optimization. The I_{ij} values are calculated based on single-orbital entropies calculated with the pCCD-LCCSD method.

Toward Reliable Dipole Moments without Single Excitations: The Role of Orbital Rotations and Dynamical Correlation

Rahul Chakraborty, Matheus Morato F. de Moraes,* Katharina Boguslawski,* Artur Nowak, Julian Świerczyński, and Paweł Tecmer*



Cite This: *J. Chem. Theory Comput.* 2024, 20, 4689–4702



Read Online

ACCESS |



Metrics & More

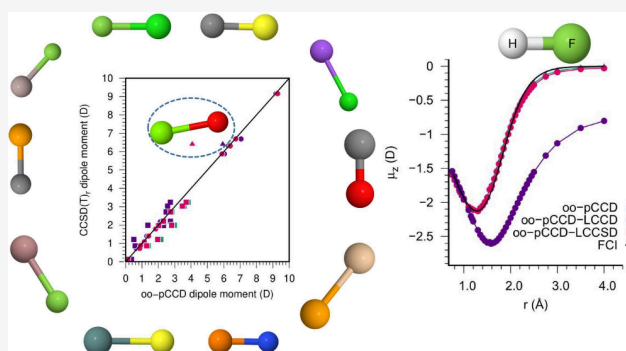


Article Recommendations



Supporting Information

ABSTRACT: The dipole moment is a crucial molecular property linked to a molecular system's bond polarity and overall electronic structure. To that end, the electronic dipole moment, which results from the electron density of a system, is often used to assess the accuracy and reliability of new electronic structure methods. This work analyses electronic dipole moments computed with the pair coupled cluster doubles (pCCD) ansätze and its linearized coupled cluster (pCCD-LCC) corrections using the canonical Hartree–Fock and pCCD-optimized (localized) orbital bases. The accuracy of pCCD-based dipole moments is assessed against experimental and CCSD(T) reference values using relaxed and unrelaxed density matrices and different basis set sizes. Our test set comprises molecules of various bonding patterns and electronic structures, exposing pCCD-based methods to a wide range of electron correlation effects. Additionally, we investigate the performance of pCCD-in-DFT dipole moments of some model complexes. Finally, our work indicates the importance of orbital relaxation in the pCCD model and shows the limitations of the linearized couple cluster corrections in predicting electronic dipole moments of multiple-bonded systems. Most importantly, pCCD with a linearized CCD correction can reproduce the dipole moment surfaces in singly bonded molecules, which are comparable to the multireference ones.



1. INTRODUCTION

The electric dipole moment is the major component of electrostatic interactions, which plays a significant role in many areas of chemistry, physics, and biology.¹ The electronic component of the molecular dipole moment contains many finer details about the electronic structure and bonding patterns in molecules² and contributes to interpreting spectroscopic data.^{3,4} Dipole moment surfaces, on the other hand, provide information about the change in bond polarity,⁵ intensities of the rovibrational transitions⁶. The reliable determination of this fundamental property is, thus, of preliminary importance for both experimental and theoretical domains. To that end, the quantum chemical modeling of electronic dipole moment provides a common testing ground for approximate wave function models.^{7–11} They can be compared with experimental results that are readily available for many small molecules. For example, the dipole moment was benchmarked against quantum chemical methods like Hartree–Fock theory, second-order Møller–Plesset (MP2) perturbation theory, coupled-cluster (CC) methods, multi-reference methods, and density functional theory (DFT) approximations.^{12,13} Specifically, coupled cluster-based ansätze have been extensively tested for dipole moment properties^{14,15} and remain an active research field.^{16,17} Maroulis and co-

workers^{18–24} performed numerous coupled cluster based studies, including the quantum chemistry gold standard—coupled cluster singles and doubles with perturbative triples (CCSD(T)), on electronic properties of different system types ranging from small di- and triatomic to organic molecules.²⁵ Studies by Mazziotti and co-workers^{26–28} have shown alternate routes for evaluation of electric properties using variational reduced density matrices. The elimination of the need for any reference wave function in this approach has great promise for determining electric property in systems with multireference characters. Ground and excited state dipole moments of full configuration interaction (FCI) quality can be reproduced with configuration interaction using a perturbative selection made iteratively (CIPSI) algorithm.^{29–31}

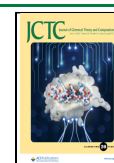
Although the electric dipole moment can be easily determined through density matrices, its sensitivity toward the accuracy of the electron density poses a real challenge to

Received: April 9, 2024

Revised: May 15, 2024

Accepted: May 17, 2024

Published: May 29, 2024



various quantum chemical methods.^{32–34} First, orbital relaxation has been shown to have a profound role in this regard.^{15,35} Second, some molecules require the inclusion of triple (or higher) excitations in the wave function expansion to obtain reliable dipole moments.^{36,37} The above aspects are the source of the well-known struggle approximate quantum chemistry methods face in an accurate description of the dipole moment of the CO molecule.^{5,38–43}

There are new families of geminal-based methods^{34,44–49} that are yet to be thoroughly tested for dipole moment properties. Some of the most promising ones are those based on the pCCD ansätze.^{50–53} They have seen recent successes in treating strongly correlated systems with mean-field-like scaling. pCCD has the feature of using its optimized orbital basis without defining active spaces.^{51,54–56} The size-extensive and size-consistent nature of orbital-optimized pCCD has motivated a wide range of studies for covalent molecules,^{54–63} noncovalent systems,^{64,65} and excited states,^{66–69} including organic systems.^{70,71} Perturbation theory-based, and linearized coupled cluster (LCC) corrections have also been successfully added to the pCCD wave function to improve the description of dynamic correlation.^{72–77}

To the best of our knowledge, little is known about the performance of the pCCD family of methods for ground-state electronic properties like dipole moments. However, there have been studies for such properties with the antisymmetric product of strongly orthogonal geminals (APSG)^{78,79} and other pair-coupled approximate wave function methods. The natural orbital functional theory formulated by Piris and co-workers (PNOFi, $i = 1, 6$) is noteworthy in this respect.^{80,81} Specifically, the PNOF5 is similar to the APSG approach⁸² and, thus, indirectly related to pCCD.⁵³ The coupled electron pair approximation(0) (CEPA(0)) and its orbital optimized variant⁸³ have similarities with the LCC approach. CEPA-based methods have been tested for dipole moments of various molecules.^{84–86}

This work aims to assess the performance of pCCD-type methods in quantifying the electric dipole moments of diatomics of the main group elements, and some larger complexes. The selected diatomic systems represent various bonding patterns (metal–nonmetal, nonmetal–nonmetal, metalloid–nonmetal, metal–metal van der Waals interaction). However, the pCCD framework restricts us to molecules with singlet ground states. Our work focuses on the effects of orbital optimization within pCCD and the inclusion of dynamic correlation. We use linearized coupled-cluster methods for the latter on top of the Hartree–Fock and pCCD wave function: doubles (pCCD-LCCD) and singles and doubles (pCCD-LCCSD) models.^{87,88} Furthermore, we probe the sensitivity of pCCD-based methods for dipole moments regarding different basis set sizes. We compare our electronic dipole moment values with the CCSD and CCSD(T) methods using relaxed and unrelaxed density matrices and experimental values. CCSD(T) has been well tested against dipole moments for a range of chemical specie like molecules of main group elements⁸⁹ and for transition metal compounds.⁹⁰ Specifically, in the large-scale benchmarking study by Liu et al., the average error for CCSD(T) dipole moment with respect to experimental values was found to be ≈ 0.15 D, showing even better performance for molecules with only main-group atoms. Finally, we extend our studies to pCCD-based static embedding calculations, where we obtain the embedding potential through the DFT approach (pCCD-in-DFT).^{91,92}

Precisely, we assess the performance of the pCCD-in-DFT embedding model for the electronic dipole moments of weakly hydrogen-bonded binary complexes such as CO–HF, CO–HCl, N₂–HF, N₂–HCl, and the H₂O⋯Rg [Rg = He, Ne, Ar, Kr] van der Waals complexes. The electronic structures of these complexes have been studied with various quantum chemical methods and thus represent a good reference point.^{93–96} Additionally, the weak interactions present in these molecules provide a good testing ground for the static embedding approach. In summary, this work reports the performance of some unique pCCD-based models (with and without orbital optimization) with and without dynamic energy corrections for dipole moment calculations.

2. THEORY

2.1. pCCD and Related Methods. Limiting the cluster operator to pair-excitations in the coupled cluster ansätze produces the pCCD ansätze,

$$|\Psi_{\text{pCCD}}\rangle = \exp\left(\sum_i^{\text{occ}} \sum_a^{\text{virt}} t_i^a \hat{a}_a^\dagger \hat{a}_i\right) |0\rangle = e^{\hat{T}_{\text{pCCD}}} |0\rangle \quad (1)$$

where $\hat{a}_p^\dagger$ and $\hat{a}_p$ ($\hat{a}_{\bar{p}}^\dagger$ and $\hat{a}_{\bar{p}}$) are the creation and annihilation operators for α -spin (and β -spin) electrons. $\hat{T}_{\text{pCCD}}$ is the pair-excitation cluster operator and $|0\rangle$ is a reference independent particle model, usually the Hartree–Fock wave function. The pCCD model misses a significant fraction of the dynamic electron correlation effects. In this work, we use a posteriori linearized coupled cluster⁸⁷ (LCC) corrections on top of the pCCD wave function to compensate for that. In the LCC correction, the exponential coupled cluster ansätze with a pCCD reference wave function is used as

$$|\Psi\rangle = \exp(\hat{T}') |\Psi_{\text{pCCD}}\rangle \quad (2)$$

where $\hat{T}' = \sum_\nu t_\nu \hat{\tau}_\nu$ is a cluster operator containing excitation operators $\hat{\tau}_\nu$ of various levels. The "''" in the cluster operator indicates that the pair excitations present in pCCD are excluded. The corresponding energy equation is

$$\hat{H} \exp(\hat{T}') |\Psi_{\text{pCCD}}\rangle = E \exp(\hat{T}') |\Psi_{\text{pCCD}}\rangle \quad (3)$$

where $\hat{H}$ is the electronic Hamiltonian of the system. In the LCC framework, the associated Baker–Campbell–Hausdorff expansion is restricted to the second term, i.e.,

$$(\hat{H} + [\hat{H}, \hat{T}']) |\Psi_{\text{pCCD}}\rangle = E |\Psi_{\text{pCCD}}\rangle \quad (4)$$

When we include both single and double excitations (for the pCCD-LCCSD model), $\hat{T}'$ reads,

$$\hat{T}' = \hat{T}_1 + \hat{T}_2 = \sum_i^{\text{occ}} \sum_a^{\text{virt}} t_i^a \hat{E}_{ai} + \frac{1}{2} \sum_{i,j}^{\text{occ}} \sum_{a,b}^{\text{virt}} t_{ij}^{ab} \hat{E}_{ai} \hat{E}_{bj} \quad (5)$$

where

$$\hat{E}_{ai} = \hat{a}_a^\dagger \hat{a}_i + \hat{a}_{\bar{a}}^\dagger \hat{a}_{\bar{i}} \quad (6)$$

is the singlet excitation operator. Note that the "''" in the second sum of the above equation excludes the cases where $i = j$ and simultaneously $a = b$, while terms where $i = j \wedge a \neq b$ or $i \neq j \wedge a = b$ are still included. Elimination of $\hat{T}_1$ amplitudes from $\hat{T}'$ in eq 5 leads to the pCCD-LCCD model. Both pCCD-LCC

variants have been successfully used for various molecules, providing a moderate balance between dynamic and non-dynamic electron correlation effects.^{62,64,75,87,97}

2.2. Density Matrices from pCCD and Related Methods. Elements of the 1-electron reduced density matrix (1-RDM) obtained from any wave function Ψ can be expressed as

$$\gamma_q^p = \langle \Psi | \hat{a}_p^\dagger \hat{a}_q | \Psi \rangle \quad (7)$$

For truncated CC models, the 1-electron molecular response properties are calculated using the derivative approach as a response to a small external perturbation related to the property in question (such as dipole moments). In this approach, the response density matrices are often used.^{98–101} Accordingly, the elements of the pCCD response 1-RDM are defined as

$${}^{\text{pCCD}}\gamma_q^p = \langle 0 | (1 + \Lambda_{\text{pCCD}}) e^{-\hat{T}_{\text{pCCD}}} \hat{a}_p^\dagger \hat{a}_q e^{\hat{T}_{\text{pCCD}}} | 0 \rangle \quad (8)$$

where $\Lambda_{\text{pCCD}} = \sum_{ia} \lambda_a^i \hat{a}_i^\dagger \hat{a}_a^\dagger \hat{a}_a \hat{a}_i$ is the electron-pair de-excitation operator.

On the other hand, the response 1-RDM from the pCCD-LCC wave functions can be constructed using the reference response 1-RDM of pCCD from eq 8 and the correlation contribution of the LCC correction on top of the pCCD wave function calculated using the so-called Λ -equations,⁸⁸

$${}^{\text{LCC}}\gamma_q^p = \langle 0 | (1 + \Lambda'_{\text{LCC}}) \{ e^{-\hat{T}' - \hat{T}_{\text{pCCD}}} \{ \hat{a}_p^\dagger \hat{a}_q \} e^{\hat{T}_{\text{pCCD}} + \hat{T}'} \} | 0 \rangle \quad (9)$$

where $\Lambda'_{\text{LCC}} = \Lambda'_1 + \Lambda'_2$ or $\Lambda'_{\text{LCC}} = \Lambda'_2$, for the LCCSD and LCCD models respectively, and

$$\Lambda'_n = \frac{1}{(n!)^2} \sum_{ij \dots} \sum_{ab \dots} \lambda_{ab \dots}^{ij \dots} \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_a \hat{a}_b \dots \quad (10)$$

is the de-excitation operator, where all electron-pair de-excitation are to be excluded as they do not enter the LCC equations (again, indicated by the "..."). For the LCC response density matrices, only terms that are at most linear in $\hat{T}'_1$ and $\hat{T}'_2$ are to be considered. This is indicated by $\{\dots\}_L$ in Eq. 9. The final 1-RDM from oo-pCCD-LCC(S)D approaches is the sum of the relaxed oo-pCCD and unrelaxed LCC(S)D contributions,

$$\gamma_q^p = {}^{\text{pCCD}}\gamma_q^p + {}^{\text{LCC}}\gamma_q^p \quad (11)$$

As evident from eq 8, ${}^{\text{pCCD}}\gamma_q^p$ contains both the contribution from the reference determinant and the pCCD-correlation part, ${}^{\text{pCCD}}\gamma_q^p = {}^{\text{ref}}\gamma_q^p + {}^{\text{corr(pCCD)}}\gamma_q^p$, while ${}^{\text{LCC}}\gamma_q^p$ accounts for the LCC correlation part only. It is to be noted that orbital relaxation due to the LCC correction is not considered in this work.

2.3. Dipole Moment Calculation. The total dipole moment of a molecule is defined as

$$\mu_\alpha = \sum_{i=1}^{N_{\text{nuc}}} Z_i \mathbf{R}_{i\alpha} - \int \rho(r) \mathbf{r}_\alpha dr \quad (12)$$

where the first term accounts for nuclear and the second for electronic contributions. In eq 12, α denotes the axial direction (x, y or z), Z_i charge of the i -th nucleus, N_{nuc} the number of

nuclei in the molecular structure, and $\mathbf{R}$ and $\mathbf{r}$ correspond to the nuclear and electronic coordinates, respectively.

After introducing an atomic orbital (AO) basis set, one α -component of the dipole moment is evaluated from

$$\mu_\alpha = \sum_{i=1}^{N_{\text{nuc}}} Z_i \mathbf{R}_{i\alpha} - \sum_{\mu} \sum_{\nu} \gamma_{\nu}^{\mu} \langle \nu | \hat{\mathbf{r}}_\alpha | \mu \rangle \quad (13)$$

where γ_{ν}^{μ} is the density matrix in the AO and $\langle \nu | \hat{\mathbf{r}}_\alpha | \mu \rangle = \int \chi_{\nu}^*(r) \mathbf{r}_\alpha \chi_{\mu}(r) dr$ are the dipole moment integrals expressed in the AO basis $\{\chi_{\nu}\}$.¹⁰² Since all pCCD-based methods work in the molecular orbital (MO) basis and hence the corresponding 1-RDMs are defined for the molecular orbitals, we need to perform an AO-MO transformation step of the dipole moment integrals or the 1-RDMs, respectively.

3. COMPUTATIONAL DETAILS

3.1. Structures. The geometries of the main group diatomic molecules were taken from Liu et al.¹⁰³ and references therein. Their bond lengths are collected in Table S1 of the SI. Each diatomic molecule is placed along the z-axis.

The structures of the CO–HF, CO–HCl, N₂–HF, and N₂–HCl^{93,104,105} were optimized with the CCSD(T) method and the augmented Dunning-type correlation consistent basis sets of quadruple- ζ quality (aug-cc-pVQZ).^{106,107} The molecules were placed along the z-axis, as shown in Figure 1a along with the optimized bond lengths.

The bond parameters of the H₂O...Rg [Rg = He, Ne, Ar, Kr] complexes were taken from Haskopoulos et al.⁹⁶ Following the original work, these complexes were kept in the xz plane with the center of mass of H₂O at the origin and the oxygen atom on the negative z-axis (see Figure 1b). The equilibrium bond

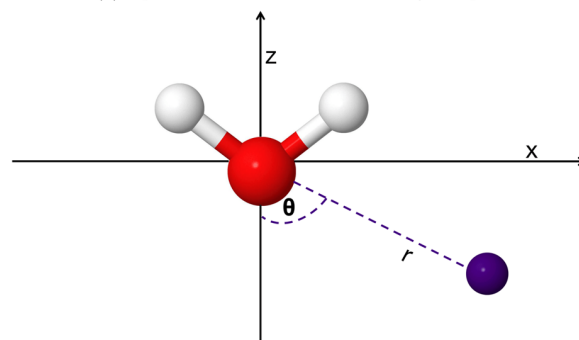
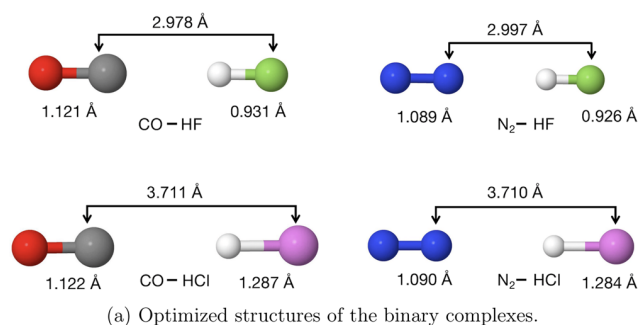


Figure 1. Structural representations of the complexes studied in this work.

Table 1. Error Analysis for the Dataset of 20 Main Group Diatomics Studied in This Work^a

Method	Exp.		CCSD(T) _r									
			CCSD(T) _u		full data set		singly bonded		multiply bonded		w/o MgO	
	MUE	RMSE	MUE	RMSE	MUE	RMSE	MUE	RMSE	MUE	RMSE	MUE	RMSE
pCCD	0.437	0.640	0.471	0.782	0.393	0.585	0.131	0.155	0.577	0.807	0.437	0.512
pCCD-LCCD	0.356	0.535	0.365	0.663	0.288	0.460	0.091	0.109	0.429	0.638	0.309	0.374
pCCD-LCCSD	0.530	0.753	0.382	0.547	0.465	0.730	0.105	0.138	0.754	1.024	0.585	0.672
CCSD _u	0.180	0.276	0.136	0.307	0.063	0.102	0.032	0.061	0.086	0.131	0.059	0.068
CCSD(T) _u	0.194	0.300	—	—	0.087	0.217	0.011	0.016	0.149	0.306	0.070	0.081
oo-pCCD	0.363	0.604	0.323	0.525	0.336	0.637	0.078	0.097	0.550	0.896	0.375	0.526
oo-pCCD-LCCD	0.345	0.581	0.284	0.474	0.302	0.605	0.068	0.081	0.493	0.852	0.309	0.429
oo-pCCD-LCCSD	0.373	0.476	0.248	0.307	0.283	0.352	0.092	0.136	0.399	0.465	0.393	0.441
CCSD _r	0.191	0.262	0.169	0.289	0.085	0.119	0.038	0.062	0.122	0.156	0.116	0.144

^aErrors are calculated in Debye using the aug-cc-pVTZ basis for all methods. MUE and RMSE stand for mean unsigned error $\left[\frac{1}{N} \sum_i |\mu_{\text{Method},i} - \mu_{\text{ref},i}| \right]$ and root mean square error $\left[\sqrt{\left(\sum_i (\mu_{\text{Method},i} - \mu_{\text{ref},i})^2 / N \right)} \right]$, respectively, where N is the number of molecules in the dataset. For CCSD(T)_r, reference data, MUE and RMSE are divided for the full data set, all singly-bonded, and multiply-bonded systems (with and without the outlier MgO).

parameters of these 4 complexes are given in Table S8 of the SI.

3.2. pCCD-Based Dipole Moment. The pCCD-based dipole moment calculations were carried out in a developer version (v1.4.0dev) of the PYBEST software package.^{61,108,109} The dipole moments were calculated with the Dunning family of basis sets with and without augmentation, that is, (aug-)cc-pVnZ, for $n = D, T$ and Q with optimized general contractions.^{106,107,110,111} Henceforth, the orbital optimized pCCD and the LCC corrections on top of it are called oo-pCCD and oo-pCCD-LCC(S)D. Consequently, the pCCD and pCCD-LCC(S)D will refer to pCCD and a posteriori LCC corrections within a canonical (Hartree–Fock (HF)) orbital basis.

Cholesky decomposition of the two-electron repulsion integrals with a threshold of 10^{-5} was used for all systems. Pipek–Mezey orbital localization¹¹² was used to speed up the orbital optimization process for all systems. In all pCCD and oo-pCCD based calculations, all nonvalence orbitals were kept frozen to match the MOLPRO reference results (vide infra).

3.2.1. pCCD-in-DFT. The embedding potentials were generated within the Amsterdam Modeling Suite (AMS2022)^{113–115} and then extracted with the help of the PyADF¹¹⁶ scripting framework. In all DFT-in-DFT calculations, the triple- ζ double polarization (TZ2P) basis set,¹¹⁷ the PW91^{118,119} exchange–correlation functional, and the PW91k¹²⁰ kinetic energy functional were used. More details about the DFT-in-DFT frozen density embedding (FDE) setup used here to obtain the embedding potential are described in our previous work.⁹² For each embedding calculation, two sets of calculations were performed, in which the system and environment were swapped, and their dipole moment results were added together.

3.3. Reference Dipole Moment Calculations. All reference values were obtained using the MOLPRO package version 19.^{121–123} The reference dipole moments were obtained using the CCSD and CCSD(T) methods^{124–127} (relaxed and unrelaxed density matrices) and the same family of basis sets used in pCCD and oo-pCCD based calculations with PyBEST. In this work, CCSD_u and CCSD(T)_u refer to dipole moments with unrelaxed densities, whereas CCSD_r and CCSD(T)_r are for the same with relaxed densities. The CCSD

and CCSD(T) dipole moments are calculated with the CC response formalism as implemented in the MOLPRO software package.

4. RESULTS AND DISCUSSION

4.1. Dipole Moment of Main Group Diatomics.

4.1.1. Statistical Analysis. We start our analysis with the diatomic molecules and the basis set dependence. Table S2 of the SI collects all the dipole moments computed with different quantum chemistry methods and (aug-)cc-pVnZ [$n = D, T, Q$] basis sets with and without augmented functions. All basis sets provide qualitatively similar results. The most significant differences are observed between the cc-pVDZ and cc-pVTZ basis sets, and between the standard and augmented series. The differences within the augmented series are significantly smaller. Table S3 collects the mean unsigned errors (MUE) and root-mean-square errors (RMSE) for all the methods considered in this work in all basis sets with respect to experimental dipole moments. We observe that triple- ζ and quadruple- ζ basis sets produce similar errors. MUE and RMSE increase slightly from aug-cc-pVTZ to aug-cc-pVQZ for oo-pCCD and oo-pCCD-LCCD. However, the opposite is seen for oo-pCCD-LCCSD. In short, not much accuracy is gained by increasing the size of the basis set from triple- ζ to quadruple- ζ in terms of dipole moments, as has been observed in previous works with traditional coupled cluster methods.¹² To that end, we used the aug-cc-pVTZ as the basis set of choice for further investigations. In addition, we should stress that the dipole moment results are more or less independent of the frozen core approximation (cf. Table S4 of the SI).

Table 1 summarizes the MUE and the RMSE of our pCCD-based methods with respect to the experimental data and the reference theoretical CCSD(T)_r and CCSD(T)_u values. The data from Table 1 shows that, on average, the orbital optimization within the pCCD reference function improves the overall performance of the pCCD-based dipole moments with respect to experiment and reference theoretical data. Including LCC on top of pCCD further refines the dipole moment values toward the reference. From a numerical perspective, the MUEs for pCCD and pCCD-LCCSD improve by ≈ 0.1 D upon the addition of orbital optimization. However,

pCCD-LCCD statistics do not show much improvement with the same.

In Figure 2, we show the percentage errors (with sign) in dipole moments obtained with pCCD-based methods for

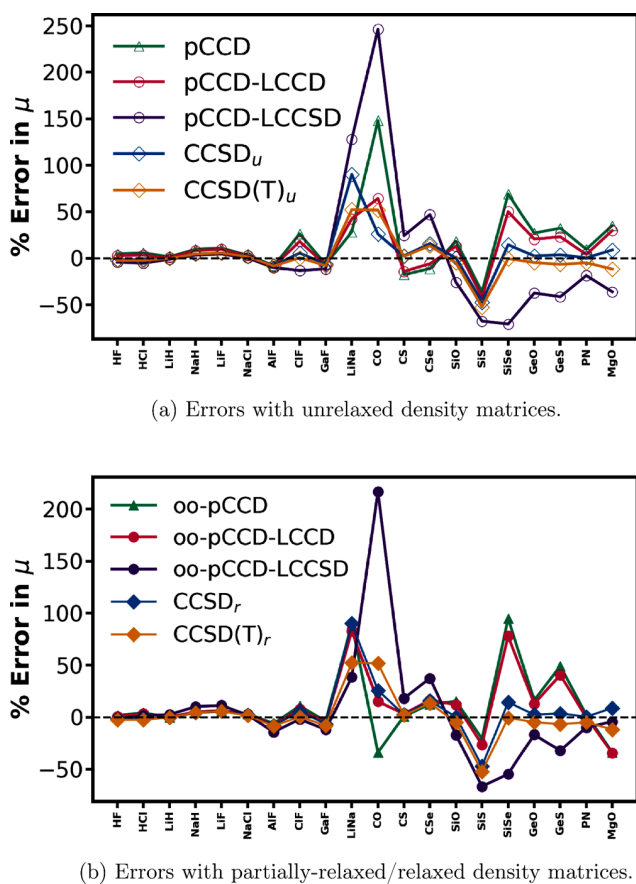


Figure 2. Percentage errors in all methods using aug-cc-pVTZ basis with respect to the experimental dipole moment values for all molecules in the data set.

individual molecules, with respect to the experimental values. Figure 2a shows the performance of pCCD and its variants without orbital optimization, i.e., with completely unrelaxed densities, whereas Figure 2b depicts the same for oo-pCCD and subsequent LCC variants, with relaxed densities achieved through orbital optimization within pCCD. Here, it is important to remember that the oo-pCCD-LCC density matrices are only partially relaxed. In this plot, we see a clear distinction between the behavior of simple singly bonded molecules and the molecules with significant multiple-bond characters. As evident from Figure 2b, the second class of molecules shows higher relative errors with all pCCD-based methods. We also observe that LCCD values remain in close vicinities of the pCCD ones for most of the molecules. Exceptions to this occur for molecules, again, with multiple bond characters (see also last columns in Table 1). The LCCSD values, on the other hand, differ significantly from their counterparts for almost all molecules. Of particular interest is the MgO molecule, where the oo-pCCD-LCCSD seems to perform even better than CCSD(T)_r with respect to the experiment. The impact of the character of the bond on the dipole moment values obtained with pCCD-based methods is also evident in the violin plots in Figure 3. Specifically, Figure

3a and Figure 3b show the distribution (skewness) of the errors in dipole moments with pCCD-based methods with respect to CCSD(T)_r and experimental values, respectively. As can be seen, the multiply bonded molecules show a significantly higher spread of errors than the singly bonded molecules. For the latter, the interquartile ranges are distributed closely around the median. If the orbitals are optimized within pCCD, the median and spread are shifted closer to the reference. Moreover, an LCCSD correction introduces outliers and features a broader interquartile range. For multiply bonded systems, the skewness of errors is right-shifted for (oo-)pCCD and (oo-)pCCD-LCCD, while (oo-)pCCD-LCCSD yields left-shifted ones. Furthermore, (oo-)pCCD-LCCD reduces the interquartile range and shifts the median closer to the reference, while (oo-)pCCD-LCCSD introduces a strong asymmetry, moving the median below the reference point.

Overall, though our statistical analysis shows the utility of adding dynamic correlation with LCC corrections in the pCCD framework, a case-by-case analysis reveals that this is not a black-box tool for all molecules regarding the calculation of dipole moments. That motivates us to conduct a deeper analysis of the performance of pCCD-based methods for different types of molecules and bonding patterns in the next section.

4.1.2. In-Depth Comparison with Reference Theoretical Methods. Figure 4a shows the correlation between the reference CCSD(T)_r and the CCSD_r dipole moments (both with relaxed density matrices). We observe an excellent agreement between the two methods for singly bonded molecules (represented by circles). The correlation worsens for multiply bonded systems (marked by squares), underlining the importance of triple excitations. Figure 4b shows good agreement between CCSD(T) results using relaxed and unrelaxed density matrices. The only exception is the MgO molecule (denoted by a triangular shape in Figure 4), for which relaxation has a more profound effect.

By comparing the pCCD-based dipole moments with CCSD(T)_r, we observe a set of characteristic features for each molecule type. Molecules with negligible relaxation effects and triple excitations dependence (mainly singly bonded) provide a very satisfactory agreement between all pCCD-based methods and reference results (cf. Figure 4c-d). Although the variation among pCCD-based methods is slight, we note that the pCCD-LCCSD variant using the canonical orbitals leads to the smallest errors. On the contrary, when orbital-optimized pCCD orbitals are employed, the LCCD correction is the most reliable and results in the smallest errors. Surprisingly, the LCCSD correction on top of oo-pCCD increases the error in some cases.

The MgO molecule presents the most challenging test case for pCCD-LCCD and pCCD-LCCSD methods (cf. Figure 4c). The oo-pCCD-LCCD dipole moment is similar to the pCCD-LCCSD using canonical HF orbitals, which suggests that the orbital relaxation has recovered the effect of the linearized single excitations (compare Figures 4c and 4d). With the LCCSD correction on top of the oo-pCCD, the dipole moment agreement with the CCSD(T)_r reference value improves significantly. Specifically, the absolute (and relative) error in the MgO dipole moment reduces from 2.23 D (36%) to 0.46 D (7%) when moving from oo-pCCD-LCCD to oo-pCCD-LCCSD, respectively.

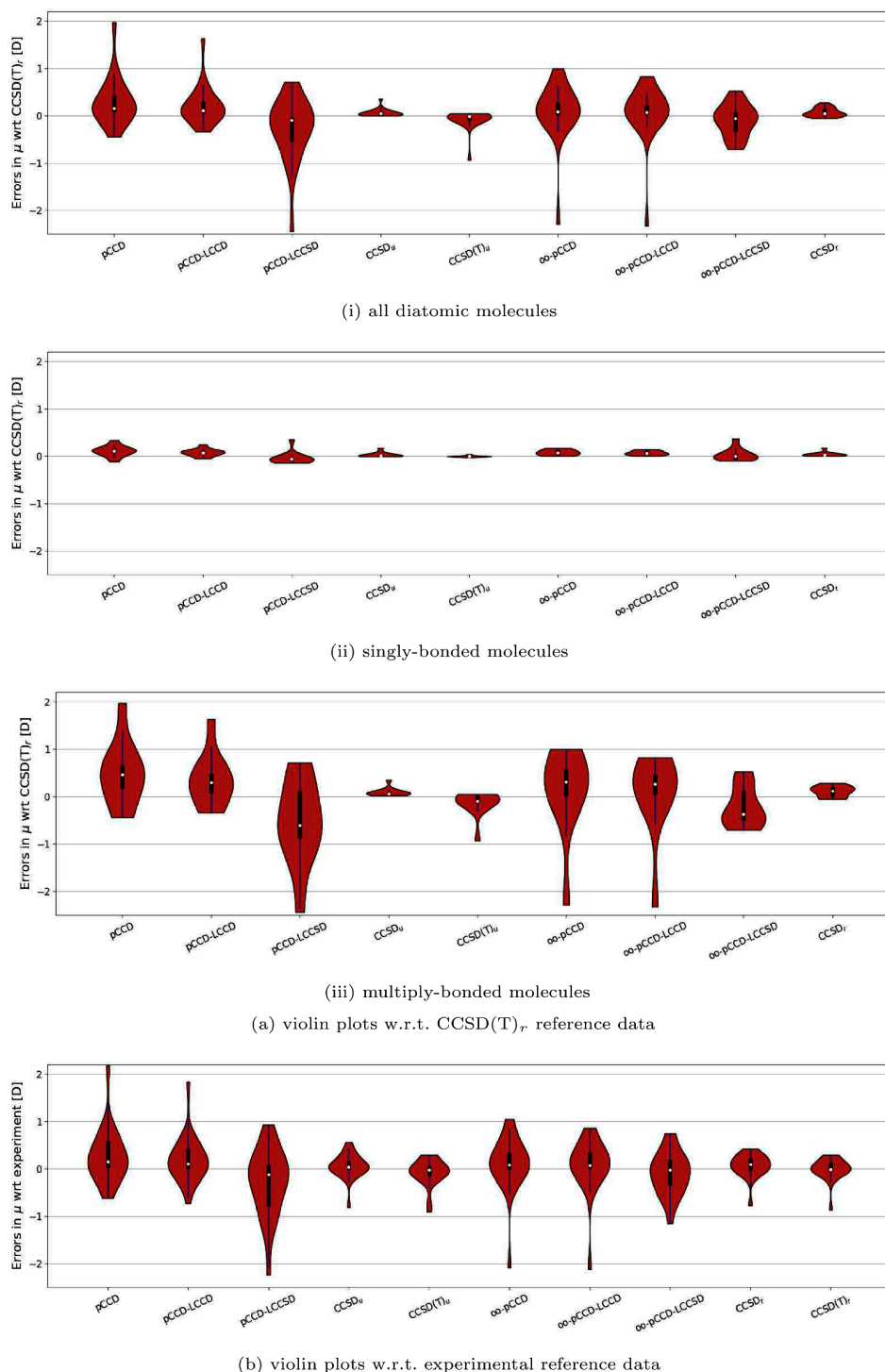


Figure 3. Violin plots illustrating errors (in D) derived from selected methods (refer to Table S2 for numerical values). All errors are reported relative to either (a) $\text{CCSD}(T)_r$ or (b) experimental reference data. A dot in each violin plot represents the median value, while the blue line indicates the 1.5 interquartile range and the black bar the quartile range, respectively.

The diatomic molecules with a large contribution of triple excitations to the dipole moment show a similar, but smaller, swing in dipole values between the pCCD-LCCD and pCCD-LCCSD as the one seen for the MgO molecule. However, as the main change in dipole moments is not due to an orbital relaxation effect, the oo-pCCD variation leads to a dipole value closer to the pCCD than the pCCD-LCCSD one. Consequently, the oo-pCCD-LCCD and oo-pCCD-LCCSD results

approach the reference from opposite directions. Although the orbital optimization improves the results, the oo-pCCD-LCCD and oo-pCCD-LCCSD dipole moment values have similar but substantial errors. The only exception is the carbon-containing compounds; in these cases, the oo-pCCD-LCCSD error to the $\text{CCSD}(T)_r$ is higher than the oo-pCCD-LCCD. Once some of the studied systems require triple excitations, as concluded during the analysis of Figure 4a, none of the investigated

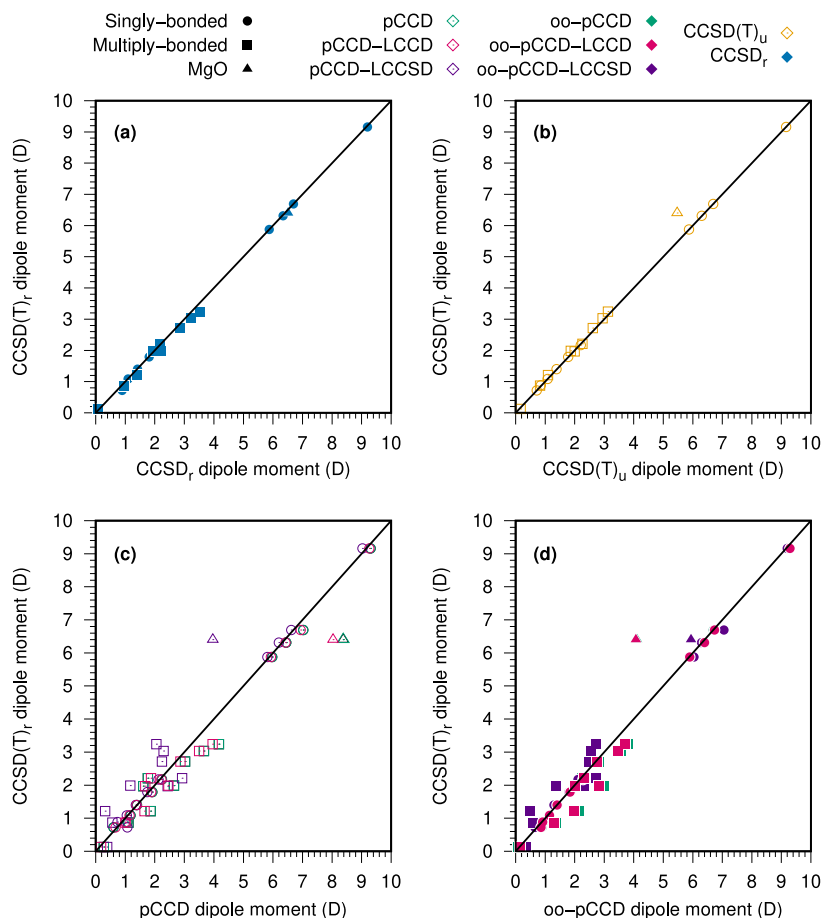


Figure 4. Correlation between the reference CCSD(T)_r dipole moments (in D) and other CC-based methods. (a) relaxed CCSD; (b) unrelaxed CCSD(T); (c) pCCD and pCCD with LCC corrections; and (d) oo-pCCD and oo-pCCD with LCC corrections.

pCCD-based approaches can recover this effect, and, therefore, such an error is expected.

Based on this analysis, the variation of dipole moment values among pCCD, oo-pCCD, and pCCD-LCCSD results can be used to estimate the magnitude of the orbital relaxation and triple excitations for the dipole moment. Systems, where the three values agree with each other have a small dependence on orbital relaxation and triple excitations. Thus, either the pCCD-LCCSD or oo-pCCD-LCCD leads to minor errors with respect to the CCSD(T)_r reference, that is, a relative average error of around 4%. When oo-pCCD and pCCD-LCCSD are similar, orbital relaxation is required, and the oo-pCCD-LCCSD value should be preferable. Lastly, for distinct oo-pCCD and pCCD-LCCSD values, pCCD-based methods would require a larger excitation order to be reliable. In these cases, excluding the carbon-containing molecules, both oo-pCCD-LCCD and oo-pCCD-LCCSD methods have a relative average error of around 30%. Including the carbon-based ones, the oo-pCCD-LCCD error decreases to 21%, while the oo-pCCD-LCCSD one increases up to 43%.

4.2. Dipole Moment Surfaces with pCCD-Based Methods. Dipole moment surfaces (DMS) are essential for estimating rovibrational spectroscopic parameters of molecules. Here, we focus on the DMS of two main group diatomic molecules, HF and CO. Their DMSs have been widely studied^{5,128} in previous theoretical works and, thus, represent suitable test cases for the investigated pCCD-based methods in

different bond length regions. In this work, the diatomics AB are placed along the z -axis with A (the less electronegative atom) at the origin and B on the positive z -axis. Then, the bond between the two atoms of AB is stretched along the positive z -axis for constructing the DMS. Hence, a positive μ_z value will indicate A^-B^+ polarity, whereas a negative μ_z indicates the same as A^+B^- .

4.2.1. Hydrogen Fluoride (HF). Figure 5 shows the DMS of the HF molecule in the aug-cc-pVTZ basis, calculated with oo-pCCD-based methods. We also included the CCSD and

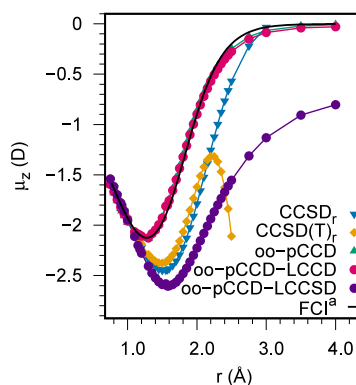


Figure 5. Dipole moment surface of HF in aug-cc-pVTZ basis. ^aFCI/cc-pVDZ DMS is taken from Samanta and Köhn.¹²⁹

CCSD(T) DMSs (both with relaxed densities, i.e., CCSD_r and CCSD(T)_r) and the FCI DMS (determined for the cc-pVDZ basis set)¹²⁹ for comparison. Around the equilibrium distance ($r_e = 0.917$ Å), all oo-pCCD variants agree well with CCSD_r and CCSD(T)_r, as discussed for singly bonded systems in section 4.1. Passed that region, significant deviations are observed between the curves of oo-pCCD variants and the conventional CC curves. Orbital relaxation has become essential in that region. The CCSD_r and CCSD(T)_r dipole moment values significantly deviate from the FCI results. As discussed by Samanta and Köhn,¹²⁹ in this region, the CCSD is unable to compensate the ionic contribution of the Hartree–Fock reference wave function. Although the inclusion of full triple excitations (CCSDT) can improve the CCSD poor modeling, it is not a reasonable zeroth-order wave function for the inclusion of triple excitations perturbatively. This poor description by CC methods during bond-stretching is reinforced by the change in the DMS behavior beyond 2.00 Å and the lack of convergence of coupled perturbed Hartree–Fock (CPHF) calculations for CCSD(T)_r at 2.25 Å. Therefore, the CCSD(T)_r dipole moment values are not reliable beyond this point.

The oo-pCCD and oo-pCCD-LCCD DMS lie on top of each other for almost the entire bond length region, indicating the lower significance of the doubles correction on top of pCCD. In good agreement with the previous FCI results,¹²⁹ both the oo-pCCD and oo-pCCD-LCCD dipole moment curves show turnings at around 1.30–1.35 Å and present a much shallower DMS compared to the other methods from Figure 5. These results indicate that the oo-pCCD and oo-pCCD-LCCD can model the HF dipole moment at the bond stretching and dissociation regions. Both have the right shape at larger interatomic distances and converge to the proper asymptotic limit. The oo-pCCD-LCCSD curve, on the other hand, overlaps with the CC curves to a slightly longer bond distance. It also turns at a greater bond length (around 1.60–1.65 Å), showing closer agreement with the turning of CC curves (around 1.50–1.55 Å). At stretched bond lengths, the oo-pCCD-LCCSD curve remains below the CC curve and does not converge to the correct asymptotic limit. That indicates that the linearized singles correction on top of the oo-pCCD wave function modifies the dipole toward the CCSD results but overshoots it at stretched geometries.

4.2.2. Carbon Monoxide (CO). We focused on the region from 0.75 to 1.50 Å in the CO DMS study. The HF and pCCD wave function optimization beyond that region is very challenging⁶² and will likely not provide reliable dipole moments. In this range of interest of internuclear distances, the CCSD(T)_r shows a remarkable agreement with the fitted MRCISD+Q dipole values using the finite-field approach and aug-cc-pCV6D basis set¹³⁰ as shown in Figure 6. As discussed in section 4.1, for the CO case, triple excitations are relevant from the equilibrium distance (around 1.13 Å) onward. That is indicated by the growing splitting between CCSD_r and CCSD(T)_r dipole values in Figure 6.

Similar to what we observed for HF curves, oo-pCCD-LCCSD overestimates the CO dipole value for large equilibrium distances and has small errors only at the repulsive region (see Figure 5). To that end, the oo-pCCD-LCCSD DMS of CO is not reliable. On the other hand, the oo-pCCD DMS matches the CCSD_r from 0.90 to 1.25 Å and the oo-pCCD-LCCD DMS resembles the shape of CCSD(T)_r up to 1.28 Å. Throughout the internuclear distances, the average

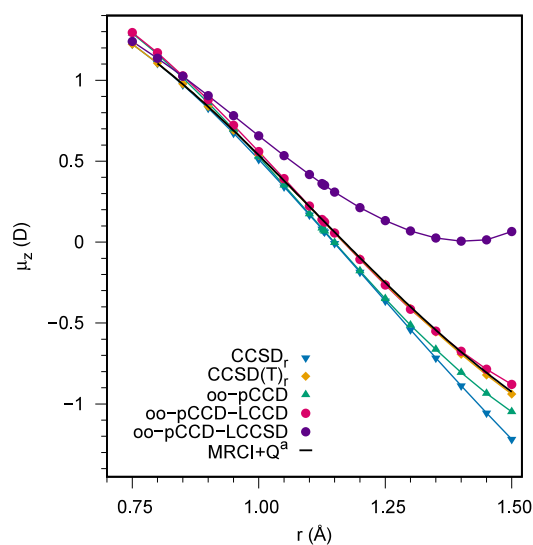


Figure 6. Dipole moment surface of CO in aug-cc-pVTZ basis. ^aThe MRCI+Q/aug-cc-pCV6Z values have been taken from Balashov et al.¹³⁰

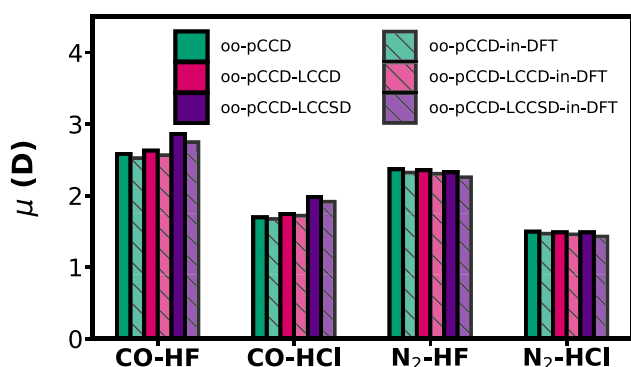
absolute error in the dipole moment of oo-pCCD-LCCD compared to the MRCI+Q reference is about 0.023 D (or around 4% considering relative errors). Thus, the oo-pCCD-LCCD provides comparable DMS with the computationally more expensive multireference and CCSD(T)_r calculations.

4.3. Dipole Moments from pCCD-Based Static Embedding. Dipole moments are often used to assess the performance of DFT-based embedding approaches.¹³¹ The calculated dipole moments are susceptible to electron density changes caused by environmental effects and, thus, are valuable measures for validating the quality of the embedding potential.^{132,133} To that end, we investigate the performance of recently implemented pCCD-in-DFT static embedding models⁹² for two sets of weakly interacting systems: linear hydrogen-bonded binary complexes and coplanar water complexes with noble gases. Their structural parameters are presented in Figures 1a and 1b, respectively. Building on the experience gained in the previous section and knowing the importance of orbital relaxation in oo-pCCD, we solely focused on orbital-optimized variants. The supramolecular oo-pCCD-LCCSD dipole moments show low error with respect to the CCSD(T)_r data (shown in Table S9 of the SI) and, thus, provide a reliable supramolecular reference except for CO-HF and CO-HCl, where oo-pCCD-LCCD performs better, similarly to the observer for the isolated CO molecule in section 4.2.2.

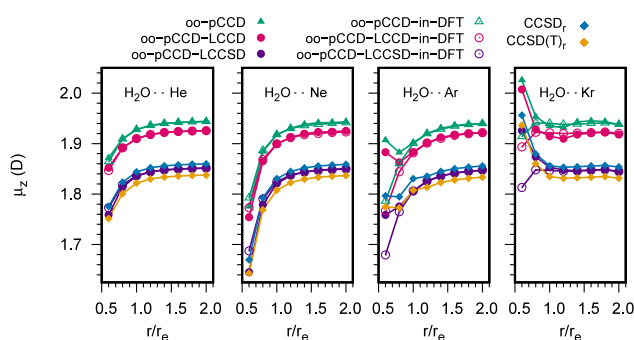
Table 2 collects dipole moments obtained from various pCCD models with and without embedding and the difference between them. Figure 7 summarizes the performance of the orbital optimized pCCD-based embedding models for dipole moments of weakly hydrogen-bonded complexes (the binary complexes, see also Figure 1a). The static embedding approach produces dipole moments closer to the respective supramolecular values with both oo-pCCD and oo-pCCD-LCC methods. Interestingly, the difference in embedding and supramolecular dipole moment values is lower with oo-pCCD and oo-pCCD-LCCD compared to oo-pCCD-LCCSD. This is most likely attributed to the limitations of oo-pCCD-LCCSD when individual fragments possess multiple bonds, as we observed for the diatomics (vide supra).

Table 2. Dipole Moment (μ in D) from aug-cc-pVTZ oo-pCCD and oo-pCCD-in-DFT Types of Methods and Their Differences^a

Complex	oo-pCCD			oo-pCCD-LCCD			oo-pCCD-LCCSD		
	supra.	emb.	error	supra.	emb.	error	supra.	emb.	error
CO–HF	2.586	2.528	−0.058	2.630	2.569	−0.061	2.866	2.752	−0.114
CO–HCl	1.698	1.677	−0.021	1.745	1.723	−0.022	1.983	1.918	−0.065
N ₂ –HF	2.371	2.325	−0.046	2.357	2.309	−0.048	2.329	2.260	−0.069
N ₂ –HCl	1.497	1.471	−0.026	1.488	1.461	−0.027	1.488	1.431	−0.057
H ₂ O···He	1.928	1.928	0.000	1.910	1.910	0.000	1.836	1.836	0.000
H ₂ O···Ne	1.918	1.918	0.000	1.899	1.900	0.001	1.821	1.824	0.003
H ₂ O···Ar	1.905	1.904	−0.001	1.887	1.885	−0.002	1.810	1.810	0.000
H ₂ O···Kr	1.940	1.948	0.008	1.922	1.930	0.008	1.863	1.856	−0.007

^aThe errors are calculated as $\mu_{\text{emb.}} - \mu_{\text{supra.}}$.**Figure 7.** Dipole moments (μ in D) of the binary complexes from oo-pCCD variants and the corresponding embedding approaches in the aug-cc-pVTZ basis set.

We also study the dipole moments of the van der Waal's complexes between H₂O and the first four inert gases. Here, the performance of the static embedding approach is even better for all oo-pCCD variants. This is to be expected as, for these complexes, the electronic properties are dominated by the highly polar H₂O molecule, and it is easier to estimate them with embedding. As far as the supramolecular results in comparison to CCSD(T)_r are concerned, oo-pCCD-LCCSD shows the best performance, with errors comparable to CCSD_r (bottom part of Table S9 of the SI). Most importantly, the changes in the dipole moment with change in the inert gas molecule (decrease from He to Ar and then increase for Kr) are captured by all oo-pCCD-based methods (supramolecular and embedding). Figure 8 shows the change in dipole

**Figure 8.** Distance dependence of the calculated dipole moment components of the H₂O···Rg [Rg = He, Ne, Ar, and Kr] complexes in aug-cc-pVTZ basis.

moments of the H₂O···Rg complexes with the distance between H₂O and the inert gas atom. For these curves, the distance between H₂O and the Rg atom is increased in multiples of the equilibrium distances, keeping the angles the same for the respective structures. Here, we plot the major component of the dipole, that is μ_z . A plot for μ_x is shown in Figure S2 of the SI. For H₂O···He and H₂O···Ne, the supramolecular trends in the changes in the dipole are well-reproduced by the embedding methods throughout the distances scanned. For H₂O···Ar and H₂O···Kr, the embedding methods differ from the supramolecular variants significantly at shorter distances. We anticipate that this is caused by the shortcoming of the kinetic energy functional, which has been observed for other complexes with Ar and Kr.¹³² The nonparallelity errors (difference between highest error and lowest error between embedding and supramolecular curves) are 0.121, 0.116, and 0.079 (H₂O···Ar), and 0.112, 0.114, and 0.112 (H₂O···Kr) for oo-pCCD-in-DFT, oo-pCCD-LCCD-in-DFT, and oo-pCCD-LCCSD-in-DFT respectively.

Barring the initial points for H₂O···Ar and H₂O···Kr, the oo-pCCD-LCCSD curves (both supra and embedding) are between those of CCSD_r and CCSD(T)_r for all systems. To conclude, the performance of both oo-pCCD-LCCSD and oo-pCCD-LCCSD-in-DFT is encouraging for these systems, keeping in mind the low computational cost of the static embedding approach.

5. CONCLUSIONS AND OUTLOOK

In this work, we investigated the performance of various pCCD-based methods for predicting dipole moments. Our study shows that orbital optimization is essential and improves the overall performance of pCCD-based methods. Altogether, the best performance is obtained for the oo-pCCD-LCCD method, which is comparable to CCSD in predicting dipole moments. Specifically, oo-pCCD-LCCD approaches CCSD accuracy in dipole moments for singly bonded systems, while it reproduces the DMSs obtained by multireference methods. Thus, we demonstrated that reliable dipole moments can also be obtained without explicitly including single excitations in the wave function model.

For equilibrium structures, oo-pCCD-LCCD provides good agreement with the CCSD(T)_r dipole moment values for singly bonded systems—for instance, HF, AlF, and LiNa. For multiply bonded systems (such as SiO, GeS, and PN), the oo-pCCD-LCCD performance deteriorates (errors w.r.t. CCSD(T)_r are up to around 30%). The only exception is systems containing the carbon atom, where the relative errors drop below 5%. The oo-pCCD-LCCD approach is also noticeably

good in the modeling of DMSs. Specifically, for the HF molecule, oo-pCCD-LCCD provides excellent agreement with FCI even in the region where CCSD (and CCSD(T)) fail. For carbon monoxide (up to a distance of 1.50 Å), the agreement among oo-pCCD-LCCD, CCSD(T), and MRCISD+Q results is remarkable.

On the contrary, the presence of linearized singles in the LCC correction on top of the pCCD reference worsens the performance when multiply bonded diatomic molecules are considered. That is particularly true for the investigated DMSs, where the LCCSD correction provides erroneous dipole moments. The presence of singles, however, improves the description of van der Waals complexes as singles are crucial for dispersion interactions.⁶⁴ All pCCD-in-DFT models provide similar results for supramolecular and embedded dipole moments. As expected, for van der Waals complexes, the oo-pCCD-LCCSD provides the best agreement with coupled cluster reference data.

Finally, this work provides a reference point for further improvements of pCCD-based models. Specifically, our in-depth analysis of dipole moments demonstrates that when oo-pCCD provides a good reference function (like van der Waals and single-bonded systems), the LCCD (for singly bonded systems) and LCCSD (for van der Waals interactions) corrections can improve the electric properties of the system. We point out cases (e.g., multiple-bonded systems) where oo-pCCD does not produce reliable dipole moments, despite giving qualitatively correct potential energy surfaces as observed in previous works.^{62,87,134} For such molecules, LCCD does not improve the overall description, and pCCD-LCCSD tends to overcorrect dipole moments. It remains to be checked if using other than response density matrices (which are linear in nature) will bring some improvements. Furthermore, it needs to be determined whether frozen-pair or tailored variants of pCCD-based models⁶² will correct for deficiencies in the investigated LCC corrections.

■ ASSOCIATED CONTENT

Data Availability Statement

The data underlying this study are available in the published article and its Supporting Information. The PyBEST code is available on Zenodo at <https://zenodo.org/records/10069179> and on PyPI at <https://pypi.org/project/pybest/>.

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.4c00471>.

Experimental dipole moments and bond lengths, dipole moments from different basis sets, dipole moments without a frozen core approximation, data for DMSs of HF and CO, CO potential energy surface, structural parameters for the H₂O...Rg systems, and additional comparisons and graphs for embedding results (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Matheus Morato F. de Moraes — Institute of Physics, Faculty of Physics, Astronomy, and Informatics, Nicolaus Copernicus University in Toruń, 87-100 Toruń, Poland; orcid.org/0000-0003-3423-7949; Email: matheusmorat@gmail.com

Katharina Boguslawski — Institute of Physics, Faculty of Physics, Astronomy, and Informatics, Nicolaus Copernicus University in Toruń, 87-100 Toruń, Poland; orcid.org/

[0000-0001-7793-1151](https://orcid.org/0000-0001-7793-1151); Email: k.boguslawski@fizyka.umk.pl

Paweł Tecmer — Institute of Physics, Faculty of Physics, Astronomy, and Informatics, Nicolaus Copernicus University in Toruń, 87-100 Toruń, Poland; orcid.org/0000-0001-6347-878X; Email: ptecmer@fizyka.umk.pl

Authors

Rahul Chakraborty — Institute of Physics, Faculty of Physics, Astronomy, and Informatics, Nicolaus Copernicus University in Toruń, 87-100 Toruń, Poland; orcid.org/0000-0003-4086-6573

Artur Nowak — Institute of Physics, Faculty of Physics, Astronomy, and Informatics, Nicolaus Copernicus University in Toruń, 87-100 Toruń, Poland

Julian Świerczyński — Institute of Engineering and Technology, Faculty of Physics, Astronomy, and Informatics, Nicolaus Copernicus University in Toruń, 87-100 Toruń, Poland

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.jctc.4c00471>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

R.C. and P.T. acknowledge financial support from the OPUS research grant from the National Science Centre, Poland (Grant No. 2019/33/B/ST4/02114). P.T. acknowledges the scholarship for outstanding young scientists from Poland's Ministry of Science and Higher Education. R.C. thanks Dr. Marta Gałynska for many interesting discussions on the topic. Funded/Co-funded by the European Union (ERC, DRESSED-pCCD, 101077420). Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

■ REFERENCES

- (1) Honig, B.; Nicholls, A. Classical electrostatics in biology and chemistry. *Science* **1995**, 268, 1144–1149.
- (2) Frenking, G.; Loschen, C.; Krapp, A.; Fau, S.; Strauss, S. H. Electronic structure of CO—An exercise in modern chemical bonding theory. *J. Comput. Chem.* **2007**, 28, 117–126.
- (3) Neugebauer, J.; Reiher, M.; Kind, C.; Hess, B. A. Quantum chemical calculation of vibrational spectra of large molecules—Raman and IR spectra for Buckminsterfullerene. *J. Comput. Chem.* **2002**, 23, 895–910.
- (4) Fedorov, D. A.; Barnes, D. K.; Varganov, S. A. Ab initio calculations of spectroscopic constants and vibrational state lifetimes of diatomic alkali-alkaline-earth cations. *J. Chem. Phys.* **2017**, 147, DOI: [10.1063/1.4986818](https://doi.org/10.1063/1.4986818).
- (5) Buldakov, M.; Cherepanov, V.; Koryukina, E.; Kalugina, Y. N. On some aspects of changing the sign of the dipole moment functions of diatomic molecules. *J. Phys. B: Atom. Mol. Opt. Phys.* **2009**, 42, 105102.
- (6) Lodi, L.; Tolchenov, R. N.; Tennyson, J.; Lynas-Gray, A.; Shirin, S. V.; Zobov, N. F.; Polyansky, O. L.; Császár, A. G.; van Stralen, J. N.; Visscher, L. A new ab initio ground-state dipole moment surface for the water molecule. *J. Chem. Phys.* **2008**, 128, DOI: [10.1063/1.2817606](https://doi.org/10.1063/1.2817606).
- (7) Cohen, H. D.; Roothaan, C. Electric dipole polarizability of atoms by the Hartree—Fock method. I. Theory for closed-shell systems. *J. Chem. Phys.* **1965**, 43, S34–S39.

- (8) Yoshimine, M.; McLean, A. Ground states of linear molecules: Dissociation energies and dipole moments in the Hartree–Fock approximation. *Int. J. Quantum Chem.* **1967**, *1*, 313–326.
- (9) Green, S. Electric Dipole Moment of Diatomic Molecules by Configuration Interaction. I. Closed-Shell Molecules. *J. Chem. Phys.* **1971**, *54*, 827–832.
- (10) Boyd, D. B. Electron density and dipole moment analysis of valence-electron wave functions. *J. Am. Chem. Soc.* **1972**, *94*, 64–70.
- (11) Diercksen, G. H.; Kellö, V.; Sadlej, A. J. Perturbation theory of the electron correlation effects for atomic and molecular properties. VII. Complete fourth-order MBPT study of the dipole moment and dipole polarizability of H₂O. *J. Chem. Phys.* **1983**, *79*, 2918–2923.
- (12) Hickey, A. L.; Rowley, C. N. Benchmarking quantum chemical methods for the calculation of molecular dipole moments and polarizabilities. *J. Phys. Chem. A* **2014**, *118*, 3678–3687.
- (13) Liu, X.; McKemmish, L.; Pérez-Ríos, J. The performance of CCSD(T) for the calculation of dipole moments in diatomics. *Phys. Chem. Chem. Phys.* **2023**, *25*, 4093–4104.
- (14) Noga, J.; Urban, M. On expectation value calculations of one-electron properties using the coupled cluster wave functions. *Theor. Chem. Acc.* **1988**, *73*, 291–306.
- (15) Korona, T.; Pflüger, K.; Werner, H.-J. The effect of local approximations in coupled-cluster wave functions on dipole moments and static dipole polarisabilities. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2059–2065.
- (16) Bozkaya, U.; Soydaş, E.; Filiz, B. State-of-the-art computations of dipole moments using analytic gradients of high-level density-fitted coupled-cluster methods with focal-point approximations. *J. Comput. Chem.* **2020**, *41*, 769–779.
- (17) Traore, D.; Toulouse, J.; Giner, E. Basis-set correction for coupled-cluster estimation of dipole moments. *J. Chem. Phys.* **2022**, *156*, 174101.
- (18) Maroulis, G. A systematic study of basis set, electron correlation, and geometry effects on the electric multipole moments, polarizability, and hyperpolarizability of HCl. *J. Chem. Phys.* **1998**, *108*, 5432–5448.
- (19) Maroulis, G.; Thakkar, A. J. Multipole moments, polarizabilities, and hyperpolarizabilities for N₂ from fourth-order many-body perturbation theory calculations. *J. Chem. Phys.* **1988**, *88*, 7623–7632.
- (20) Maroulis, G.; Thakkar, A. J. Erratum: Multipole moments, polarizabilities, and hyperpolarizabilities for N₂ from fourth-order many-body perturbation theory calculations [J. Chem. Phys. 88, 7623 (1988)]. *J. Chem. Phys.* **1988**, *89*, 6558–6558.
- (21) Maroulis, G.; Bishop, D. M. Electric moments, polarizabilities and hyperpolarizabilities for BH(X¹Σ⁺) and CH⁺(X¹Σ⁺). *Chem. Phys.* **1985**, *96*, 409–418.
- (22) Maroulis, G. Electric moments, polarizabilities and hyperpolarizabilities for carbon disulfide (S = C=S) from accurate SCF calculations. *Chem. Phys. Lett.* **1992**, *199*, 250–256.
- (23) Maroulis, G. Hyperpolarizability of H₂O. *J. Chem. Phys.* **1991**, *94*, 1182–1190.
- (24) Maroulis, G. Electric multipole moment, dipole and quadrupole (hyper) polarizability derivatives for HF (X¹Σ⁺). *J. Mol. Struct.: THEOCHEM* **2003**, *633*, 177–197.
- (25) Maroulis, G.; Xenides, D.; Hohm, U.; Loose, A. Dipole, dipole–quadrupole, and dipole–octupole polarizability of adamantane, C₁₀H₁₆, from refractive index measurements, depolarized collision-induced light scattering, conventional ab initio and density functional theory calculations. *J. Chem. Phys.* **2001**, *115*, 7957–7967.
- (26) Gidofalvi, G.; Mazziotti, D. A. Computation of dipole, quadrupole, and octupole surfaces from the variational two-electron reduced density matrix method. *J. Chem. Phys.* **2006**, *125*. DOI: 10.1063/1.2355490.
- (27) Gidofalvi, G.; Mazziotti, D. A. Molecular properties from variational reduced-density-matrix theory with three-particle N-representability conditions. *J. Chem. Phys.* **2007**, *126*. DOI: 10.1063/1.2423008.
- (28) Mazziotti, D. A. Parametrization of the two-electron reduced density matrix for its direct calculation without the many-electron wave function: Generalizations and applications. *Phys. Rev. A* **2010**, *81*, 062515.
- (29) Chrayteh, A.; Blondel, A.; Loos, P.-F.; Jacquemin, D. Mountaineering strategy to excited states: highly accurate oscillator strengths and dipole moments of small molecules. *J. Chem. Theory Comput.* **2021**, *17*, 416–438.
- (30) Sarkar, R.; Boggio-Pasqua, M.; Loos, P.-F.; Jacquemin, D. Benchmarking TD-DFT and wave function methods for oscillator strengths and excited-state dipole moments. *J. Chem. Theory Comput.* **2021**, *17*, 1117–1132.
- (31) Damour, Y.; Quintero-Monsebaiz, R.; Caffarel, M.; Jacquemin, D.; Kossoski, F.; Scemama, A.; Loos, P.-F. Ground-and excited-state dipole moments and oscillator strengths of full configuration interaction quality. *J. Chem. Theory Comput.* **2023**, *19*, 221–234.
- (32) Korona, T.; Jezierski, B. One-electron properties and electrostatic interaction energies from the expectation value expression and wave function of singles and doubles coupled cluster theory. *J. Chem. Phys.* **2006**, *125*. DOI: 10.1063/1.2364489.
- (33) Fecteau, C.-É.; Fortin, H.; Cloutier, S.; Johnson, P. A. Reduced density matrices of Richardson–Gaudin states in the Gaudin algebra basis. *J. Chem. Phys.* **2020**, *153*, 164117.
- (34) Johnson, P. A.; Fortin, H.; Cloutier, S.; Fecteau, C.-É. Transition density matrices of Richardson–Gaudin states. *J. Chem. Phys.* **2021**, *154*, 124125.
- (35) Salter, E.; Sekino, H.; Bartlett, R. J. Property evaluation and orbital relaxation in coupled cluster methods. *J. Chem. Phys.* **1987**, *87*, 502–509.
- (36) Klopper, W.; Noga, J.; Koch, H.; Helgaker, T. Multiple basis sets in calculations of triples corrections in coupled-cluster theory. *Theor. Chem. Acc.* **1997**, *97*, 164–176.
- (37) Liu, X.; McKemmish, L.; Pérez-Ríos, J. The performance of CCSD(T) for the calculation of dipole moments in diatomics. *Phys. Chem. Chem. Phys.* **2023**, *25*, 4093–4104.
- (38) Kellö, V.; Noga, J.; Diercksen, G. H.; Sadlej, A. J. A study of the performance of high-level correlated methods: the energy, dipole moment, and polarizability functions of CO. *Chem. Phys. Lett.* **1988**, *152*, 387–392.
- (39) Werner, H.-J. MCSCF calculation of the dipole moment function of CO. *Mol. Phys.* **1981**, *44*, 111–123.
- (40) Scuseria, G. E.; Miller, M. D.; Jensen, F.; Geertsens, J. The dipole moment of carbon monoxide. *J. Chem. Phys.* **1991**, *94*, 6660–6663.
- (41) Barnes, L. A.; Liu, B.; Lindh, R. Bond length, dipole moment, and harmonic frequency of CO. *J. Chem. Phys.* **1993**, *98*, 3972–3977.
- (42) Schautz, F.; Flad, H.-J. Quantum Monte Carlo study of the dipole moment of CO. *J. Chem. Phys.* **1999**, *110*, 11700–11707.
- (43) Meshkov, V. V.; Ermilov, A. Y.; Stolyarov, A. V.; Medvedev, E. S.; Ushakov, V. G.; Gordon, I. E. Semi-empirical dipole moment of carbon monoxide and line lists for all its isotopologues revisited. *J. Quant. Spectrosc. Radiat. Transfer* **2022**, *280*, 108090.
- (44) Hapka, M.; Pernal, K.; Jensen, H. J. A. An efficient implementation of time-dependent linear-response theory for strongly orthogonal geminal wave function models. *J. Chem. Phys.* **2022**, *156*, 174102.
- (45) Johnson, P. A.; Ayers, P. W.; Limacher, P. A.; De Baerdemacker, S.; Van Neck, D.; Bultinck, P. A size-consistent approach to strongly correlated systems using a generalized antisymmetrized product of nonorthogonal geminals. *Comput. Chem. Theory* **2013**, *1003*, 101–113.
- (46) Johnson, P. A.; Limacher, P. A.; Kim, T. D.; Richer, M.; Miranda-Quintana, R. A.; Heidar-Zadeh, F.; Ayers, P. W.; Bultinck, P.; De Baerdemacker, S.; Van Neck, D. Strategies for extending geminal-based wavefunctions: Open shells and beyond. *Comput. Theor. Chem.* **2017**, *1116*, 207–219.
- (47) Kim, T. D.; Miranda-Quintana, R. A.; Richer, M.; Ayers, P. W. Flexible ansatz for N-body configuration interaction. *Comput. Theor. Chem.* **2021**, *1202*, 113187.

- (48) Johnson, P. A.; Fecteau, C.-É.; Berthiaume, F.; Cloutier, S.; Carrier, L.; Gratton, M.; Bultinck, P.; De Baerdemacker, S.; Van Neck, D.; Limacher, P.; Ayers, P. W. Richardson–Gaudin mean-field for strong correlation in quantum chemistry. *J. Chem. Phys.* **2020**, *153*, 104110.
- (49) Johnson, P. A.; Ayers, P. W.; De Baerdemacker, S.; Limacher, P. A.; Van Neck, D. Bivariational principle for an antisymmetrized product of nonorthogonal geminals appropriate for strong electron correlation. *Comput. Theor. Chem.* **2022**, *1212*, 113718.
- (50) Limacher, P. A.; Ayers, P. W.; Johnson, P. A.; De Baerdemacker, S.; Van Neck, D.; Bultinck, P. A new mean-field method suitable for strongly correlated electrons: computationally facile antisymmetric products of nonorthogonal geminals. *J. Chem. Theory Comput.* **2013**, *9*, 1394–1401.
- (51) Boguslawski, K.; Tecmer, P.; Ayers, P. W.; Bultinck, P.; De Baerdemacker, S.; Van Neck, D. Efficient description of strongly correlated electrons. *Phys. Rev. B* **2014**, *89*, No. 201106(R).
- (52) Stein, T.; Henderson, T. M.; Scuseria, G. E. Seniority zero pair coupled cluster doubles theory. *J. Chem. Phys.* **2014**, *140*, 214113.
- (53) Tecmer, P.; Boguslawski, K. Geminal-based electronic structure methods in quantum chemistry. Toward geminal model chemistry. *Phys. Chem. Chem. Phys.* **2022**, *24*, 23026–23048.
- (54) Boguslawski, K.; Tecmer, P.; Limacher, P. A.; Johnson, P. A.; Ayers, P. W.; Bultinck, P.; De Baerdemacker, S.; Van Neck, D. Projected seniority-two orbital optimization of the antisymmetric product of one-reference orbital geminal. *J. Chem. Phys.* **2014**, *140*, 214114.
- (55) Limacher, P. A.; Kim, T. D.; Ayers, P. W.; Johnson, P. A.; De Baerdemacker, S.; Van Neck, D.; Bultinck, P. The influence of orbital rotation on the energy of closed-shell wavefunctions. *Mol. Phys.* **2014**, *112*, 853–862.
- (56) Boguslawski, K.; Tecmer, P.; Ayers, P. W.; Bultinck, P.; De Baerdemacker, S.; Van Neck, D. Non-variational orbital optimization techniques for the AP1roG wave function. *J. Chem. Theory Comput.* **2014**, *10*, 4873–4882.
- (57) Tecmer, P.; Boguslawski, K.; Limacher, P. A.; Johnson, P. A.; Chan, M.; Verstraelen, T.; Ayers, P. W. Assessing the accuracy of new geminal-based approaches. *J. Phys. Chem. A* **2014**, *118*, 9058–9068.
- (58) Tecmer, P.; Boguslawski, K.; Ayers, P. W. Singlet ground state actinide chemistry with geminals. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14427–14436.
- (59) Garza, A. J.; Sousa Alencar, A. G.; Scuseria, G. E. Actinide chemistry using singlet-paired coupled cluster and its combinations with density functionals. *J. Chem. Phys.* **2015**, *143*, 244106.
- (60) Boguslawski, K.; Tecmer, P.; Legeza, Ö. Analysis of two-orbital correlations in wavefunctions restricted to electron-pair states. *Phys. Rev. B* **2016**, *94*, 155126.
- (61) Boguslawski, K.; Leszczyk, A.; Nowak, A.; Brzęk, F.; Żuchowski, P. S.; Kędziera, D.; Tecmer, P. Pythonic Black-box Electronic Structure Tool (PyBEST). An open-source Python platform for electronic structure calculations at the interface between chemistry and physics. *Comput. Phys. Commun.* **2021**, *264*, 107933.
- (62) Leszczyk, A.; Máté, M.; Legeza, Ö.; Boguslawski, K. Assessing the accuracy of tailored coupled cluster methods corrected by electronic wave functions of polynomial cost. *J. Chem. Theory Comput.* **2022**, *18*, 96–117.
- (63) Leszczyk, A.; Dome, T.; Tecmer, P.; Kędziera, D.; Boguslawski, K. Resolving the π -assisted U–N σ f-bond formation using quantum information theory. *Phys. Chem. Chem. Phys.* **2022**, *24*, 21296–21307.
- (64) Brzęk, F.; Boguslawski, K.; Tecmer, P.; Zuchowski, P. S. Benchmarking the accuracy of seniority-zero wave function methods for noncovalent interactions. *J. Chem. Theory Comput.* **2019**, *15*, 4021–4035.
- (65) Garza, A. J.; Bulik, I. W.; Henderson, T. M.; Scuseria, G. E. Range separated hybrids of pair coupled cluster doubles and density functionals. *Phys. Chem. Chem. Phys.* **2015**, *17*, 22412–22422.
- (66) Boguslawski, K. Targeting excited states in all-trans polyenes with electron-pair states. *J. Chem. Phys.* **2016**, *145*, 234105.
- (67) Boguslawski, K. Erratum: Targeting excited states in all-trans polyenes with electron-pair states. *J. Chem. Phys.* **2017**, *147*, 139901.
- (68) Kossoski, F.; Marie, A.; Scemama, A.; Caffarel, M.; Loos, P.-F. Excited states from state-specific orbital-optimized pair coupled cluster. *J. Chem. Theory Comput.* **2021**, *17*, 4756–4768.
- (69) Boguslawski, K. Open-shell extensions to closed-shell pCCD. *Chem. Commun.* **2021**, *57*, 12277–12280.
- (70) Jahani, S.; Boguslawski, K.; Tecmer, P. The relationship between structure and excited-state properties in polyanilines from geminal-based methods. *RSC Adv.* **2023**, *13*, 27898–27911.
- (71) Tecmer, P.; Galińska, M.; Szczuczko, L.; Boguslawski, K. Geminal-based strategies for modeling large building blocks of organic electronic materials. *J. Phys. Chem. Lett.* **2023**, *14*, 9909–9917.
- (72) Limacher, P.; Ayers, P.; Johnson, P.; De Baerdemacker, S.; Van Neck, D.; Bultinck, P. Simple and Inexpensive Perturbative Correction Schemes for Antisymmetric Products of Nonorthogonal Geminals. *Phys. Chem. Chem. Phys.* **2014**, *16*, 5061–5065.
- (73) Henderson, T. M.; Bulik, I. W.; Stein, T.; Scuseria, G. E. Seniority-based coupled cluster theory. *J. Chem. Phys.* **2014**, *141*, 244104.
- (74) Boguslawski, K. Targeting Doubly Excited States with Equation of Motion Coupled Cluster Theory Restricted to Double Excitations. *J. Chem. Theory Comput.* **2019**, *15*, 18–24.
- (75) Boguslawski, K.; Tecmer, P. Benchmark of dynamic electron correlation models for seniority-zero wavefunctions and their application to thermochemistry. *J. Chem. Theory Comput.* **2017**, *13*, 5966–5983.
- (76) Limacher, P. A. Orbital energies for seniority-zero wave functions. *J. Chem. Theory Comput.* **2015**, *11*, 3629–3635.
- (77) Nowak, A.; Tecmer, P.; Boguslawski, K. Assessing the accuracy of simplified coupled cluster methods for electronic excited states in f0 actinide compounds. *Phys. Chem. Chem. Phys.* **2019**, *21*, 19039–19053.
- (78) Peinel, G. Calculation of dipole moments with a CNDO-APSG procedure. *Mol. Phys.* **1975**, *29*, 641–643.
- (79) Ogilvie, J.; Rodwell, W.; Tipping, R. Dipole moment functions of the hydrogen halides. *J. Chem. Phys.* **1980**, *73*, 5221–5229.
- (80) Piris, M. A natural orbital functional based on an explicit approach of the two-electron cumulant. *Int. J. Quantum Chem.* **2013**, *113*, 620–630.
- (81) Piris, M. Interacting pairs in natural orbital functional theory. *J. Chem. Phys.* **2014**, *141*, 044107.
- (82) Pernal, K. The equivalence of the Piris Natural Orbital Functional 5 (PNOF5) and the antisymmetrized product of strongly orthogonal geminal theory. *Comput. Theor. Chem.* **2013**, *1003*, 127–129.
- (83) Bozkaya, U.; Sherrill, C. D. Orbital-optimized coupled-electron pair theory and its analytic gradients: Accurate equilibrium geometries, harmonic vibrational frequencies, and hydrogen transfer reactions. *J. Chem. Phys.* **2013**, *139*, 054104.
- (84) Meyer, W.; Rosmus, P. PNO–CI and CEPA studies of electron correlation effects. III. Spectroscopic constants and dipole moment functions for the ground states of the first-row and second-row diatomic hydrides. *J. Chem. Phys.* **1975**, *63*, 2356–2375.
- (85) Werner, H.-J.; Meyer, W. PNO–CI and PNO–CEPA studies of electron correlation effects: V. Static dipole polarizabilities of small molecules. *Mol. Phys.* **1976**, *31*, 855–872.
- (86) Werner, H.-J.; Rosmus, P. Theoretical dipole moment functions of the HF, HCl, and HBr molecules. *J. Chem. Phys.* **1980**, *73*, 2319–2328.
- (87) Boguslawski, K.; Ayers, P. W. Linearized Coupled Cluster Correction on the Antisymmetric Product of 1-Reference Orbital Geminals. *J. Chem. Theory Comput.* **2015**, *11*, 5252–5261.
- (88) Nowak, A.; Legeza, Ö.; Boguslawski, K. Orbital entanglement and correlation from pCCD-tailored coupled cluster wave functions. *J. Chem. Phys.* **2021**, *154*, 084111.
- (89) Bak, K. L.; Gauss, J.; Helgaker, T.; Jørgensen, P.; Olsen, J. The accuracy of molecular dipole moments in standard electronic structure calculations. *Chem. Phys. Lett.* **2000**, *319*, 563–568.

- (90) Aoto, Y. A.; de Lima Batista, A. P.; Kohn, A.; de Oliveira-Filho, A. G. How to arrive at accurate benchmark values for transition metal compounds: Computation or experiment? *J. Chem. Theory Comput.* **2017**, *13*, 5291–5316.
- (91) Severo Pereira Gomes, A.; Jacob, C. R. Quantum-chemical embedding methods for treating local electronic excitations in complex chemical systems. *Annu. Rep. Prog. Chem., Sect. C* **2012**, *108*, 222.
- (92) Chakraborty, R.; Boguslawski, K.; Tecmer, P. Static embedding with pair coupled cluster doubles based methods. *Phys. Chem. Chem. Phys.* **2023**, *25*, 25377–25388.
- (93) Curtiss, L. A.; Pochatko, D. J.; Reed, A. E.; Weinhold, F. Investigation of the differences in stability of the $\text{OC} \cdots \text{HF}$ and $\text{CO} \cdots \text{HF}$ complexes. *J. Chem. Phys.* **1985**, *82*, 2679–2687.
- (94) Bacskay, G. B.; Kerdraon, D. I.; Hush, N. S. Quantum chemical study of the HCl molecule and its binary complexes with CO , C_2H_2 , C_2H_4 , PH_3 , H_2S , HCN , H_2O and NH_3 : hydrogen bonding and its effect on the ^{35}Cl nuclear quadrupole coupling constant. *Chem. Phys.* **1990**, *144*, 53–69.
- (95) Woon, D. E.; Dunning Jr, T. H.; Peterson, K. A. Ab initio investigation of the N_2 –HF complex: Accurate structure and energetics. *J. Chem. Phys.* **1996**, *104*, 5883–5891.
- (96) Haskopoulos, A.; Maroulis, G. Interaction electric hyperpolarizability effects in weakly bound $\text{H}_2\text{O} \cdots \text{Rg}$ ($\text{Rg} = \text{He, Ne, Ar, Kr}$ and Xe) complexes. *J. Phys. Chem. A* **2010**, *114*, 8730–8741.
- (97) Tecmer, P.; Boguslawski, K.; Borkowski, M.; Żuchowski, P. S.; Kędziera, D. Modeling the electronic structures of the ground and excited states of the ytterbium atom and the ytterbium dimer: A modern quantum chemistry perspective. *Int. J. Quantum Chem.* **2019**, *119*, No. e25983.
- (98) Salter, E.; Trucks, G. W.; Bartlett, R. J. Analytic energy derivatives in many-body methods. I. First derivatives. *J. Chem. Phys.* **1989**, *90*, 1752–1766.
- (99) Kraka, E.; Gauss, J.; Cremer, D. Determination and use of response densities. *J. Mol. Struct.: THEOCHEM* **1991**, *234*, 95–126.
- (100) Bokhan, D.; Trubnikov, D. N.; Bartlett, R. J. Electric multipole moments calculation with explicitly correlated coupled-cluster wavefunctions. *J. Chem. Phys.* **2016**, *144*, 234107.
- (101) Bozkaya, U. Orbital-optimized linearized coupled-cluster doubles with density-fitting and Cholesky decomposition approximations: an efficient implementation. *Phys. Chem. Chem. Phys.* **2016**, *18*, 11362–11373.
- (102) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular electronic-structure theory*; Wiley: New York, 2000.
- (103) Liu, X.; Meijer, G.; Pérez-Ríos, J. A data-driven approach to determine dipole moments of diatomic molecules. *Phys. Chem. Chem. Phys.* **2020**, *22*, 24191–24200.
- (104) Soper, P.; Legon, A.; Flygare, W. Microwave rotational spectrum, molecular geometry, and intermolecular interaction potential of the hydrogen-bonded dimer $\text{OC} \cdots \text{HCl}$. *J. Chem. Phys.* **1981**, *74*, 2138–2142.
- (105) Cummins, P. L.; Bacskay, G. B.; Hush, N. S. Ab initio quantum chemical studies of the electronic properties of the hydrogen bonded N_2 –HF, N_2 –HCl, $(\text{HCN})_2$ and NH_3 –HCN complexes. *Chem. Phys.* **1987**, *115*, 325–337.
- (106) Dunning, T., Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (107) Peterson, K. A.; Dunning Jr, T. H. Accurate correlation consistent basis sets for molecular core-valence correlation effects: The second row atoms Al–Ar, and the first row atoms B–Ne revisited. *J. Chem. Phys.* **2002**, *117*, 10548–10560.
- (108) Boguslawski, K.; Brzęk, F.; Chakraborty, R.; Cieślak, K.; Jahani, S.; Leszczyk, A.; Nowak, A.; Sujkowski, E.; Świerczyński, J.; Ahmadkhani, S.; Kędziera, D.; Kriebel, M. H.; Żuchowski, P. S.; Tecmer, P. PyBEST: improved functionality and enhanced performance. *Comput. Phys. Commun.* **2024**, *297*, 109049.
- (109) Kriebel, M. H.; Tecmer, P.; Gałyńska, M.; Leszczyk, A.; Boguslawski, K. Accelerating Pythonic coupled cluster implementations: a comparison between CPUs and GPUs. *J. Chem. Theory Comput.* **2024**, *20*, 1130–1142.
- (110) DeYonker, N. J.; Peterson, K. A.; Wilson, A. K. Systematically convergent correlation consistent basis sets for molecular core-valence correlation effects: The third-row atoms gallium through krypton. *J. Phys. Chem. A* **2007**, *111*, 11383–11393.
- (111) Hill, J. G.; Peterson, K. A. Gaussian basis sets for use in correlated molecular calculations. XI. Pseudopotential-based and all-electron relativistic basis sets for alkali metal (K–Fr) and alkaline earth (Ca–Ra) elements. *J. Chem. Phys.* **2017**, *147*, 244106.
- (112) Pipek, J.; Mezey, P. G. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.* **1989**, *90*, 4916–4926.
- (113) te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931–967.
- (114) AMS2022, *Theoretical Chemistry*; Vrije Universiteit: Amsterdam, The Netherlands, 2022; see <http://www.scm.com> (accessed March 19, 2024).
- (115) Fonseca Guerra, C.; Snijders, J. G.; te Velde, G.; Baerends, E. J. Towards an order-N DFT method. *Theor. Chem. Acc.* **1998**, *99*, 391–403.
- (116) Jacob, C. R.; Beyhan, S. M.; Bulo, R. E.; Gomes, A. S. P.; Götz, A. W.; Kiewisch, K.; Sikkema, J.; Visscher, L. Software News and Updates PyADF—A Scripting Framework for Multiscale Quantum Chemistry. *J. Comput. Chem.* **2011**, *32*, 2328.
- (117) van Lenthe, E.; Baerends, E. J. Optimized Slater-type basis sets for the elements 1–118. *J. Comput. Chem.* **2003**, *24*, 1142–1156.
- (118) Burke, K.; Perdew, J. P.; Wang, Y. Derivation of a generalized gradient approximation: The PW91 density functional. *Electronic Density Functional Theory: recent progress and new directions* **1998**, 81–111.
- (119) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.
- (120) Lembarki, A.; Chermette, H. Obtaining a gradient-corrected kinetic-energy functional from the Perdew–Wang exchange functional. *Phys. Rev. A* **1994**, *50*, 5328.
- (121) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: A General Purpose Quantum Chemistry Program Package. *WIREs Comput. Mol. Sci.* **2012**, *2*, 242–253.
- (122) Werner, H.-J.; Knowles, P. J.; Manby, F. R.; Black, J. A.; Doll, K.; Heßelmann, A.; Kats, D.; Kohn, A.; Korona, T.; Kreplin, D. A.; Ma, Q.; Miller, T. F.; Mitrushchenkov, A.; Peterson, K. A.; Polyak, I.; Rauhut, G.; Sibaev, M. The Molpro quantum chemistry package. *J. Chem. Phys.* **2020**, *152*, 144107.
- (123) Werner, H.-J.; Knowles, P. J.; et al. MOLPRO, version 2020.2.1, a package of *ab initio* programs. 2020; see <http://www.molpro.net> (accessed March 1, 2024).
- (124) Hampel, C.; Peterson, K. A.; Werner, H.-J. A comparison of the efficiency and accuracy of the quadratic configuration interaction (QCISD), coupled cluster (CCSD), and Brueckner coupled cluster (BCCD) methods. *Chem. Phys. Lett.* **1992**, *190*, 1–12.
- (125) Deegan, M. J.; Knowles, P. J. Perturbative corrections to account for triple excitations in closed and open shell coupled cluster theories. *Chem. Phys. Lett.* **1994**, *227*, 321–326.
- (126) Kats, D.; Manby, F. R. Communication: The distinguishable cluster approximation. *J. Chem. Phys.* **2013**, *139*. DOI: 10.1063/1.4813481.
- (127) Kats, D. Communication: The distinguishable cluster approximation. II. The role of orbital relaxation. *J. Chem. Phys.* **2014**, *141*. DOI: 10.1063/1.4892792.
- (128) Lykhin, A. O.; Truhlar, D. G.; Gagliardi, L. Dipole moment calculations using multiconfiguration pair-density functional theory and hybrid multiconfiguration pair-density functional theory. *J. Chem. Theory Comput.* **2021**, *17*, 7586–7601.
- (129) Samanta, P. K.; Köhn, A. First-order properties from internally contracted multireference coupled-cluster theory with particular focus on hyperfine coupling tensors. *J. Chem. Phys.* **2018**, *149*, 064101.

- (130) Balashov, A. A.; Bielska, K.; Li, G.; Kyuberis, A. A.; Wójtewicz, S.; Domysławska, J.; Ciuryło, R.; Zobov, N. F.; Lisak, D.; Tennyson, J.; et al. Measurement and calculation of CO (7–0) overtone line intensities. *J. Chem. Phys.* **2023**, *158*, 234306.
- (131) Jacob, C. R.; Wesolowski, T. A.; Visscher, L. Orbital-free embedding applied to the calculation of induced dipole moments in CO₂... X (X = He, Ne, Ar, Kr, Xe, Hg) van der Waals complexes. *J. Chem. Phys.* **2005**, *123*, 174104.
- (132) Tecmer, P.; van Lingen, H.; Gomes, A. S. P.; Visscher, L. The electronic spectrum of CUONg₄ (Ng = Ne, Ar, Kr, Xe): New insights in the interaction of the CUO molecule with noble gas matrices. *J. Chem. Phys.* **2012**, *137*, 084308.
- (133) Olejniczak, M.; Antušek, A.; Jaszuński, M. Relativistic frozen density embedding calculations of solvent effects on the nuclear magnetic resonance shielding constants of transition metal nuclei. *Int. J. Quantum Chem.* **2021**, *121*, No. e26789.
- (134) Bytautas, L.; Scuseria, G. E.; Ruedenberg, K. Seniority number description of potential energy surfaces: Symmetric dissociation of water, N₂, C₂, and Be₂. *J. Chem. Phys.* **2015**, *143*. DOI: 10.1063/1.4929904.

Toward Reliable Dipole Moments without Single Excitations: The Role of Orbital Rotations and Dynamical Correlation

Rahul Chakraborty¹, Matheus Morato F. de Moraes¹, Katharina Boguslawski¹,
Artur Nowak¹, Julian Świerczyński², and Paweł Tecmer¹

¹*Institute of Physics, Faculty of Physics, Astronomy and Informatics,
Nicolaus Copernicus University in Torun,
Grudziadzka 5, 87-100 Torun, Poland*

²*Institute of Engineering and Technology, Faculty of Physics, Astronomy, and Informatics,
Nicolaus Copernicus University in Toruń, 87-100 Toruń, Poland*

Corresponding authors: matheusmorat@gmail.com, k.boguslawski@fizyka.umk.pl,
ptecmer@fizyka.umk.pl

Supplementary Information

Table S1: Experimental bond lengths and dipole moments of diatomics studied in this work.

Type	Molecule	Bond length (Å)	μ (D)
Singly-bonded	HF	0.917	1.827
	HCl	1.275	1.109
	LiH	1.596	5.882
	NaH	1.889	6.400
	LiF	1.564	6.327
	NaCl	2.361	9.002
	AlF	1.654	1.515
	ClF	1.628	0.850
	GaF	1.774	2.400
	LiNa	2.810	0.470
Multiply-bonded	CO	1.128	0.112
	CS	1.535	1.958
	CSe	1.676	1.990
	SiO	1.510	3.098
	SiS	1.730	1.740
	SiSe	2.058	1.100
	GeO	1.625	3.282
	GeS	2.012	2.000
	PN	1.491	2.751
A ²⁺ B ²⁻	MgO	1.749	6.200

Table S2: Dipole moments (μ in D) of main-group diatomics in different methods. Frozen core approximation has been used here for all the methods. CC_u and CC_r denote unrelaxed and relaxed dipole moments of the corresponding CC methods.

Molecule	Basis	pCCD	pCCD -LCCD	pCCD -LCCSD	CCSD _u	CCSD(T) _u	oo-pCCD	oo-pCCD -LCCD	oo-pCCD -LCCSD	CCSD _r	CCSD(T) _r
HF	cc-pVDZ	1.915	1.899	1.814	1.822	1.817	1.861	1.857	1.861	1.830	1.819
	cc-pVTZ	1.915	1.889	1.790	1.813	1.800	1.872	1.861	1.838	1.823	1.807
	cc-pVQZ	1.907	1.880	1.789	1.815	1.801	1.866	1.852	1.832	1.826	1.809
	aug-cc-pVDZ	1.923	1.890	1.746	1.787	1.781	1.857	1.845	1.800	1.803	1.788
	aug-cc-pVTZ	1.915	1.880	1.749	1.794	1.782	1.860	1.845	1.806	1.808	1.791
	aug-cc-pVQZ	1.913	1.878	1.758	1.802	1.788	1.857	1.842	1.815	1.814	1.797
HCl	cc-pVDZ	1.386	1.378	1.339	1.329	1.312	1.344	1.346	1.404	1.323	1.312
	cc-pVTZ	1.242	1.228	1.173	1.184	1.172	1.215	1.211	1.239	1.182	1.174
	cc-pVQZ	1.210	1.195	1.145	1.150	1.139	1.194	1.189	1.196	1.150	1.142
	aug-cc-pVDZ	1.212	1.199	1.126	1.144	1.128	1.198	1.196	1.190	1.146	1.132
	aug-cc-pVTZ	1.175	1.147	1.049	1.093	1.081	1.158	1.149	1.126	1.097	1.086
	aug-cc-pVQZ	1.163	1.138	1.061	1.108	1.096	1.153	1.146	1.139	1.110	1.099
LiH	cc-pVDZ	5.903	5.857	5.681	5.735	5.735	5.721	5.753	5.951	5.735	5.735
	cc-pVTZ	5.931	5.889	5.764	5.848	5.848	5.829	5.855	6.015	5.848	5.848
	cc-pVQZ	5.964	5.924	5.799	5.855	5.855	5.854	5.876	6.025	5.856	5.856

Continued on next page

Table S2 – *Continued from previous page*

Molecule	Basis	pCCD	pCCD -LCCD	pCCD -LCCSD	CCSD _u	CCSD(T) _u	oo-pCCD	oo-pCCD -LCCD	oo-pCCD -LCCSD	CCSD _r	CCSD(T) _r
LiH	aug-cc-pVDZ	5.995	5.950	5.782	5.911	5.911	5.912	5.936	6.088	5.912	5.912
	aug-cc-pVTZ	5.972	5.932	5.811	5.872	5.872	5.873	5.894	6.037	5.873	5.873
	aug-cc-pVQZ	5.982	5.944	5.818	5.868	5.868	5.869	5.889	6.031	5.869	5.869
NaH	cc-pVDZ	6.959	6.871	6.557	6.423	6.423	6.610	6.657	7.011	6.424	6.424
	cc-pVTZ	6.959	6.871	6.557	6.615	6.615	6.610	6.657	7.011	6.616	6.616
	cc-pVQZ	7.002	6.913	6.571	6.665	6.665	6.679	6.721	7.049	6.666	6.666
	aug-cc-pVDZ	7.029	6.940	6.533	6.678	6.678	6.679	6.724	7.065	6.679	6.679
	aug-cc-pVTZ	7.028	6.940	6.617	6.692	6.692	6.693	6.734	7.058	6.693	6.693
	aug-cc-pVQZ	7.027	6.940	6.588	6.687	6.687	6.696	6.736	7.052	6.689	6.689
LiF	cc-pVDZ	6.464	6.427	5.991	6.195	6.173	6.363	6.350	6.140	6.232	6.179
	cc-pVTZ	6.406	6.372	6.063	6.273	6.245	6.358	6.345	6.194	6.296	6.251
	cc-pVQZ	6.413	6.384	6.139	6.285	6.260	6.379	6.367	6.247	6.304	6.266
	aug-cc-pVDZ	6.490	6.469	6.250	6.369	6.359	6.461	6.452	6.337	6.389	6.364
	aug-cc-pVTZ	6.451	6.424	6.202	6.322	6.304	6.411	6.399	6.295	6.341	6.311
	aug-cc-pVQZ	6.442	6.415	6.204	6.324	6.304	6.407	6.396	6.317	6.342	6.310
NaCl	cc-pVDZ	9.324	9.264	8.841	9.012	8.949	9.252	9.230	9.038	9.030	8.955
	cc-pVTZ	9.182	9.128	8.834	9.047	8.997	9.201	9.184	9.057	9.057	9.000
	cc-pVQZ	9.225	9.177	8.956	9.138	9.099	9.201	9.194	9.209	9.146	9.100
	aug-cc-pVDZ	9.350	9.321	9.088	9.231	9.199	9.346	9.339	9.225	9.242	9.203
	aug-cc-pVTZ	9.305	9.263	9.028	9.191	9.158	9.305	9.291	9.199	9.200	9.161
	aug-cc-pVQZ	9.264	9.224	9.038	9.179	9.146	9.278	9.266	9.203	9.186	9.148
AlF	cc-pVDZ	1.559	1.498	1.123	1.280	1.261	1.519	1.469	1.111	1.329	1.278
	cc-pVTZ	1.372	1.348	1.226	1.286	1.265	1.375	1.347	1.179	1.314	1.279
	cc-pVQZ	1.383	1.359	1.300	1.366	1.351	1.403	1.379	1.275	1.387	1.365
	aug-cc-pVDZ	1.433	1.418	1.352	1.430	1.431	1.496	1.475	1.314	1.458	1.442
	aug-cc-pVTZ	1.394	1.383	1.359	1.394	1.387	1.424	1.401	1.300	1.416	1.401
	aug-cc-pVQZ	1.370	1.342	1.315	1.400	1.392	1.409	1.387	1.325	1.419	1.406
ClF	cc-pVDZ	1.303	1.265	0.962	1.094	1.073	1.109	1.091	1.059	1.128	1.091
	cc-pVTZ	1.119	1.069	0.799	0.968	0.921	0.979	0.954	0.907	1.002	0.951
	cc-pVQZ	1.059	1.002	0.741	0.915	0.861	0.929	0.900	0.841	0.948	0.895
	aug-cc-pVDZ	1.073	1.013	0.738	0.884	0.865	0.875	0.853	0.832	0.923	0.890
	aug-cc-pVTZ	1.072	1.004	0.738	0.898	0.854	0.943	0.917	0.833	0.934	0.886
	aug-cc-pVQZ	1.072	1.000	0.730	0.896	0.845	0.900	0.872	0.823	0.929	0.879
GaF	cc-pVDZ	2.255	2.192	1.830	2.021	1.963	2.260	2.207	1.844	2.030	1.943
	cc-pVTZ	2.172	2.130	1.926	2.098	2.043	2.216	2.181	1.963	2.079	2.009
	cc-pVQZ	2.198	2.167	2.057	2.195	2.147	2.26	2.228	2.084	2.170	2.112
	aug-cc-pVDZ	2.301	2.281	2.187	2.318	2.290	2.397	2.370	2.172	2.307	2.261
	aug-cc-pVTZ	2.244	2.222	2.122	2.244	2.207	2.309	2.278	2.114	2.219	2.171
	aug-cc-pVQZ	2.224	2.192	2.113	2.251	2.212	2.295	2.265	2.138	2.223	2.176
LiNa	cc-pVDZ	0.515	0.559	0.939	0.898	0.880	0.898	0.869	0.648	0.900	0.879
	cc-pVTZ	0.562	0.600	0.976	0.892	0.721	0.889	0.858	0.651	0.894	0.725
	cc-pVQZ	0.570	0.609	1.001	0.898	0.671	0.895	0.862	0.661	0.900	0.677
	aug-cc-pVDZ	0.558	0.576	0.846	0.892	0.864	0.886	0.857	0.644	0.895	0.863
	aug-cc-pVTZ	0.603	0.670	1.071	0.894	0.716	0.893	0.860	0.652	0.896	0.720
	aug-cc-pVQZ	0.590	0.646	1.069	0.898	0.653	0.901	0.868	0.661	0.900	0.661
CO	cc-pVDZ	0.144	0.069	0.494	0.252	0.262	0.124	0.186	0.474	0.172	0.221
	cc-pVTZ	0.164	0.086	0.446	0.178	0.210	0.089	0.146	0.397	0.106	0.167
	cc-pVQZ	0.189	0.110	0.406	0.135	0.169	0.078	0.132	0.352	0.065	0.125
	aug-cc-pVDZ	0.265	0.165	0.421	0.172	0.182	0.073	0.132	0.386	0.097	0.141

Continued on next page

Table S2 – *Continued from previous page*

Molecule	Basis	pCCD	pCCD -LCCD	pCCD -LCCSD	CCSD _u	CCSD(T) _u	oo-pCCD	oo-pCCD -LCCD	oo-pCCD -LCCSD	CCSD _r	CCSD(T) _r
CO	aug-cc-pVTZ	0.278	0.184	0.388	0.141	0.170	0.074	0.129	0.355	0.070	0.126
	aug-cc-pVQZ	0.280	0.191	0.368	0.128	0.161	0.078	0.131	0.344	0.059	0.117
CS	cc-pVDZ	1.583	1.641	2.350	1.935	1.928	1.785	1.860	2.241	1.843	1.879
	cc-pVTZ	1.639	1.692	2.397	1.959	1.949	1.917	1.968	2.273	1.878	1.918
	cc-pVQZ	1.685	1.728	2.419	1.979	1.963	1.963	2.009	2.288	1.898	1.937
	aug-cc-pVDZ	1.615	1.677	2.423	2.021	2.014	1.851	1.924	2.323	1.936	1.971
	aug-cc-pVTZ	1.611	1.678	2.433	2.010	1.999	1.977	2.026	2.313	1.928	1.969
	aug-cc-pVQZ	1.626	1.680	2.416	2.000	1.984	1.983	2.028	2.305	1.919	1.959
CSe	cc-pVDZ	1.838	1.912	2.849	2.255	2.215	2.078	2.176	2.671	2.126	2.148
	cc-pVTZ	1.825	1.908	2.911	2.275	2.233	2.197	2.274	2.726	2.153	2.185
	cc-pVQZ	1.857	1.927	2.927	2.268	2.222	2.226	2.297	2.718	2.146	2.178
	aug-cc-pVDZ	1.793	1.890	2.919	2.335	2.293	2.167	2.262	2.741	2.214	2.234
	aug-cc-pVTZ	1.768	1.871	2.926	2.304	2.260	2.240	2.315	2.737	2.183	2.213
	aug-cc-pVQZ	1.763	1.848	2.878	2.277	2.230	2.238	2.309	2.718	2.155	2.187
SiO	cc-pVDZ	3.309	3.135	1.893	2.534	2.405	3.070	2.944	2.112	2.697	2.506
	cc-pVTZ	3.456	3.297	2.173	2.881	2.698	3.375	3.268	2.390	3.009	2.791
	cc-pVQZ	3.610	3.447	2.310	3.089	2.908	3.531	3.430	2.573	3.207	2.998
	aug-cc-pVDZ	3.613	3.438	2.162	2.976	2.850	3.506	3.398	2.461	3.115	2.933
	aug-cc-pVTZ	3.662	3.502	2.301	3.109	2.943	3.574	3.472	2.572	3.232	3.032
	aug-cc-pVQZ	3.686	3.527	2.343	3.157	2.984	3.584	3.484	2.626	3.273	3.073
SiS	cc-pVDZ	1.016	0.937	0.427	0.731	0.632	1.365	1.238	0.418	0.794	0.684
	cc-pVTZ	1.091	1.002	0.565	0.903	0.801	1.396	1.293	0.564	0.947	0.840
	cc-pVQZ	1.044	0.975	0.646	0.952	0.862	1.380	1.285	0.623	0.993	0.898
	aug-cc-pVDZ	1.088	0.995	0.517	0.853	0.768	1.433	1.325	0.532	0.907	0.815
	aug-cc-pVTZ	1.111	1.010	0.563	0.920	0.829	1.380	1.283	0.582	0.965	0.869
	aug-cc-pVQZ	1.106	1.021	0.638	0.955	0.869	1.292	1.212	0.661	0.997	0.904
SiSe	cc-pVDZ	1.786	1.609	0.362	0.997	0.840	2.174	1.983	0.488	1.167	0.977
	cc-pVTZ	1.786	1.609	0.362	1.243	1.064	2.174	1.983	0.488	1.388	1.183
	cc-pVQZ	1.797	1.628	0.465	1.337	1.166	2.190	2.007	0.571	1.473	1.278
	aug-cc-pVDZ	1.828	1.616	0.278	1.156	1.010	2.254	2.048	0.462	1.310	1.134
	aug-cc-pVTZ	1.861	1.652	0.321	1.261	1.095	2.146	1.960	0.501	1.404	1.212
	aug-cc-pVQZ	1.860	1.659	0.399	1.339	1.171	2.178	1.998	0.573	1.474	1.282
GeO	cc-pVDZ	3.752	3.524	1.697	2.712	2.516	3.214	3.072	2.149	2.926	2.663
	cc-pVTZ	3.893	3.699	2.027	3.092	2.833	3.569	3.451	2.502	3.260	2.959
	cc-pVQZ	4.069	3.857	2.067	3.311	3.051	3.739	3.628	2.697	3.462	3.168
	aug-cc-pVDZ	4.120	3.921	2.175	3.287	3.096	3.805	3.687	2.669	3.467	3.210
	aug-cc-pVTZ	4.168	3.957	2.052	3.364	3.126	3.815	3.703	2.732	3.521	3.241
	aug-cc-pVQZ	4.174	3.961	2.043	3.406	3.157	3.819	3.709	2.776	3.553	3.270
GeS	cc-pVDZ	2.514	2.325	0.964	1.730	1.523	2.900	2.675	1.141	1.896	1.672
	cc-pVTZ	2.625	2.432	1.131	1.993	1.765	2.956	2.779	1.298	2.120	1.887
	cc-pVQZ	2.608	2.436	1.264	2.116	1.902	2.978	2.812	1.419	2.233	2.014
	aug-cc-pVDZ	2.685	2.482	1.133	2.008	1.814	3.070	2.882	1.348	2.151	1.944
	aug-cc-pVTZ	2.646	2.459	1.170	2.075	1.866	2.983	2.812	1.362	2.201	1.985
	aug-cc-pVQZ	2.647	2.470	1.250	2.137	1.929	2.981	2.819	1.435	2.253	2.039
PN	cc-pVDZ	2.578	2.486	1.968	2.394	2.290	2.493	2.422	2.155	2.474	2.368
	cc-pVTZ	2.868	2.736	2.140	2.638	2.482	2.704	2.634	2.377	2.714	2.582
	cc-pVQZ	2.984	2.844	2.247	2.762	2.595	2.795	2.726	2.488	2.833	2.700
	aug-cc-pVDZ	2.931	2.794	2.223	2.688	2.581	2.782	2.715	2.419	2.771	2.660
	aug-cc-pVTZ	3.023	2.871	2.240	2.765	2.614	2.821	2.752	2.478	2.843	2.716

Continued on next page

Table S2 – *Continued from previous page*

Molecule	Basis	pCCD	pCCD -LCCD	pCCD -LCCSD	CCSD _u	CCSD(T) _u	oo-pCCD	oo-pCCD -LCCD	oo-pCCD -LCCSD	CCSD _r	CCSD(T) _r
PN	aug-cc-pVQZ	3.062	2.904	2.290	2.807	2.643	2.830	2.762	2.523	2.880	2.749
MgO	cc-pVDZ	8.366	8.041	4.423	4.934	3.838	4.158	4.122	5.758	4.600	4.952
	cc-pVTZ	8.377	8.036	3.962	5.996	4.696	4.109	4.075	5.940	5.721	5.706
	cc-pVQZ	8.430	8.089	3.752	6.576	5.256	3.949	3.923	5.996	6.316	6.196
	aug-cc-pVDZ	8.366	8.041	4.423	6.573	5.341	4.158	4.122	5.758	6.349	6.294
	aug-cc-pVTZ	8.377	8.036	3.962	6.759	5.470	4.109	4.075	5.940	6.513	6.404
	aug-cc-pVQZ	8.430	8.089	3.752	6.911	5.608	3.949	3.923	5.996	6.664	6.510

TableS3: Error analysis w.r.t. experimental values for the dataset of 20 main group diatomics studied in this work. Errors are calculated as $|\mu_{\text{Method}} - \mu_{\text{Exp}}|$

Quantity	Basis	pCCD	pCCD -LCCD	pCCD -LCCSD	CCSD _u	CCSD(T) _u	oo-pCCD	oo-pCCD -LCCD	oo-pCCD -LCCSD	CCSD _r	CCSD(T) _r
MUE	cc-pVDZ	0.385	0.325	0.600	0.303	0.399	0.334	0.347	0.493	0.278	0.313
	cc-pVTZ	0.387	0.316	0.552	0.181	0.273	0.339	0.325	0.409	0.183	0.204
	cc-pVQZ	0.412	0.336	0.520	0.169	0.203	0.363	0.348	0.365	0.181	0.143
	aug-cc-pVDZ	0.426	0.344	0.498	0.170	0.222	0.355	0.337	0.403	0.164	0.157
	aug-cc-pVTZ	0.437	0.356	0.530	0.180	0.194	0.363	0.345	0.373	0.191	0.148
	aug-cc-pVQZ	0.442	0.363	0.520	0.193	0.173	0.376	0.357	0.349	0.205	0.143
RMSE	cc-pVDZ	0.599	0.513	0.787	0.441	0.655	0.574	0.560	0.610	0.457	0.452
	cc-pVTZ	0.601	0.506	0.766	0.260	0.442	0.585	0.565	0.512	0.266	0.290
	cc-pVQZ	0.631	0.531	0.763	0.255	0.322	0.627	0.606	0.463	0.251	0.234
	aug-cc-pVDZ	0.633	0.530	0.688	0.269	0.336	0.607	0.579	0.505	0.252	0.258
	aug-cc-pVTZ	0.640	0.535	0.753	0.276	0.300	0.604	0.581	0.476	0.262	0.244
	aug-cc-pVQZ	0.651	0.545	0.764	0.290	0.271	0.636	0.614	0.447	0.276	0.240

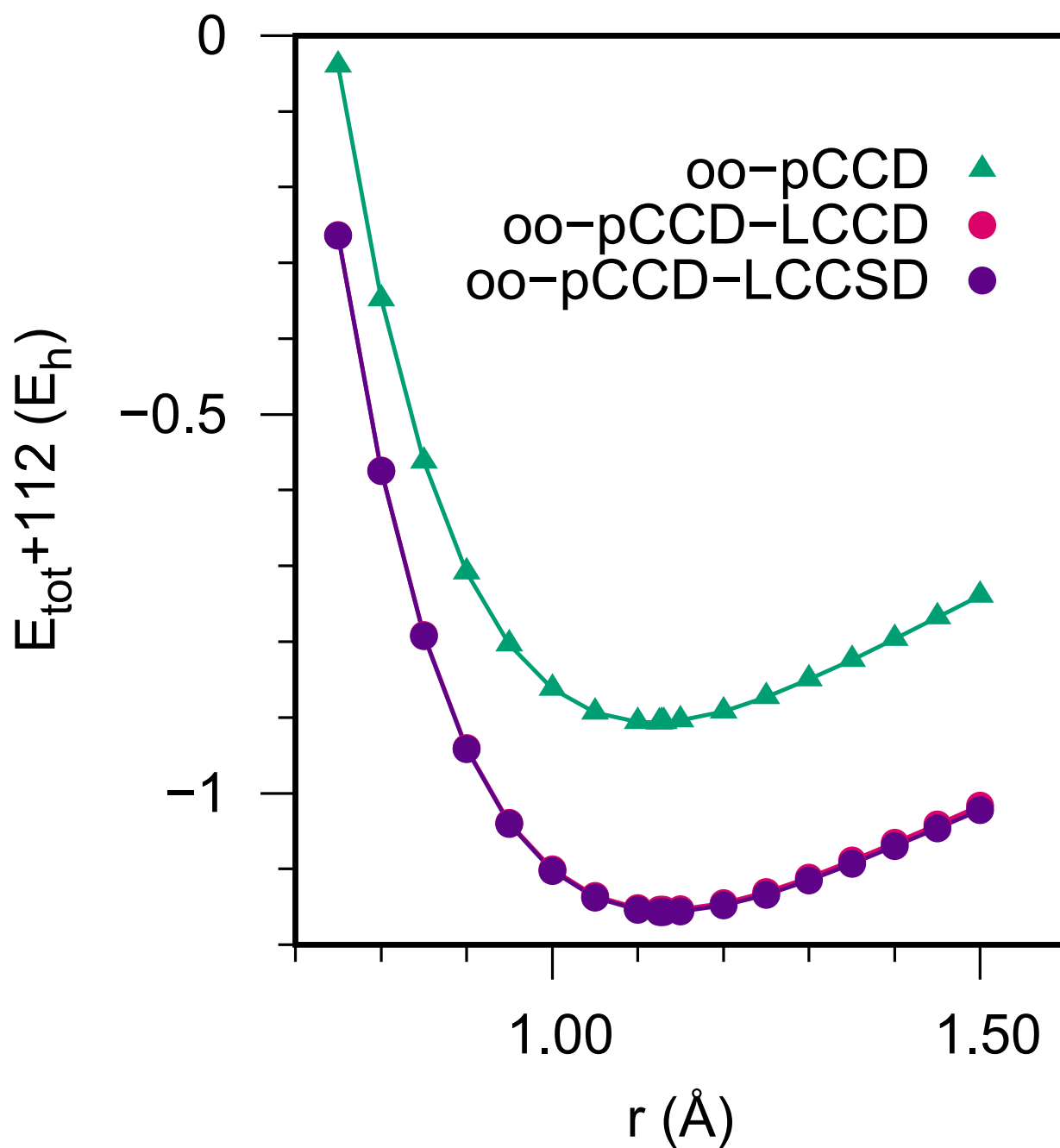


Figure S1: Potential energy surface of CO with oo-pCCD method and its dynamic correlation correction variants in aug-cc-pVTZ basis.

Table S4: Dipole moment (μ in D) of diatomic molecules in aug-cc-pVQZ basis set without frozen core. CC_u and CC_r denote unrelaxed and relaxed dipole moments of the corresponding CC methods.

Species	pCCD	pCCD-LCCD	pCCD-LCCSD	CCSD _u	CCSD(T) _u	oo-pCCD	oo-pCCD-LCCD	oo-pCCD-LCCSD	CCSD _r	CCSD(T) _r
HF	1.913	1.878	1.764	1.806	1.792	1.855	1.840	1.819	1.818	1.800
HCl	1.163	1.138	1.055	1.104	1.092	1.151	1.144	1.133	1.107	1.095
LiH	5.984	5.945	5.798	5.854	5.851	5.912	5.927	5.994	5.856	5.851
NaH	7.028	6.946	6.452	6.577	6.551	6.740	6.778	6.960	6.578	6.553
LiF	6.444	6.416	6.186	6.311	6.288	6.413	6.399	6.290	6.328	6.295
NaCl	9.265	9.232	8.990	9.144	9.102	9.285	9.274	9.153	9.154	9.108
AlF	1.371	1.345	1.379	1.439	1.432	1.394	1.373	1.374	1.456	1.446
ClF	1.072	1.001	0.744	0.908	0.853	0.897	0.868	0.839	0.940	0.887
GaF	2.224	2.188	2.090	2.270	2.252	2.295	2.261	2.117	2.309	2.267
LiNa	0.587	0.649	0.800	0.678	0.653	0.897	0.882	0.473	0.676	0.662
CO	0.280	0.191	0.376	0.131	0.166	0.084	0.139	0.352	0.062	0.122
CS	1.629	1.684	2.469	2.024	2.006	1.996	2.043	2.349	1.939	1.981
CSe	1.764	1.850	2.891	2.252	2.204	2.266	2.334	2.710	2.134	2.170
SiO	3.687	3.528	2.378	3.193	3.013	3.686	3.576	2.545	3.304	3.102
SiS	1.107	1.022	0.655	0.976	0.885	1.364	1.273	0.630	1.016	0.922
SiSe	1.899	1.683	0.422	1.428	1.252	2.068	1.892	0.630	1.553	1.350
GeO	4.176	3.957	1.822	3.377	3.114	3.862	3.735	2.401	3.550	3.252
GeS	2.649	2.473	1.056	2.081	1.850	3.081	2.901	1.007	2.214	1.982
PN	3.063	2.907	2.327	2.839	2.664	2.826	2.760	2.559	2.907	2.773
MgO	8.430	8.088	4.139	6.859	5.526	3.510	3.551	5.232	6.660	6.484

Figure S2: Distance dependence of the calculated dipole moment components in aug-cc-pVTZ basis for the $H_2O \cdots Rg$ [$Rg = He, Ne, Ar, \text{ and } Kr$] complexes.

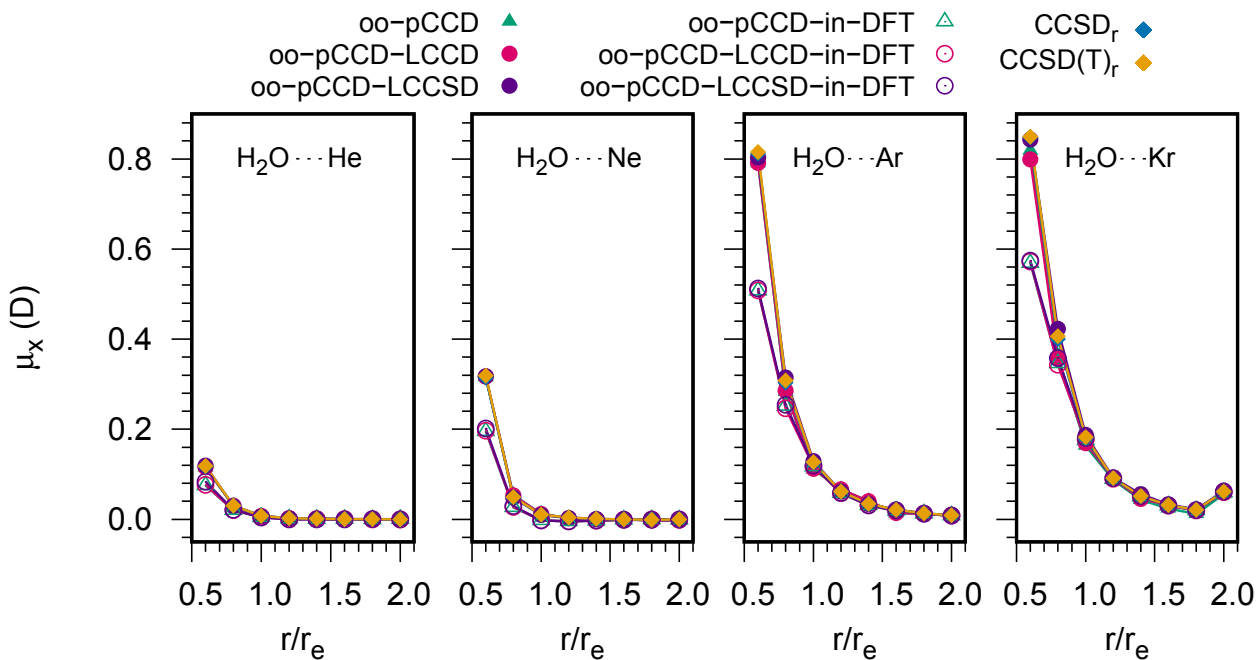


Table S5: Dipole moment surface of the HF molecule in aug-cc-pVTZ basis.

r (Å)	oo-pCCD	oo-pCCD-LCCD	oo-pCCD-LCCSD	CCSD _r	CCSD(T) _r
0.75	-1.606	-1.594	-1.541	-1.548	-1.538
0.80	-1.681	-1.668	-1.616	-1.625	-1.613
0.85	-1.758	-1.745	-1.697	-1.703	-1.688
0.90	-1.835	-1.820	-1.778	-1.781	-1.765
0.917	-1.860	-1.845	-1.806	-1.808	-1.791
0.95	-1.908	-1.892	-1.862	-1.860	-1.841
1.00	-1.974	-1.958	-1.946	-1.938	-1.917
1.10	-2.080	-2.065	-2.112	-2.088	-2.060
1.15	-2.100	-2.089	-2.200	-2.158	-2.127
1.175	-2.111	-2.100	-2.239	-2.191	-2.158
1.20	-2.118	-2.108	-2.278	-2.223	-2.188
1.30	-2.134	-2.127	-2.407	-2.334	-2.291
1.35	-2.088	-2.085	-2.475	-2.378	-2.330
1.40	-2.048	-2.048	-2.524	-2.412	-2.358
1.45	-1.992	-1.995	-2.563	-2.436	-2.374
1.50	-1.921	-1.928	-2.589	-2.447	-2.377
1.55	-1.801	-1.812	-2.581	-2.446	-2.366
1.575	-1.791	-1.803	-2.605	-2.440	-2.354
1.60	-1.742	-1.756	-2.603	-2.430	-2.338
1.65	-1.637	-1.655	-2.590	-2.399	-2.294
1.70	-1.526	-1.547	-2.565	-2.354	-2.234
1.75	-1.412	-1.436	-2.528	-2.293	-2.158
1.80	-1.296	-1.323	-2.482	-2.217	-2.067
1.85	-1.182	-1.211	-2.426	-2.126	-1.963
1.90	-1.071	-1.102	-2.364	-2.023	-1.851
1.95	-0.965	-0.997	-2.298	-1.909	-1.734
2.00	-0.866	-0.898	-2.227	-1.786	-1.619
2.05	-0.773	-0.806	-2.155	-1.656	-1.512
2.10	-0.687	-0.720	-2.082	-1.523	-1.421
2.15	-0.609	-0.642	-2.009	-1.388	-1.353
2.20	-0.538	-0.571	-1.937	-1.253	-1.317
2.25	-0.474	-0.507	-1.867	-1.122	-1.320
2.30	-0.417	-0.449	-1.799	-0.996	-1.368
2.35	-0.366	-0.397	-1.734	-0.877	-1.469
2.40	-0.321	-0.351	-1.671	-0.765	-1.625
2.45	-0.281	-0.311	-1.611	-0.662	-1.838
2.50	-0.246	-0.275	-1.554	-0.567	-2.110
2.75	-0.127	-0.151	-1.311	-0.219	—
3.00	-0.067	-0.086	-1.131	-0.033	—
3.50	-0.022	-0.040	-0.906	—	—
4.00	-0.009	-0.030	-0.803	—	—

Table S6: Dipole moment surface of the CO molecule in aug-cc-pVTZ basis.

r (Å)	oo-pCCD	oo-pCCD-LCCD	oo-pCCD-LCCSD	CCSD _r	CCSD(T) _r
0.75	1.291	1.295	1.240	1.224	1.223
0.80	1.162	1.170	1.135	1.105	1.106
0.85	1.014	1.027	1.025	0.972	0.978
0.90	0.860	0.879	0.905	0.828	0.838
0.95	0.697	0.722	0.781	0.674	0.691
1.00	0.527	0.559	0.656	0.511	0.536
1.05	0.352	0.392	0.534	0.342	0.378
1.10	0.174	0.224	0.417	0.169	0.217
1.125	0.085	0.140	0.362	0.081	0.137
1.128	0.074	0.129	0.355	0.070	0.127
1.13	0.068	0.123	0.351	0.063	0.120
1.15	-0.003	0.056	0.309	-0.008	0.056
1.20	-0.179	-0.107	0.213	-0.186	-0.103
1.25	-0.349	-0.265	0.132	-0.364	-0.259
1.30	-0.513	-0.415	0.069	-0.541	-0.410
1.35	-0.664	-0.55	0.025	-0.716	-0.555
1.40	-0.806	-0.675	0.006	-0.887	-0.693
1.45	-0.935	-0.786	0.014	-1.055	-0.821
1.50	-1.048	-0.879	0.066	-1.218	-0.939

Table S7: Potential energy surface of the CO molecule in aug-cc-pVTZ basis.

r (Å)	oo-pCCD	oo-pCCD-LCCD	oo-pCCD-LCCSD
0.75	-112.039847	-112.263221	-112.264186
0.80	-112.347900	-112.573941	-112.574974
0.85	-112.562476	-112.791373	-112.792613
0.90	-112.708294	-112.940263	-112.941638
0.95	-112.803726	-113.038946	-113.040487
1.00	-112.862264	-113.100881	-113.102623
1.05	-112.893868	-113.136013	-113.137993
1.10	-112.905944	-113.151733	-113.153993
1.125	-112.906435	-113.154085	-113.156501
1.128	-112.906275	-113.154165	-113.156602
1.13	-112.906164	-113.15419	-113.156638
1.15	-112.903983	-113.153519	-113.156103
1.20	-112.892097	-113.145475	-113.148426
1.25	-112.873341	-113.130644	-113.134004
1.30	-112.850002	-113.111302	-113.115103
1.35	-112.823797	-113.089152	-113.093311
1.40	-112.796049	-113.065542	-113.070078
1.45	-112.767685	-113.041273	-113.046226
1.50	-112.739429	-113.016702	-113.022521

Table S8: Equilibrium bond parameters of the $\text{H}_2\text{O} \cdots \text{Rg}$ [$\text{Rg} = \text{He, Ne, Ar, Kr}$] complexes used in this work. [HM10]

Species	r_e (Å)	θ_e
$\text{H}_2\text{O} \cdots \text{He}$	3.186	79°
$\text{H}_2\text{O} \cdots \text{Ne}$	3.228	76°
$\text{H}_2\text{O} \cdots \text{Ar}$	3.651	100°
$\text{H}_2\text{O} \cdots \text{Kr}$	3.884	109.3°

Table S9: Supramolecular dipole moments (μ in D) of the binary complexes in aug-cc-pVTZ basis. CC_u and CC_r denote unrelaxed and relaxed dipole moments of the corresponding CC methods. The errors with respect to reference $\text{CCSD}(\text{T})_r$ values are calculated as $\epsilon = \mu_{\text{Method}} - \mu_{\text{CCSD}(\text{T})_r}$.

Species	oo-pCCD	ϵ	oo-pCCD-LCCD	ϵ	oo-pCCD-LCCSD	ϵ	CCSD_r	ϵ	$\text{CCSD}(\text{T})_r$
CO-HF	2.586	-0.030	2.630	-0.014	2.866	-0.250	2.571	-0.045	2.616
CO-HCl	1.698	-0.007	1.745	0.040	1.983	0.278	1.652	-0.053	1.705
$\text{N}_2\text{-HF}$	2.371	0.047	2.357	0.043	2.329	0.005	2.335	0.011	2.324
$\text{N}_2\text{-HCl}$	1.497	0.064	1.488	0.055	1.466	0.033	1.441	0.008	1.433
$\text{H}_2\text{O} \cdots \text{He}$	1.928	0.107	1.910	0.089	1.835	0.032	1.844	0.023	1.821
$\text{H}_2\text{O} \cdots \text{Ne}$	1.918	0.110	1.899	0.091	1.821	0.013	1.831	0.023	1.808
$\text{H}_2\text{O} \cdots \text{Ar}$	1.905	0.104	1.887	0.086	1.810	0.009	1.822	0.021	1.801
$\text{H}_2\text{O} \cdots \text{Kr}$	1.940	0.096	1.922	0.078	1.863	0.019	1.865	0.021	1.844

References

- [HM10] Anastasios Haskopoulos and George Maroulis. “Interaction electric hyperpolarizability effects in weakly bound $\text{H}_2\text{O} \cdots \text{Rg}$ ($\text{Rg} = \text{He}, \text{Ne}, \text{Ar}, \text{Kr}$ and Xe) complexes.” In: *J. Phys. Chem. A* 114.33 (2010), pp. 8730–8741.

Expectation Value-pCCD-Based Methods for Single-Electron Properties

Rahul Chakraborty, Somayeh Ahmadkhani,* Julian Świerczyński, Katharina Boguslawski, and Paweł Tecmer*



Cite This: *J. Phys. Chem. A* 2025, 129, 6713–6732



Read Online

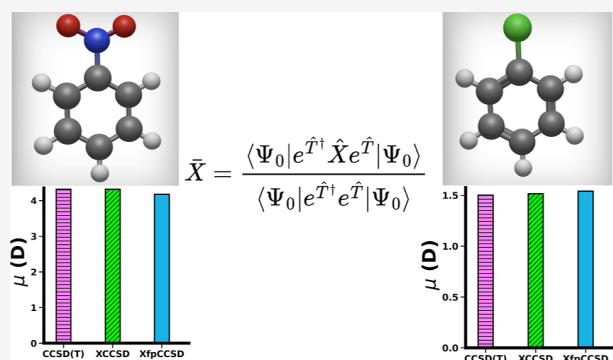
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Expectation-value-coupled cluster theory (XCC) offers a simple avenue for molecular property evaluation. However, its potential has not been fully explored for the new computationally inexpensive CC models, such as pair-coupled cluster doubles (pCCD) and post-pCCD extensions. To that end, we implemented and explored one-electron reduced density matrices in the explicitly connected commutator expansion of the expectation value framework [*J. Chem. Phys.* **2006**, 125, 184109] using pCCD, frozen pair Coupled Cluster (fpCC), and frozen pair linearized Coupled Cluster (fpLCC) variants. The expectation-value-based density matrices are calculated directly using the cluster amplitudes and are computationally cheaper than the corresponding response CC densities, as we bypass solving the computationally expensive Λ -equations. The performance of this approach, when combined with the pCCD-based methods, is assessed against the dipole and quadrupole moments of molecules of a varying chemical nature. We benchmarked our results against the response of CCSD(T) using Hartree–Fock canonical orbitals and variationally optimized pCCD orbitals. Our study highlights that localized pCCD orbitals are a good choice for computing one-electron properties of organic molecules.



1. INTRODUCTION

In the last decades, coupled cluster (CC) theory^{1–6} has become the standard tool for wave function theory (WFT)-based quantum chemical studies.^{5,7,8} Concurrent with the theoretical development of various flavors of CC theory, tremendous progress has also been made toward large-scale applications.^{9–13} In this aspect, CC with domain-based local pair natural orbital (DLPNO), which exploits the localized short-range nature of electron correlation, has been extensively used for large chemical systems.^{13–16} Molecular property calculation with this approach has also been devised.¹⁷ Still, the performance of standard single-reference CC methods leaves room for improvement in cases of strongly correlated systems.^{18,19} Although multireference CC models have been developed,^{6,20,21} they are computationally demanding and dependent on active space choices.

Customizing the CC operator to specifically focus on static correlation, albeit at the expense of losing some dynamic correlation, has been used to circumvent this problem.^{22–24} The antisymmetric product of the 1-reference orbital geminal (AP1roG) method developed by Limacher et al.^{24,25} and later given the moniker pair coupled cluster doubles (pCCD) by Scuseria and co-workers^{26,27} has shown promise along these lines.¹⁹ This approach delineates the zero-seniority block of the Hilbert space and thus performs well for static correlation.^{28,29}

The method offers features such as polynomial scaling, size consistency, and size extensivity (achieved through a variational orbital optimization protocol^{25,30–33}). Augmented with various dynamic correlation correction techniques such as frozen-pair coupled cluster (fpCC)³⁴ and frozen-pair linearized coupled cluster (fpLCC),³⁵ pCCD can treat the electron correlation problem in a balanced way. Through numerous studies, pCCD-based methods have established their reliability for a range of chemical problems, including ground- and excited-state properties.^{19,36–48}

The evaluation of molecular properties, especially multipole moments, using CC theory still constitutes a very active field of research. The calculation of properties using any CC wave function, including pCCD and post-pCCD ones, is not straightforward due to the nonlinear relation between the wave function and the cluster amplitudes. The standard approach in this case has been to calculate the properties as a response of any operator $\hat{X}$ to external perturbation through

Received: June 4, 2025

Revised: June 26, 2025

Accepted: June 27, 2025

Published: July 16, 2025



CC energy derivatives. A generalized form of the Hellmann–Feynman theorem for nonvariational wave functions provides the so-called $\hat{\Lambda}$ operator approach.^{49–52}

$$\bar{X}_{\text{resp}} = \langle 0 | (1 + \hat{\Lambda}) e^{-\hat{T}} \hat{X} e^{\hat{T}} | 0 \rangle \quad (1)$$

Here, $\hat{\Lambda}$ is a de-excitation operator whose amplitudes are solved as zeroth-order Lagrange multipliers, and $\hat{T}$ is the excitation operator. The first-order energy derivative and property calculations using this method are generally known as the linear–response CC method. For example, the dipole moment can be calculated as a response of the atom or molecule to an external electric field. In our previous work, we evaluated the performance of pCCD-based methods in terms of response properties for ground-state electronic dipole moments.⁵³ We observed that orbital optimization improves the overall statistics in favor of pCCD and post-pCCD methods by introducing the effect of partial inclusion of singles (since any variational orbital optimization protocol is equivalent to adding unitary single orbital rotations). The linearized CCD, added to the orbital-optimized pCCD wave function (referred to as oo-pCCD-LCCD), showed the best performance and provided satisfactory agreement with a standard CCSD approach. However, the singles from the linearized Baker–Campbell–Hausdorff (BCH) equation were found to be unsatisfactory in this regard. The nature and extent of the correction produced by LCC also depend to a great extent on the quality of the reference orbital-optimized pCCD wave function. For molecules where the pCCD dipole moments in an optimized-orbital basis are worse than those on a canonical basis, oo-pCCD-LCCD does not produce any significant improvement in dipole moment, whereas oo-pCCD-LCCSD overcompensates the error.

This leaves room for other dynamic correlation correction schemes, especially fpCC, to be tested for their performance, in terms of molecular properties. However, response density matrices are currently not available for the fpCC-type methods. An alternate approach is to use the definition of the expectation value of any observable related to any general operator $\hat{X}$ in a state with wave function Ψ ,

$$\bar{X} = \frac{\langle \Psi | \hat{X} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \quad (2)$$

If the CC wave function is written as $|\Psi\rangle = e^{\hat{T}} |\Psi_0\rangle$, where $|\Psi_0\rangle$ is the reference determinant, the corresponding expectation value expression reads.

$$\bar{X} = \frac{\langle \Psi_0 | e^{\hat{T}^\dagger} \hat{X} e^{\hat{T}} | \Psi_0 \rangle}{\langle \Psi_0 | e^{\hat{T}^\dagger} e^{\hat{T}} | \Psi_0 \rangle} \quad (3)$$

A major roadblock to directly implementing eq 3 is the problem with the truncation of the infinite series of commutators produced by the expansion of the de-excitation operator $\hat{T}^\dagger$. In this work, we take the route proposed by Jezierski and Moszyński,⁵⁴ where they use a different excitation operator $\hat{S}$ ($\hat{S} = \hat{S}_1 + \hat{S}_2 + \hat{S}_3 + \dots$) defined as

$$e^{\hat{S}} |\Psi_0\rangle = \frac{e^{\hat{T}^\dagger} e^{\hat{T}} |\Psi_0\rangle}{\langle \Psi_0 | e^{\hat{T}^\dagger} e^{\hat{T}} | \Psi_0 \rangle} \quad (4)$$

The $\hat{S}$ operator is a connected quantity, as has been shown in ref 47, and satisfies a linear equation consisting of a finite number of multiple commutators of $\hat{T}$ and $\hat{T}^\dagger$. It is to be noted that both response and expectation value approaches converge to the same full configuration interaction (FCI) limit as higher orders of excitations are included in $\hat{T}$ (consequently, in $\hat{S}$, in case of the expectation value approach). The connected (thus, size-extensive) nature of the $\hat{S}$ operator approach has been utilized in a number of works to express CC propagators and calculate various single-electron properties.^{55–57} Specifically, Korona and co-workers have shown that with respect to multipole moments and electrostatic interaction energies, the accuracy of the expectation value approach is close to the response CCSD.^{56,58,59} These works used the conventional CCSD $\hat{T}$ operator for truncating $\hat{S}$. However, this approach has not been tested with other newly devised CC ansätze methods, such as pCCD and post-pCCD methods. To that end, our work focuses on the performance and reliability of the expectation-value CC approach with pCCD and fpCC operators for truncating $\hat{S}$. This provides an alternate route for obtaining density matrices with pCCD and post-pCCD methods and for calculating molecular properties other than the response approach. More importantly, it also provides a way to avoid solving the computationally expensive Λ -equations required to get the density matrices from the response approach (cf. eq 1).

In this work, we have implemented the $\hat{S}$ operator-based expectation-value CC approach with the pCCD and post-pCCD frameworks in the PyBEST software package.^{60,61} Dipole and quadrupole moments of various sets of molecules, such as main-group diatomics and organic molecules, have been calculated on both canonical and variationally optimized pCCD orbital bases for statistical analysis. Specifically, the efficiency of the density matrices obtained with expectation-value pCCD and fpCC has been assessed and compared with expectation-value CCD and CCSD. We compare the performance of these methods with respect to the standard CCSD(*T*) relaxed dipole and quadrupole moments.

2. THEORY

2.1. pCCD and Post-pCCD Methods. In the pCCD ansatz, the cluster operator $\hat{T}_{2(p)}$ contains only the pair-double excitations. The consequent wave function looks as

$$|\Psi_{\text{pCCD}}\rangle = e^{\hat{T}_{2(p)}} |\Psi_0\rangle \quad (5)$$

$$\hat{T}_{2(p)} = \sum_i^{\text{occ}} \sum_a^{\text{virt}} t_{i\bar{i}}^{a\bar{a}} \hat{a}_a^\dagger \hat{a}_{\bar{a}}^\dagger \hat{a}_{\bar{i}} \hat{a}_i \quad (6)$$

where $\hat{a}_q^\dagger$ ($\hat{a}_{\bar{q}}^\dagger$) and $\hat{a}_q$ ($\hat{a}_{\bar{q}}$) are the electron creation and annihilation operators for α -spin (β -spin) for the q -th spin orbital, respectively, and $|\Psi_0\rangle$ is a reference wave function (usually the Hartree–Fock method). As pCCD targets the zero-seniority part of the full wave function, it performs very well in capturing a strong electron correlation of closed-shell electronic structures. When augmented with an orbital optimization protocol, pCCD becomes size-consistent and provides reliable results for strongly correlated systems.^{24,62,63}

However, a major portion of dynamic electron correlation is excluded from the pCCD wave function, as it includes only some specific configurations. A path for correcting this shortcoming is to use the tailored CC theory^{34,64–72}

Table 1. Summary of Investigated CC Models^a

method	CC operator	$ \Psi_0\rangle$	scaling	memory
CCD	$e^{\hat{T}_2}$	$ \Psi_{\text{HF}}\rangle$	$O(o^2v^4)$	$O(v^4)$
CCSD	$e^{\hat{T}_1+\hat{T}_2}$	$ \Psi_{\text{HF}}\rangle$	$O(o^2v^4)$	$O(v^4)$
pCCD	$e^{\hat{T}_{2(p)}}$	$ \Psi_{\text{SD}}\rangle$	$O(o^2v^2)$	$O(N^3)$
fpCCD	$e^{\hat{T}'_2}$	$ \Psi_{\text{pCCD}}\rangle$	$O(o^2v^4)$	$O(v^4)$
fpCCSD	$e^{\hat{T}_1+\hat{T}'_2}$	$ \Psi_{\text{pCCD}}\rangle$	$O(o^2v^4)$	$O(v^4)$
fpLCCD	$1 + \hat{T}'_2$	$ \Psi_{\text{pCCD}}\rangle$	$O(o^2v^4)$	$O(v^4)$
fpLCCSD	$1 + \hat{T}'_1 + \hat{T}'_2$	$ \Psi_{\text{pCCD}}\rangle$	$O(o^2v^4)$	$O(v^4)$

^aReference wave functions (Ψ_0), the form of the CC operator, and formal computational scalings of all the different flavors of CC studied in this work. o and v denote the number of occupied and virtual molecular orbitals, respectively, and $N = o + v$. $\hat{T}'_2$ is a doubles operator for unpaired excitation, that is $\hat{T}'_2 = \hat{T}_2 - \hat{T}_{2(p)}$. Note that we employ a single Slater determinant reference function for pCCD constructed from the optimized pCCD natural orbitals, labeled as $|\Psi_{\text{SD}}\rangle$. Memory indicates the order of the storage bottleneck (excluding a solver-dependent pre-factor).

$$|\Psi_{\text{tCC}}\rangle = e^{\hat{T}_{\text{ext}}}e^{\hat{T}_{\text{int}}}|\Psi_0\rangle \quad (7)$$

which is based on the partitioning of the full cluster operator $\hat{T}$ into $\hat{T}_{\text{int}}$ and $\hat{T}_{\text{ext}}$. First, $\hat{T}_{\text{int}}$, generally derived from a method that describes the strong correlation effect, is applied to the reference wave function $|\Psi_0\rangle$. Subsequently, $\hat{T}_{\text{ext}}$ is applied to this intermediate wave function, with the amplitudes of $\hat{T}_{\text{int}}$ held constant to introduce additional excitations and incorporate dynamic correlation into the expanded wave function. Similarly, the dynamic correlation absent in the pCCD wave function can be accounted for by a posteriori corrections, where single and unpaired double excitations are added to the wave function. One such approach is fpCC, which extends the pCCD reference wave function and adds unpaired doubles (fpCCD) or unpaired doubles along with single excitations (fpCCSD) to $|\Psi_{\text{pCCD}}\rangle$, keeping the cluster amplitudes associated with pair-doubles excitations frozen²⁷

$$|\Psi_{\text{fpCC}}\rangle = e^{\hat{T}_{\text{ext}}}|\Psi_{\text{pCCD}}\rangle \quad (8)$$

The two variants of this approach mentioned above are fpCCD, where $\hat{T}_{\text{ext}} = \hat{T}'_2$, and fpCCSD, where $\hat{T}_{\text{ext}} = \hat{T}'_1 + \hat{T}'_2$, with $\hat{T}'_2$ being the cluster operator for nonpair doubles and $\hat{T}'_1$ being the same for single excitations. These approaches have been used previously to construct potential energy surfaces and calculate the spectroscopic constants of a range of molecules. Notably, Leszczyk et al.³⁴ have shown that fpCCSD performs similarly to CCSD tailored with the density matrix renormalization group (DMRG-tCCSD).

A linearized form of the fpCC, which we denote as fpLCC (also noted as pCCD-LCC in the literature), has also been employed for dynamic correlation correction purposes,

$$|\Psi_{\text{fpLCC}}\rangle \approx (1 + \hat{T}_{\text{ext}})|\Psi_{\text{pCCD}}\rangle \quad (9)$$

In the fpLCC framework,³⁵ the associated Baker–Campbell–Hausdorff expansion is restricted to the second term, i.e.,

$$(\hat{H} + [\hat{H}, \hat{T}'])|\Psi_{\text{pCCD}}\rangle = E|\Psi_{\text{pCCD}}\rangle \quad (10)$$

Like fpCC, fpLCC can have only doubles (fpLCCD) or singles and doubles (fpLCCSD) excitation in the model. Though computationally cheaper, fpLCC can exhibit diverging points in PES and numerically unphysical response dipole moments.^{34,53}

Table 1 provides a summary of all CC models investigated in this work and the concerned reference wave functions, cluster operators, and computational scalings. Though CCD, CCSD, fpCCD, fpCCSD, fpLCCD, and fpLCCSD all appear to have the same computational scaling, they mostly differ in their convergence, which originates from the inclusion/exclusion of nonlinear terms and/or the choice of the reference determinant (HF or pCCD-optimized orbitals). It is to be noted that though pCCD itself has a scaling of $O(o^2v^2)$, the 4-index transformation of the two-electron repulsion integrals during the orbital optimization procedure takes the scaling to $O(N^5)$ unless the Cholesky decomposition or the density fitting is utilized. A subtle aspect that can be argued here is that orbital-optimized pCCD (and the dynamic correlations thereafter), in a strictly theoretical sense, is no longer a post-HF method, as the MO basis itself is entirely different. The reference Slater determinant is no longer the same as $|\Psi_{\text{HF}}\rangle$. Still, for the sake of simplicity and keeping in mind that the pCCD orbital optimization process starts from the HF reference, we treat it under the post-HF class of methods.

A separate important question regarding the property calculation is the difference in computational complexity between the expectation-value method and the response approach. The response approach brings in the additional cost of solving the response Λ -equations. This procedure has the same scaling as solving the amplitude equation, that is, $O(o^2v^4)$. The expectation-value approach entails no additional computational burden, as the equations involve already-calculated CC amplitudes. For orbital-optimized pCCD, the Λ -equations are solved anyway during the orbital optimization process, and hence, the response property calculations do not require any additional cost. However, the same is not true for fpCC and fpLCC. A direct comparison can be made for fpLCC, for which the response framework was previously implemented in PyBEST. For nitrobenzene, our largest test molecule with 261 contracted basis functions and 23 active occupied and 229 active unoccupied orbitals, the Λ -equations for the fpLCCD and fpLCCSD methods took ~ 103 and ~ 204 min, respectively (8 CPUs on Intel(R) Xeon(R) Gold 6336Y CPU at 2.40 GHz).

2.2. Expectation-Value-Based Density Matrices from pCCD and Related Methods. The particular form of the expectation CC model implemented in this work closely follows previous works by Korona.⁵⁹ Combining eqs 3 and 4, the average value of an operator can be written as

$$\bar{X} = \langle e^{\hat{S}} | e^{-\hat{T}} \hat{X} e^{\hat{T}} \rangle = \langle e^{\hat{S}} | e^{-\hat{T}} \hat{X} e^{\hat{T}} e^{-\hat{S}} \rangle \quad (11)$$

This equation can then be rewritten with nested commutators using the BCH formulation

$$e^{-\hat{T}} \hat{X} e^{\hat{T}} = \hat{X} + [\hat{X}, \hat{T}] + \frac{1}{2!} [[\hat{X}, \hat{T}], \hat{T}] + \dots \quad (12)$$

where $\hat{T} = -\hat{T}\hat{S}^\dagger$.

Following Jeziorski and Moszyński,⁵⁴ $\hat{S}$ is expanded ($S = S^{[1]} + S^{[2]} + \dots S^{[k]}$, where k is the highest level of electron excitation

in $\hat{T}$) in terms of $\hat{T}$ and $\hat{T}^\dagger$. If $\hat{T}$ is truncated at the CCSD excitation level, then we arrive at the following equation

$$\begin{aligned}\bar{X} = & \langle \hat{X} \rangle + \langle \hat{S}_1 | \hat{X} \rangle + \langle \hat{X} | \hat{T}_1 \rangle + \langle \hat{S}_2 | [\hat{X}, \hat{T}_2] \rangle \\ & + \langle \hat{S}_1 | [\hat{X}, \hat{T}_2] \rangle + \langle \hat{S}_1 | [\hat{X}, \hat{T}_1] \rangle + \langle \hat{S}_2 | [[\hat{X}, \hat{T}_1], \hat{T}_2] \rangle \\ & + \frac{1}{2} \langle \hat{S}_1^2 | [\hat{X}, \hat{T}_2] \rangle + \frac{1}{2} \langle \hat{S}_1 \hat{S}_2 | [[\hat{X}, \hat{T}_2], \hat{T}_2] \rangle \\ & + \frac{1}{2} \langle \hat{S}_1 | [[\hat{X}, \hat{T}_1], \hat{T}_1] \rangle + \frac{1}{2} \langle \hat{S}_3 | [[\hat{X}, \hat{T}_2], \hat{T}_2] \rangle \\ & + \frac{1}{2} \langle \hat{S}_1^2 | [[\hat{X}, \hat{T}_1], \hat{T}_2] \rangle + \frac{1}{12} \langle \hat{S}_1^3 | [[\hat{X}, \hat{T}_2], \hat{T}_2] \rangle\end{aligned}\quad (13)$$

Ignoring terms higher than $O(\hat{T}^2)$ and with an additional omission of $\langle [\hat{S}_1^{[1]\dagger}, [\hat{X}, \hat{T}_1]] \rangle$, we get

$$\begin{aligned}\bar{X} = & \langle \hat{X} \rangle + \langle [\hat{S}_1^{[1]\dagger}, \hat{X}] \rangle + \langle [\hat{S}_1^{[2]\dagger}, \hat{X}] \rangle + \langle [\hat{X}, \hat{T}_1] \rangle \\ & + \langle [\hat{S}_2^{[1]\dagger}, [\hat{X}, \hat{T}_2]] \rangle + \langle [\hat{S}_1^{[1]\dagger}, [\hat{X}, \hat{T}_2]] \rangle.\end{aligned}\quad (14)$$

The $\hat{S}^{[n]}$ terms are given as

$$\hat{S}_1^{[1]} = \hat{T}_1 \quad (15)$$

$$\hat{S}_2^{[1]} = \hat{T}_2 \quad (16)$$

$$\hat{S}_1^{[2]} = \hat{P}_1([\hat{T}_1^\dagger, \hat{T}_2]) \quad (17)$$

where

$\hat{P}_n(\hat{X}) = (n!)^{-2} \sum_{i_1 \dots i_n} \sum_{a_1 \dots a_n} \langle \hat{a}_{a_1} \hat{a}_{i_1}^\dagger \dots \hat{a}_{a_n} \hat{a}_{i_n}^\dagger \hat{X} \rangle \hat{a}_{a_1}^\dagger \hat{a}_{i_1} \dots \hat{a}_{a_n}^\dagger \hat{a}_{i_n}$ is the superoperator projecting on the space spanned by n -tuple excitation operators. This approximation has been labeled as XCCSD(3) (due to third-order Møller–Plesset terms in the equation). It was shown that the differences between the results produced by full XCCSD and the approximated XCCSD(3) models are insignificant due to the low contribution of the higher-order terms in eq 13. For simplicity, we do not distinguish between the two approaches, and henceforth, we refer to the approximation as the XCCSD.

The correlation contribution to the one-particle reduced density matrix γ can be derived from eq 14 by placing $\hat{X} = E_p^q$, where E_p^q is the singlet excitation operator.

$$E_p^q = \hat{a}_q^\dagger \hat{a}_p + \hat{a}_{\bar{q}}^\dagger \hat{a}_{\bar{p}} \quad (18)$$

The different blocks of γ are then given as

$$\text{occupied} - \text{virtual: } \gamma_i^a = 2(t_i^a + \sum_{jb} t_{ij}^a t_{-jb}^{ab}) = \gamma_a^i \quad (19)$$

$$\text{occupied} - \text{occupied: } \gamma_i^k = - \sum_{jab} t_{ab}^{ij} t_{ab}^{-kj} \quad (20)$$

$$\text{virtual} - \text{virtual: } \gamma_c^a = - \sum_{ijb} t_{ab}^{ij} t_{cb}^{-ij} \quad (21)$$

where $t_{-ij}^{ab} = 2t_{ij}^{ab} - t_{ij}^{ba}$ and t_{i-}^{a-} are the spin-free amplitudes of single, double, etc. excitations.

In this study, we obtain the spin-free single amplitudes t_i^a and double amplitudes t_{ij}^{ab} from the various models discussed

earlier. These amplitudes are then used in eqs 19–21 to construct the corresponding density matrices of a given CC model. Like XCCSD, we label the expectation-value-based approaches as XpCCD, XfpCC, and XfpLCC. XfpCC and XfpLCC have two flavors as discussed in Section 2.1 based on the inclusion of singles excitation. We have also investigated the density matrices obtained using the expectation-value framework for the CCD method, i.e., XCCD. For XpCCD, XCCD, XfpCCD, and XfpLCCD, the γ_i^a terms related to the occupied-virtual block of the density matrix are 0. Dipole and quadrupole moments are computed accordingly after the density matrix.

3. COMPUTATIONAL DETAILS

3.1. Calculation Protocols. The axial components of dipole and quadrupole moments are given as

Table 2. List of Chemical Species Included in the “Small” Set in This Work.^a

diatomic	polyatomic	ionic
H ₂ , LiH, NaH, LiF, NaF, NaCl, HF, HCl, LiNa, ClF, N ₂ , PN, CO, CS, CSe, SiO, SiS, SiSe, GeO, GeS, GeSe	H ₂ O, CO ₂ , NH ₃ , PH ₃	OH, NO ⁺ , CN [−]

^aThe structural parameters are given in the Supporting Information.

$$\mu_\alpha = - \int \rho(\mathbf{r}) r_\alpha dV + \sum_{i=1}^{N_i} Z_i R_{i\alpha} \quad (22)$$

and

$$\begin{aligned}Q_{\alpha\beta} = & - \frac{1}{2} \int \rho(\mathbf{r}) (3r_\alpha r_\beta - \delta_{\alpha\beta} r^2) dV \\ & + \frac{1}{2} \sum_{i=1}^{N_i} Z_i (3R_{i\alpha} R_{i\beta} - \delta_{\alpha\beta} R_i^2)\end{aligned}\quad (23)$$

respectively, where α and β denote the Cartesian axes, ρ is the electron density, Z is the nuclear charge, and r and R are the electronic and nuclear coordinates, respectively. The quadrupole moment also forms a traceless tensor ($Q_{xx} + Q_{yy} + Q_{zz} = 0$). For simplicity, we consider the Q_{zz} component for comparison.

After introducing an atomic orbital (AO) basis set, one α -component of the dipole moment is evaluated from

$$\mu_\alpha = \sum_{i=1}^{N_i} Z_i R_{i\alpha} - \sum_{\mu} \sum_{\nu} \gamma_{\nu}^{\mu} \langle \nu | \mathbf{r}_\alpha | \mu \rangle \quad (24)$$

and the two components of the quadrupole moment from

$$Q_{\alpha\beta} = \sum_{i=1}^{N_i} \frac{3}{2} Z_i R_{i\alpha\beta} - \sum_{\mu} \sum_{\nu} \gamma_{\nu}^{\mu} \langle \nu | \mathbf{q}_{\alpha\beta} | \mu \rangle \quad (25)$$

$$Q_{\alpha\alpha} = \sum_{i=1}^{N_i} \frac{1}{2} Z_i (3R_{i\alpha\alpha} - R_{rr}) \sum_{\mu} \sum_{\nu} \gamma_{\nu}^{\mu} \langle \nu | \mathbf{q}_{\alpha\alpha} | \mu \rangle \quad (26)$$

where γ_{ν}^{μ} is the density matrix in the AO basis, and $\langle \nu | \mathbf{r}_\alpha | \mu \rangle = \int \chi_{\nu}^*(\mathbf{r}) r_\alpha \chi_{\mu}(\mathbf{r}) d\mathbf{r}$ and $\langle \nu | \mathbf{q}_{\alpha\beta} | \mu \rangle = \int \chi_{\nu}^*(\mathbf{r}) \mathbf{q}_{\alpha\beta} \chi_{\mu}(\mathbf{r}) d\mathbf{r}$ are the dipole and quadrupole moment integrals, respectively, expressed in the AO basis $\{\chi_{\nu}\}$,⁷⁴ while $R_{rr} = R_{xx} + R_{yy} + R_{zz}$. Since all pCCD-based

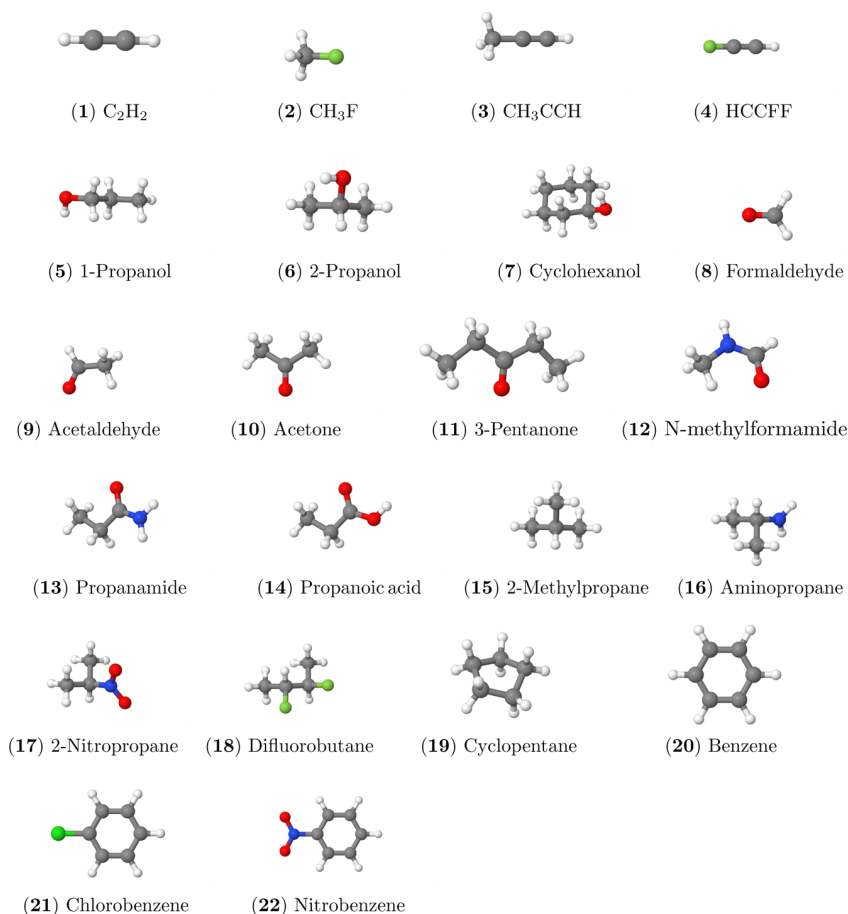


Figure 1. Molecular structures of organic molecules investigated in this work. Coordinates are given in the Supporting Information. The molecular structures were drawn using the Jmol software package.⁷³

methods work in the molecular orbital (MO) basis and hence the corresponding 1-RDMs are defined for the molecular orbitals, we need to perform an AO–MO transformation step of the dipole moment integrals or the 1-RDMs, respectively.

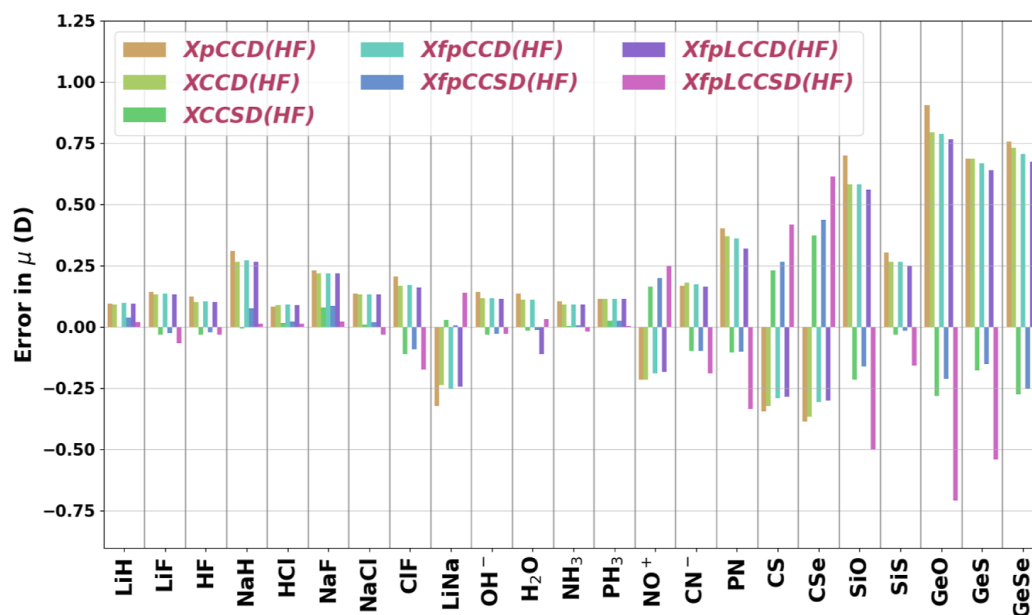
All expectation value CC approach-based calculations, with XpCCD, XCC, XfpCC, and XfpLCC methods, were implemented and performed in the developer version (v1.4.0-dev) of the PyBEST software package.^{60,61,75} The dipole and traceless quadrupole moments were calculated from the expectation value 1-RDM (γ_v^{μ}) using the Sadlej-pVTZ basis set, which is specifically designed for efficient description of electronic properties^{76,77} and shows comparable accuracy as aug-cc-pVTZ.⁷⁸ For the dipole moment surface (DMS), we have used the cc-pVDZ basis set to be able to compare our data with the reference FCI values available in the literature. For the potential dipole moment surface of the HF molecule, we also used a cc-pVDZ basis set to compare our data with the reference FCI values. In this work, we used two different molecular orbital bases, i.e., the canonical HF orbitals and the pCCD-optimized orbitals. Henceforth, all calculations with canonical HF orbitals will be marked as Method(HF), and the same calculations with pCCD optimized orbitals will be marked as Method(OO). Cholesky decomposition with a threshold of 10^{-5} was used in all calculations except for LiF and LiNa, where we faced convergence issues at the SCF level. Pipek–Mezey orbital localization⁷⁹ was used to speed up the orbital optimization process. All nonvalence orbitals were kept frozen for all post-HF calculations. Orbital relaxed CCSD and

CCSD(T) response dipole and quadrupole moment values using canonical HF molecular orbitals were obtained from the MOLPRO2019 package.^{80–82} The Supporting Information gives all of the calculated total dipole moment and quadrupole moment zz-component values.

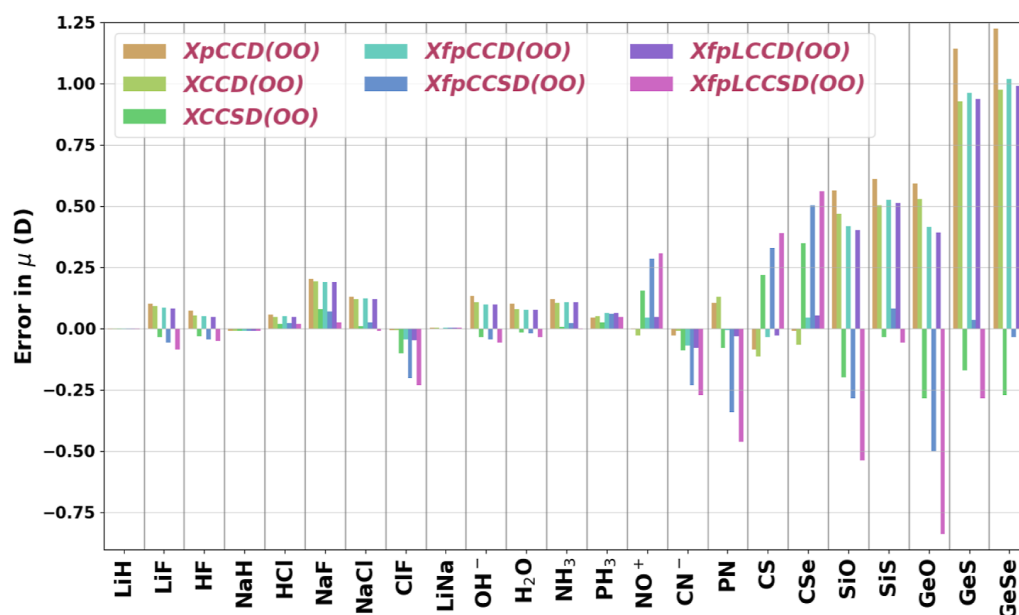
3.2. Test Set of Molecules. We selected 50 molecules covering a wide range of chemical characteristics. Of these, 28 are small molecules (listed in Table 2), including main-group diatomics and small polyatomic inorganic compounds and ions, while 22 are organic molecules (structures shown in Figure 1). Among these, 25 small and 20 organic molecules exhibit a dipole moment. The bond lengths of the main group diatomic molecules were taken from Liu et al.⁸³ and references therein. Coordinates of the larger molecules were taken from Jensen et al.⁸⁴ Each diatomic molecule is placed along the z-axis. For some organic molecules (acetaldehyde, acetone, 3-pentanone, N-methylformamide, propanamide, propanoic acid, cyclopentane, chlorobenzene, and nitrobenzene), the axes were changed in such a way that the largest absolute component of the traceless quadrupole moment always lay on the z-axis. The coordinates of all molecules are provided in the Supporting Information.

4. RESULTS AND DISCUSSION

4.1. Dipole Moment. We start by assessing the accuracy of our dipole moments computed with various pCCD-based methods for different molecular test sets at their equilibrium geometries (Subsection 3.2). Specifically, we focus on the



(a) Error comparison across methods for each molecule in the canonical MO basis.

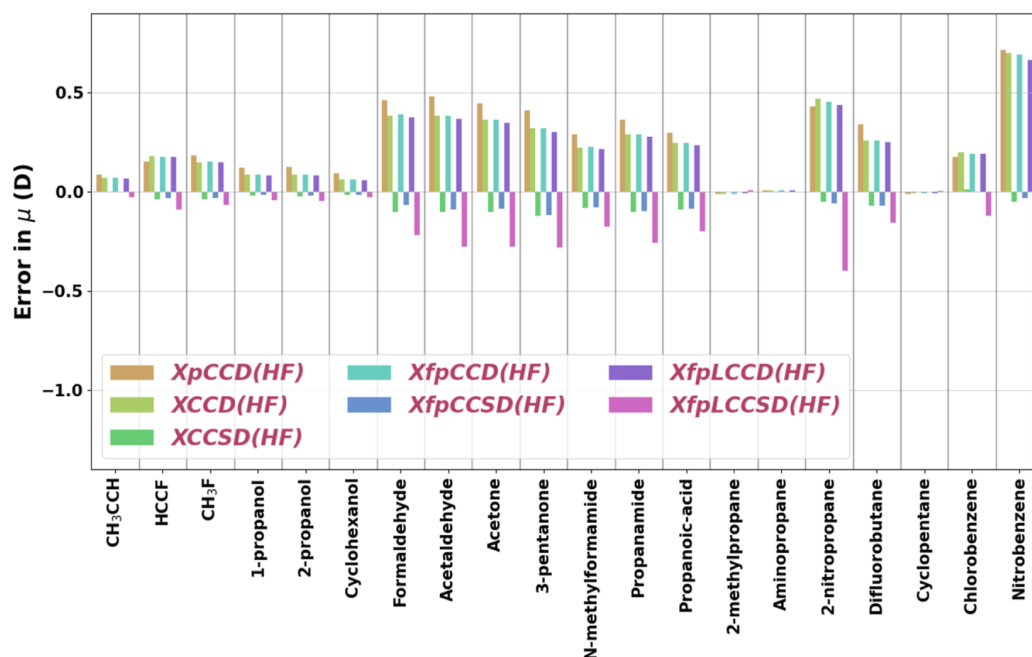


(b) Error comparison across methods for each molecule in the pCCD-optimized MO basis.

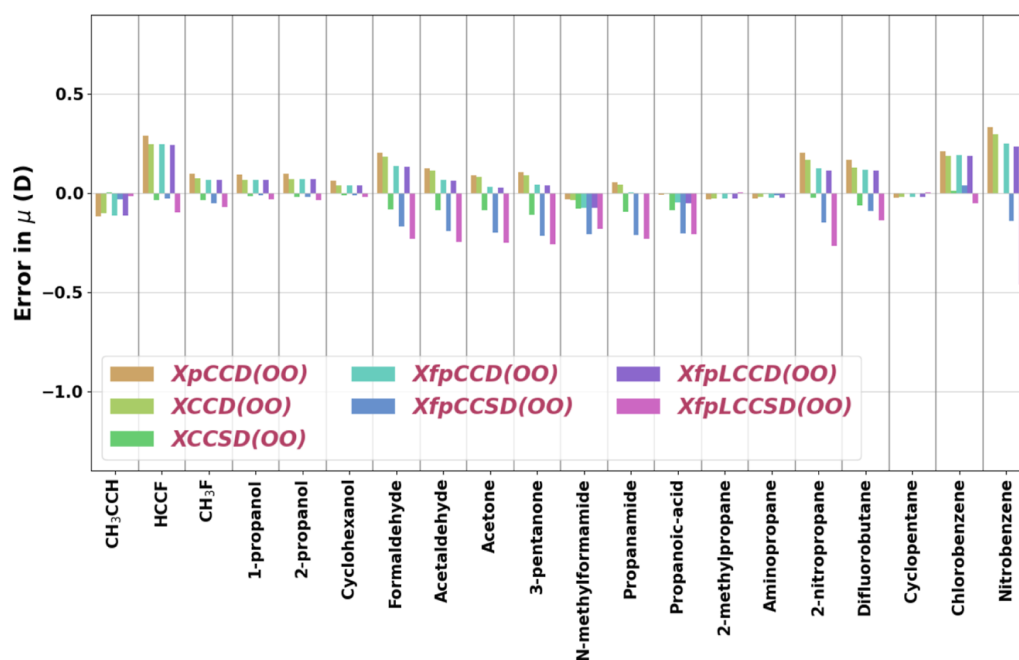
Figure 2. Signed errors with respect to CCSD(T) ($\mu_{\text{Method}} - \mu_{\text{CCSD(T)}}$) in total dipole moments (D) for all methods using the Sadlej-pVTZ basis set for the set of small molecules investigated in this work. (HF) in (a) and (OO) in (b) denote the use of canonical and pCCD-optimized molecular orbitals, respectively.

statistical parameters obtained for the “full set” (combining “small” and “organic” sets) to draw an overall picture. A detailed analysis of the different classes of compounds will be done in Subsection 4.1.1. Figures 2 and 3 show the signed errors calculated against orbital relaxed CCSD(T) response dipole moment values ($\mu_{\text{Method}} - \mu_{\text{CCSD(T)}}$) for individual chemical species for the sets of “small” and “organic”, respectively. The corresponding dipole moment numerical

values are collected in the Supporting Information. Here, a positive error indicates an overestimation of μ , whereas a negative error means vice versa. Additionally, Table S1 of the Supporting Information summarizes the statistical parameters for all the methods calculated against the orbital relaxed CCSD response dipole moment values. In our statistical analysis for dipole moment, we have excluded the most problematic molecules, CO and SiSe, as they showed abruptly high relative



(a) Error comparison across methods for each molecule in the canonical MO basis.



(b) Error comparison across methods for each molecule in the pCCD-optimized MO basis.

Figure 3. Signed errors with respect to CCSD(T) ($\mu_{\text{Method}} - \mu_{\text{CCSD(T)}}$) in total dipole moments (D) for all methods using the Sadlej-pVTZ basis set for the set of organic molecules investigated in this work. (HF) in (a) and (OO) in (b) denote the use of canonical and pCCD-optimized molecular orbitals, respectively.

errors ($(\mu_{\text{Method}} - \mu_{\text{CCSD(T)}})/\mu_{\text{CCSD(T)}}$) for some expectation-value-based methods, which leads to ambiguous statistical parameters. For example, XCCSD shows 119% and 116% relative errors for CO, in canonical and pCCD-optimized orbital basis, respectively, with respect to CCSD(T). SiSe, on the other hand, shows a relative error of 97% and 91% for XpCCD and XfpCCD with pCCD-optimized orbitals. It has been discussed in previous works that for diatomics of group

16 atoms, full triples treatment is required for a reliable prediction of the dipole moment.^{59,85} For such molecules, even the inclusion of perturbative triples may not be sufficient to produce reliable dipole moment values. Hence, particularly for these molecules, CCSD(T) is not a suitable reference.

Table 3 lists the statistical parameters with respect to the response CCSD(T) dipole moments for each of the methods in the two molecular orbital bases. Figure 4 illustrates the

Table 3. Statistical Analysis of Dipole Moments (μ in D) from Various CC- and pCCD-Based Methods w.r.t. Orbital Relaxed Response CCSD(T) Data Using the Sadlej-pVTZ Basis Set

method	RMSE (D)			ME (D)			MAE (D)			MAPE		
	full set	small	organic	full set	small	organic	full set	small	organic	full set	small	organic
XpCCD(HF)	0.356	0.383	0.323	0.225	0.195	0.259	0.285	0.305	0.261	14.9	16.4	13.2
XpCCD(OO)	0.321	0.417	0.149	0.163	0.221	0.096	0.180	0.232	0.119	12.9	13.8	11.8
XCCD(HF)	0.322	0.351	0.285	0.200	0.180	0.224	0.254	0.279	0.225	13.3	14.8	11.4
XCCD(OO)	0.265	0.343	0.127	0.134	0.181	0.080	0.154	0.201	0.100	10.9	11.8	10.0
XCCSD(HF)	0.116	0.147	0.065	-0.034	-0.020	-0.049	0.078	0.102	0.051	3.8	5.2	2.2
XCCSD(OO)	0.109	0.140	0.056	-0.030	-0.021	-0.040	0.071	0.096	0.042	3.6	4.9	2.0
XfpCCD(HF)	0.316	0.342	0.283	0.201	0.182	0.223	0.250	0.272	0.225	13.1	14.6	11.4
XfpCCD(OO)	0.262	0.343	0.113	0.123	0.179	0.059	0.145	0.193	0.089	10.8	11.8	9.7
XfpCCSD(HF)	0.116	0.149	0.058	-0.020	0.001	-0.044	0.075	0.102	0.044	3.6	5.0	2.0
XfpCCSD(OO)	0.177	0.210	0.129	-0.051	-0.014	-0.095	0.120	0.139	0.098	5.8	7.0	4.4
XfpLCCD(HF)	0.304	0.328	0.273	0.193	-0.173	0.214	0.241	0.262	0.216	12.6	14.0	11.1
XfpLCCD(OO)	0.255	0.333	0.109	0.118	0.174	0.055	0.142	0.191	0.085	10.7	11.6	9.6
XfpLCCSD(HF)	0.337	0.318	0.357	-0.139	0.085	-0.201	0.210	0.216	0.203	9.8	11.1	8.3
XfpLCCSD(OO)	0.256	0.304	0.186	-0.109	-0.084	-0.139	0.173	0.202	0.140	8.4	9.8	6.7

comparison of the RMSE for all of the methods. It is evident from Figure 4 that XCCSD produces the lowest statistical error parameters. XCCSD is also agnostic to the molecular orbital basis as its performance improves only to a small extent when pCCD-optimized orbitals are used. XfpCCSD shows the best performance among the frozen pair methods. Employing the canonical HF orbital basis, it approaches the accuracy of XCCSD. For all the methods, the statistical parameters for the “full set” improve upon using pCCD-optimized orbitals, except for fpCCSD. The role of orbital rotations, which act like pseudosingles excitations in the wave function, is prominent when we look at the significant improvements in the statistics for the methods containing only double excitations in the cluster operator, i.e., XCCD, XfpCCD, and XfpLCCD. This indicates an overcorrection due to singles excitation in the case of fpCCSD comes from both orbital rotations during orbital optimization and from the inclusion of $\hat{T}_1$ in $\hat{T}_{\text{pCC}}$. The “correction” effect of adding single excitations on the dipole moment is also prominent when we look at the opposite signs of ME for XCCD and XCCSD, or for XfpCCD and XfpCCSD. Both the doubles-only methods, XCCD and XfpCCD, overestimate the dipole moments (+ sign of ME), whereas XCCSD and XfpCCSD underestimate on average with the - sign of ME. In both cases, the values of RMSE, MAE, and MAPE also improve upon adding single excitations to the method. Hence, for both XCC and XfpCC, adding single excitations in the cluster operator takes the dipole moment values closer to the reference. However, the same does not seem to be true between fpLCCD and fpLCCSD. Introducing singles in the linearized version of the fpCC model worsens the RMSE. Interestingly, XfpLCCSD shows lower MAPE than XfpLCCD. This indicates that the performance of XfpLCCSD becomes worse compared to that of XfpLCCD primarily for molecules with a high magnitude of dipole moment.

The statistical measures include the root-mean-square-error (RMSE): $\sqrt{(\sum_i^N (\mu_{\text{Method},i} - \mu_{\text{CCSD}(T),i})^2)/N}$, mean error (ME): $\sum_i^N (\mu_{\text{Method},i} - \mu_{\text{CCSD}(T),i})/N$, mean-absolute-error (MAE): $\sum_i^N |\mu_{\text{Method},i} - \mu_{\text{CCSD}(T),i}|/N$, and mean-absolute-percentage-error (MAPE): $\sum_i^N [|\mu_{\text{Method},i} - \mu_{\text{CCSD}(T),i}|/(\mu_{\text{CCSD}(T),i})^*] 100/N$. The molecular sets are described in Section 3.2. (HF) and (OO) denote the use of canonical and pCCD-optimized orbitals, respectively. CO and SiSe are excluded from the analysis.

4.1.1. Analysis for Separate Molecule Sets. A key aspect of this analysis is to answer how the methods behave for the different molecule classes, i.e., for the set of diatomics, small inorganic polyatomics, and larger organic molecules. A quick look at the statistical averages shows the different behaviors of the methods for the two molecular classes. Figure 4 shows that for all the methods, the RMSE values are significantly lower for organic molecules. For the set of “small” molecules, the transition from canonical to pCCD-optimized orbitals worsens the RMSE values of XpCCD, XfpCCSD, and XfpLCCD. For other methods, it remains stagnant; that is, it has almost no or small improvement with pCCD-optimized orbitals in the case of the small molecules. That is also true for the other statistical parameters presented in Table 3. That motivates an in-depth inspection of the performance of the methods for the different molecular classes taken in this work.

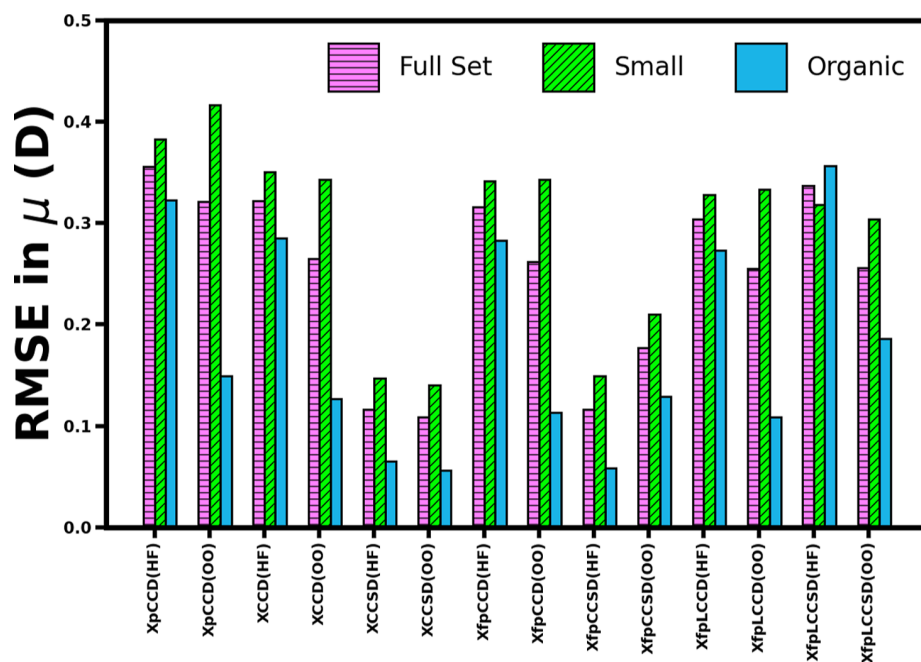
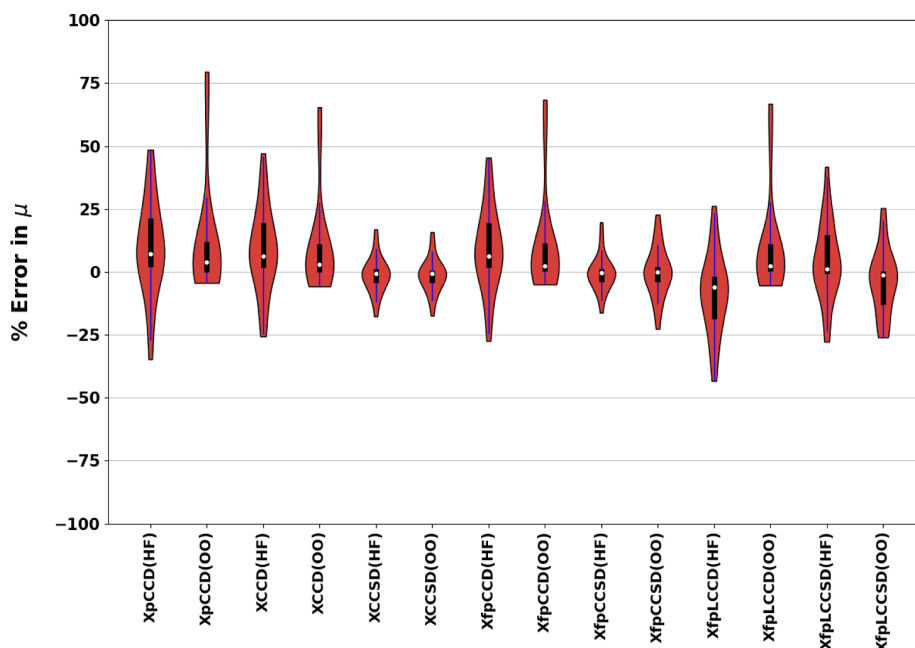


Figure 4. Comparison of CC- and pCCD-based root-mean-square-error (RMSE) of dipole moments (μ in D) w.r.t. response CCSD(T). RMSE is calculated as $\sqrt{(\sum_i^N (\mu_{\text{Method},i} - \mu_{\text{CCSD(T)},i})^2)/N}$. The sets of molecules have been described in Section 3.2. (HF) and (OO) denote the use of canonical and pCCD-optimized molecular orbitals, respectively. CO and SiSe have been excluded from the analysis.

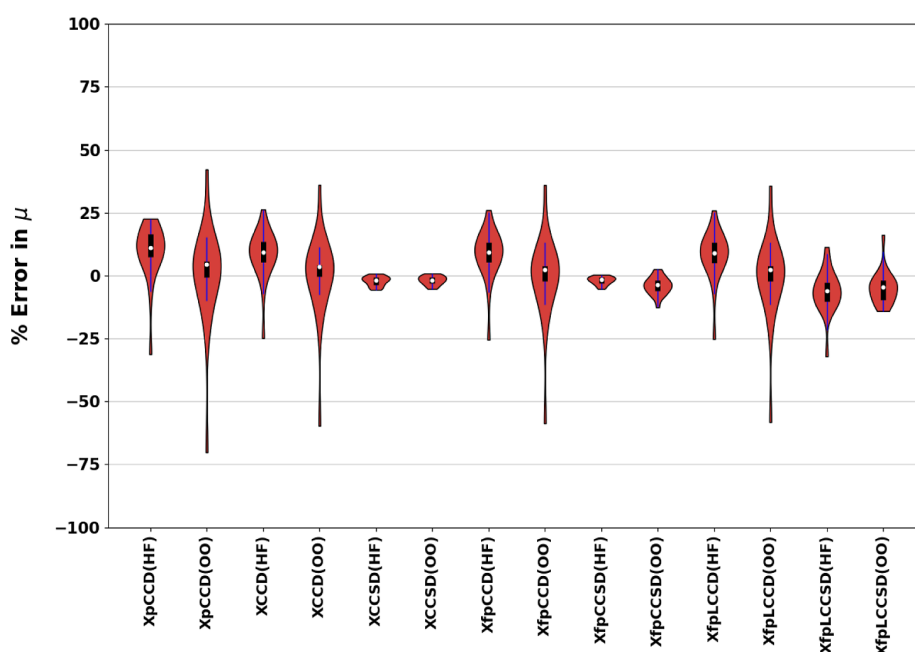
In Figure 5, we show the distribution of the signed percentage errors for (a) small molecules and ions and (b) organic molecules, respectively. For quantities like dipole and quadrupole moments, which can have a very small magnitude for some molecules, the percentage error provides a better assessment tool rather than a simple difference from the reference. In the case of small molecules (Figure 5a), most methods predict the median as positive or very close to zero, except for XfpLCCD, which shows a negative median. The distribution of errors for XCCSD and XfpCCSD using the canonical orbital basis is very similar. It shows that the similar performance between these two methods is not only a coincidence limited to statistical averaging but a molecule-specific feature. However, fpCCSD produces a slightly wider distribution of errors with a pCCD-optimized basis. On the other hand, XpCCD, XCCD, XfpCCD, and XfpLCCD all show narrower distributions (smaller interquartile ranges) once the pCCD-optimized orbitals are compared. The whiskers for all such graphs elongate, indicating a significant increment in errors in the case of some specific molecules. Examples are SiS, GeS, and GeSe molecules, for which we observe errors >50% for all methods using the pCCD-optimized MOs. Even using the canonical MO basis, the errors for these molecules are already in the range of 30–40% and orbital optimization only worsens them. This is also evident when we compare Figure 2a,b. As we can see, for most molecules, the signed errors reduce in length from the canonical to the pCCD-optimized orbital basis. However, the opposite is true for molecules on the right side of the figure, where we have the diatomics of Group 14 and Group 16 elements, analogous to CO. Notably, for fpLCCSD, the errors reduce in magnitude upon using pCCD-optimized orbitals for these three molecules. Most likely, this is a consequence of incorporating some relaxation effects into the pCCD response density matrix.

For organic molecules, on the other hand, all the statistical error parameters for all the methods in Table 3 are lower than the “small” set. pCCD-optimized orbitals further improve the numbers, except for XfpCCSD, for which the performance deteriorates again, and for XCCSD, which remains almost constant. The MAPE for XfpCCSD is remarkably small, $\approx 2\%$, using the canonical Mos, exactly similar to that of XCCSD. Figure 3a,b shows the signed errors for the individual organic molecules using canonical and pCCD-optimized MOs, respectively. Their relative errors are depicted in Figure 5b. We observe that the median and interquartile range of all the methods are much closer to the reference compared to the case of the “small” set. They are shifted closer to the reference when pCCD-optimized orbitals are used, except for fpCCSD. However, in the case of optimized orbitals, the minimum percentage error shown by all the methods containing only double excitations (XpCCD, XCCD, XfpCCD, and XfpLCCD) increases significantly. The particular molecule causing this is cyclopentane (19). Its dipole moment length is very small; CCSD(T) = 0.032 D. Even a tiny difference introduces a significant relative percentage error in such cases. In the case of XfpLCC, going from fpLCCD to fpLCCSD in the case of organic molecules, the RMSE and ME increase in both of the MO bases. This unsatisfactory statistics of XfpLCCSD for organic molecules is largely driven by the high errors shown by it, specifically for nitrobenzene (22) (−1.39 D and −0.46 D in canonical and pCCD-optimized orbital bases). However, this is not much reflected in MAPE as the reference dipole moment of nitrobenzene itself is also high (4.32 D).

4.1.2. Dipole Moment Surface of HF. In the context of prediction of one-electron molecular properties, dipole moment surfaces (DMS) are regular tools to assess a theoretical model’s performance as they reveal the molecule’s rotational–vibrational spectroscopic properties. To that end, in this work,



(a) Percentage errors in dipole moment for the set of small molecules.



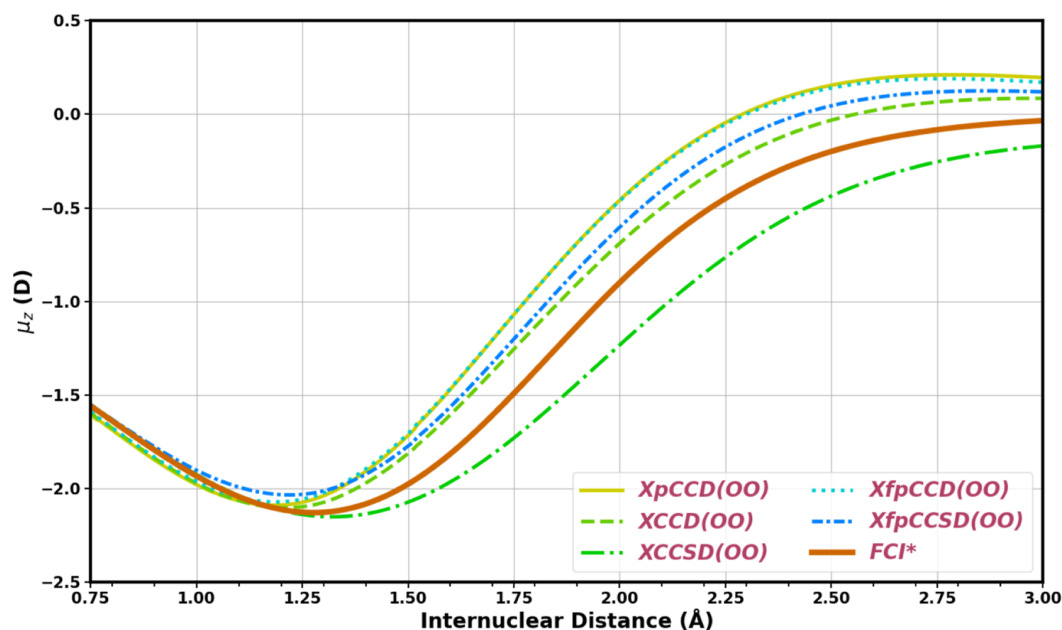
(b) Percentage errors in dipole moment for the set of organic molecules.

Figure 5. Signed percentage errors w.r.t. response CCSD(T) dipole moments $(\mu_{\text{Method}} - \mu_{\text{CCSD(T)}})/(\mu_{\text{CCSD(T)}})$ for all methods for (a) small molecules and (b) organic molecules, using the Sadlej-pVTZ basis set. CO and SiSe have been excluded. (HF) and (OO) denote the use of canonical and pCCD-optimized molecular orbitals, respectively.

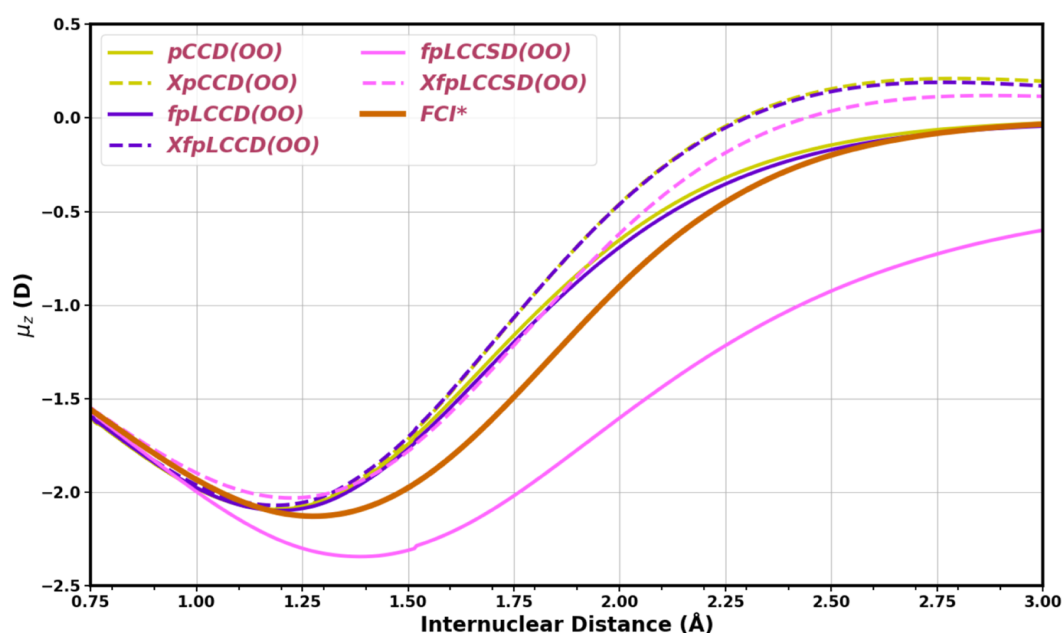
we constructed the DMS of a simple test case of the HF molecule. The molecule is placed along the z -axis with H at the origin and F on the positive z -axis, and then the bond between the two atoms is stretched along the positive z -axis for constructing the DMS. In this analysis, we used pCCD-optimized molecular orbitals. Canonical HF orbital-based pCCD and post-pCCD methods are not size-consistent and also do not produce reliable dipole moment values when the bonds are stretched. Very small strides are taken (0.005 Å),

and optimized MOs from the previous geometry are used as the initial guess for the pCCD orbital optimization cycles at the next step. The corresponding potential energy surfaces are shown in Figure S1 in the Supporting Information.

Though reliable at the equilibrium bond distances, CCSD(T) often fails to estimate dipole moments at nonequilibrium geometries.⁸⁷ Hence, for comparison, we consider here DMS constructed with FCI (subject to SciPy cubic interpolation to get congruent bond distances) at the cc-pVDZ basis set level



(a) Dipole moment surface with expectation-value methods.



(b) Comparison between expectation-value and the corresponding response approaches.

Figure 6. (a) Dipole moment surface of HF in the cc-pVDZ atomic basis set calculated with XpCCD, XCC, and XfpCC methods. (b) Dipole moment surface of HF in the cc-pVDZ atomic basis set calculated with XpCCD, XfpLCC, and their response counterparts. (OO) denotes pCCD-optimized molecular orbitals. In (b), the solid lines represent the response dipole moment surfaces, and the dashed lines show the expectation-value DMS. The FCI* curve is taken from Samanta and Köhn.⁸⁶

with canonical molecular orbitals in a previous work by Samanta and Köhn.⁸⁶ Figure 6a shows the DMS constructed with various expectation-value methods at the same basis set level. In this plot, we specifically focus on the XpCCD, XCC, and XfpCC methods. For better visualization, we do not include fpLCC curves here. Around the equilibrium bond distance (0.917 Å), all of the methods are close to each other. When the H—F bond is stretched, initially μ_z decreases with higher charge density over the F atom, but after a minimum

point, the dipole moment starts increasing. The FCI line turns at around 1.27 Å. Other than XCCSD(OO), all other expectation-value-based curves show a shorter turning point around 1.25 Å. XCCSD(OO) turns at a longer bond length and shows a flatter curve. It also converges to a lower value of μ_z than others at the bond-breaking region. This indicates that the XCCSD overestimates the ionic nature of the stretched H—F bond. The XpCCD(OO) and XfpCCD(OO) curves almost overlap with each other. The XfpCCSD(OO), on the

Table 4. Non-parallelity Errors of the Methods Compared to the FCI Reference in the cc-pVDZ Basis Set in Figure 6

method	non-parallelity error (D)
pCCD(OO)	0.38
XpCCD(OO)	0.50
XCCD(OO)	0.28
XCCSD(OO)	0.35
XfpCCD(OO)	0.48
XfpCCSD(OO)	0.30
fpLCCD(OO)	0.34
XfpLCCD(OO)	0.49
fpLCCSD(OO)	0.74
XfpLCCSD(OO)	0.29

other hand, lies almost parallel to the XCCD(OO) curve. The XfpCCSD(OO) and XCCD(OO) curves also show slightly better asymptotic behavior (converging to 0 D at large interatomic distances). Still, the inclusion of single excitations in fpCC does not seem to improve the dipole moments that much at nonequilibrium geometries. This is also evident, as the XCCD(OO) curve is closer to the FCI line compared to XfpCCSD(OO).

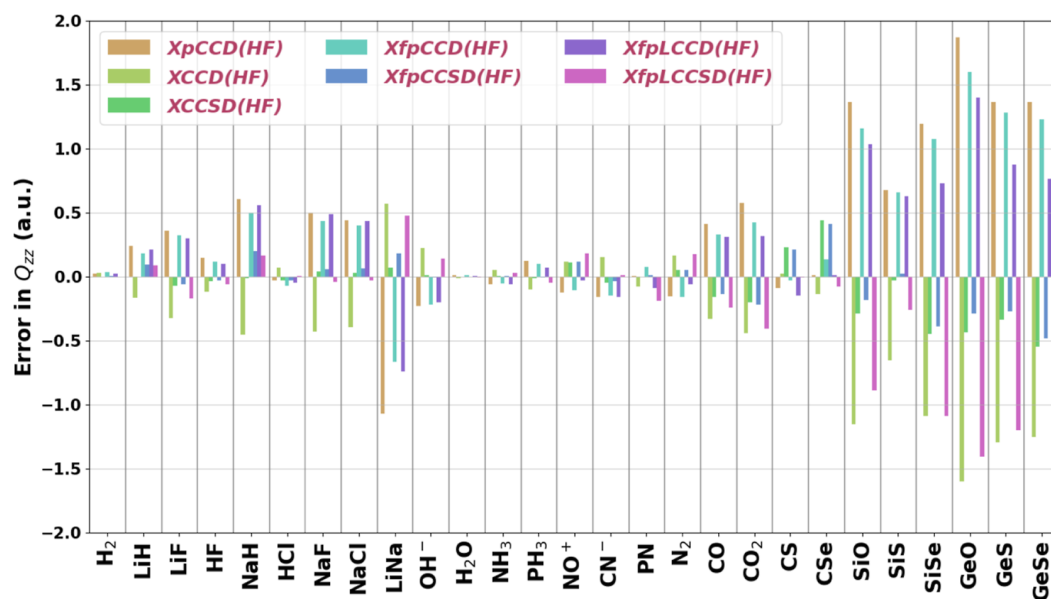
In Figure 6b, we show a comparison of the XpCCD and XfpLCC methods with the corresponding response dipole moments from each method. We note that the response pCCD and fpLCCD curves overlap with each other, similar to their respective expectation-value curves. For these two methods, the response curves show lower differences from the reference FCI line compared to the expectation-value curves. Both the pCCD and fpLCCD response DMS also converge with the reference FCI curve at the bond-breaking region. In contrast, the two fpLCCSD curves behave very differently. Specifically, the response fpLCCSD curve shows a large deviation from the FCI curve after 1 Å and the deviation only increases as we go further. The XfpLCCSD curve, on the other hand, remains closer to the FCI curve compared to the other two expectation-value curves. Table 4 provides the nonparallelity errors (maximum deviation—minimum deviation) of all the methods discussed in Figure 6. To summarize, XfpLCCSD shows the overall best performance across the bond-distance regime considered here. Importantly, it shows a marked improvement compared to the response dipole moments using fpLCCSD. Another important aspect is that XfpCCSD and XfpLCCSD perform better than XCCSD at nonequilibrium geometries.

4.2. Quadrupole Moment. In this section, we analyze the quality of the quadrupole moment values obtained using the implemented methods in this work. In Figures 7 and 8, we show the errors in the calculated quadrupole moment components of individual molecules against the response CCSD(T) values for the “small” set and for organic molecules, respectively. Table 5 details the statistical analysis of the methods for the largest component of the traceless quadrupole moment with respect to the response CCSD(T) values as the reference. In this case, we excluded ClF from the analysis, which exhibited extremely high relative errors for its quadrupole moment with pCCD, XCCD, XfpCCD, and XfpLCCD in the canonical HF orbital basis. Without any single excitation correction, these methods break down completely for ClF, giving even the order of magnitude wrong for its quadrupole moment (refer to the Supporting Information for the values). Figure 9 shows a comparison of the RMSE of different

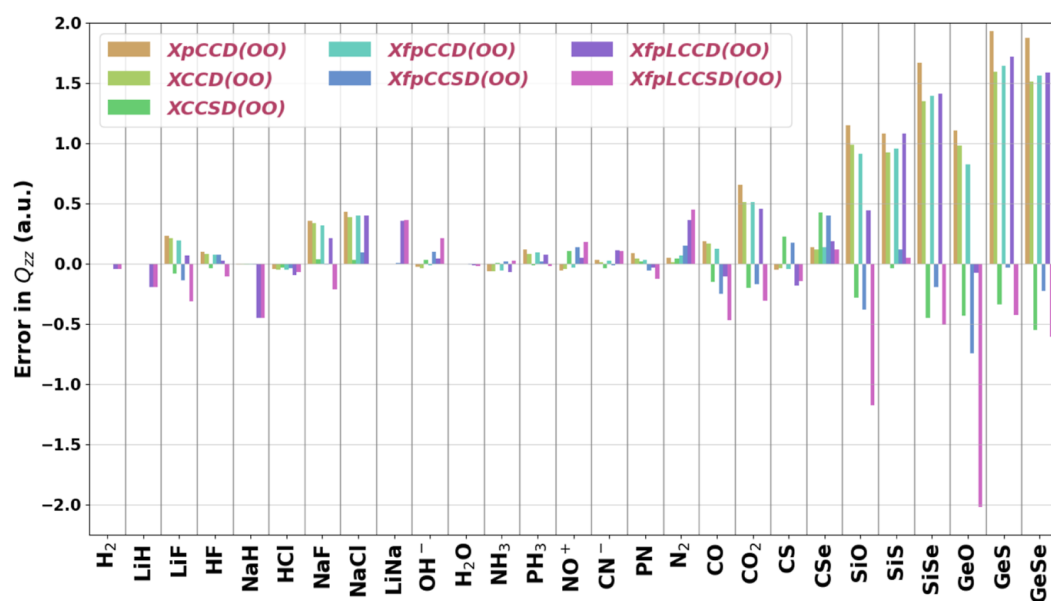
methods in the two MO bases for the three sets of molecules considered in this work. We focus on the performance of the methods for the “full set” here. The statistical parameters in this case follow a similar pattern to that we saw for the dipole moment. Inclusion of single excitations in the cluster operator vastly improves the overall performance of all the methods. The XfpCCSD quadrupole moments using the canonical MOs come out to be the most reliable, showing slightly lower average errors than XCCSD in the same MO basis. Use of pCCD-optimized orbitals improves the statistical parameters of XCCD, XfpCCD, and XfpLCCSD. However, similar to dipole moment, XfpCCSD follows the reverse trend where the optimized MOs downgrade its performance. In the case of quadrupole moment, we see the same pattern for XfpLCCD as well. Another difference to be observed here is the worsening of the results from XfpLCCD to XfpLCCSD. In the canonical MO basis, the MAE and MAPE between these two methods are almost identical. However, XfpLCCSD shows slightly higher RMSE. This indicates that XfpLCCSD has more outliers. Switching to the pCCD-optimized basis further increases the RMSE.

Moving to the “small” set, estimating the quadrupole moment in the canonical MO basis without the singles excitation seems particularly problematic, as shown by the very high RMSE of XpCCD, XCCD, XfpCCD, and XfpLCCD for this set. In terms of the magnitude of error, again, the diatomics of groups 14 and 16 turned out to be especially challenging, as we can see from Figure 7. For example, for GeO in the canonical MO basis, the errors are 1.87, −1.60, 1.60, and 1.40 a.u. for XpCCD, XCCD, XfpCCD, and XfpLCCD, respectively. The singles correction in fpLCCSD is also inadequate, as errors for XfpLCCSD continue to be high for these molecules. In Figure 10a, we show the distribution of the relative errors in Q_{zz} for the “small” set. XpCCD, XCCD, and XfpCCD have very similar distributions on the pCCD-optimized MO basis. For these methods, going from canonical to optimized MO basis narrows the relative error distribution and shifts the medians toward the reference. However, the maximum relative errors increase for all three, and the violin plots get elongated. As discussed previously, the molecules responsible for this are SiS, SiSe, GeS, and GeSe. Among the other methods, the distribution of relative error for XfpCCSD in the canonical MO basis is remarkably similar to that of XCCSD. Though pCCD-optimized MOs make the relative error distribution slightly wider for XfpCCSD, the highest error remains under 20% (for CSe). XfpLCCSD behaves differently compared to other methods because its relative error distribution improves to a small extent with pCCD-optimized MOs. The median becomes more negative in this case, although the largest errors are reduced. For example, in the canonical basis for XfpLCCSD, the highest errors are shown by SiSe (58%) and GeSe (%). In the case of pCCD-optimized MOs, the highest errors are given by CO (35%) and SiSe (33%).

For the case of organic molecules, as shown in Figure 10b, the distributions of relative errors for all of the methods are remarkably small. The effect of optimized orbital follows the same trend as in other cases. XCCSD and XfpCCSD show MAPE as low as 1–2%. The maximum relative error is found for formaldehyde with the doubles-only methods in the canonical MO basis, though these are still under 20%.



(a) Error comparison across methods for each molecule in the canonical MO basis.



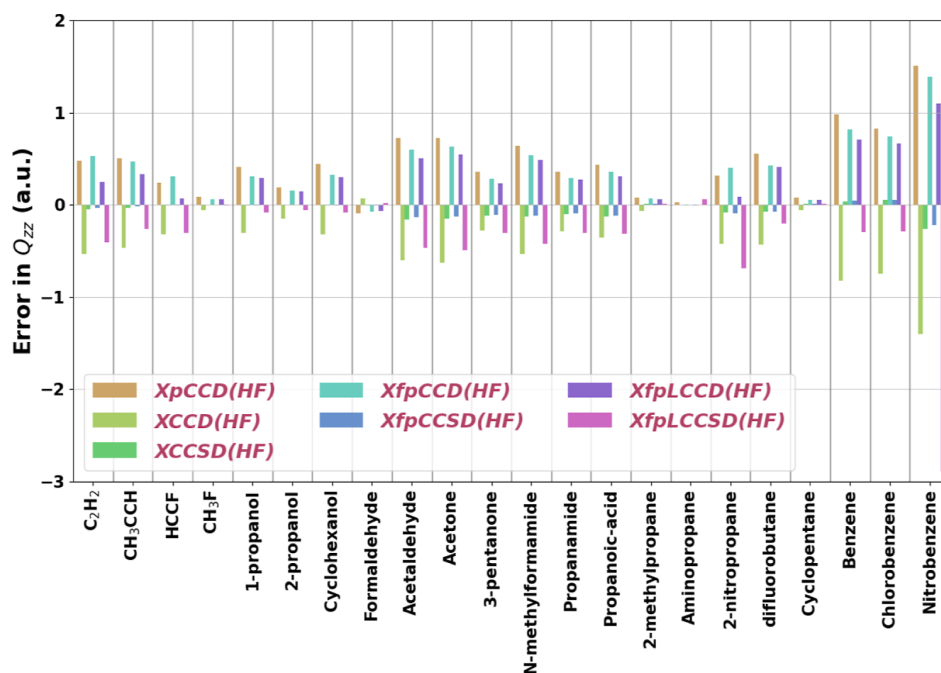
(b) Error comparison across methods for each molecule in the pCCD-optimized MO basis.

Figure 7. Signed errors with respect to CCSD(T) in the quadrupole moment component ($Q_{zz}^{\text{Method}} - Q_{zz}^{\text{CCSD(T)}}$) for all methods using the Sadlej-pVTZ basis set for the set of small molecules investigated in this work. (HF) in (a) and (OO) in (b) denote the use of canonical and pCCD-optimized molecular orbitals, respectively.

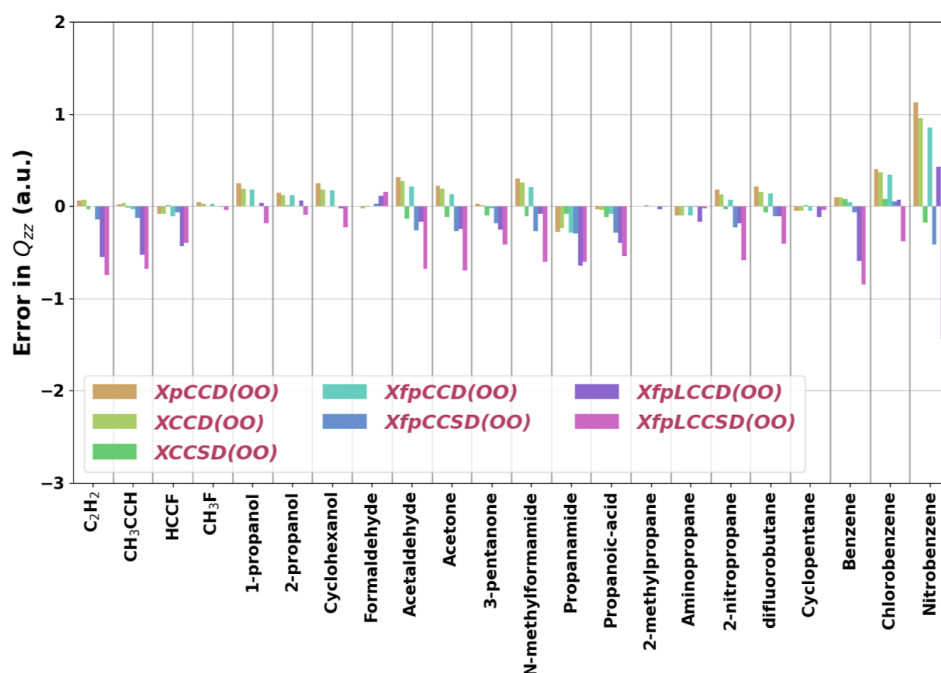
5. CONCLUSIONS

In this work, we implemented and tested the expectation-value-based single-electron property calculation approach for pCCD-based methods. Our approach offers an alternate and computationally cheaper route for computing 1-RDM-based molecular properties to more computationally expensive Λ -equations within the CC response framework. Yet, we implemented the quadrupole moment in the PyBEST software package, which allowed us to compute and benchmark pCCD-based quadrupole moments for the first time.

Our analysis of expectation value pCCD-based dipole and quadrupole moments shows that moving from the canonical HF orbitals to the pCCD-optimized orbital basis reduces errors in all cases except XfpCCSD. That, in turn, highlights that the localized pCCD orbitals are a good choice for computing properties. Moreover, our data suggest that the errors do not emerge from choosing pCCD-optimized localized orbitals but from the method itself—more precisely, from the consecutive treatment of zero and nonzero seniority



(a) Error comparison across methods for each molecule in the canonical MO basis.



(b) Error comparison across methods for each molecule in the pCCD-optimized MO basis.

Figure 8. Signed errors with respect to CCSD(T) in the quadrupole moment component ($Q_{zz}^{\text{Method}} - Q_{zz}^{\text{CCSD(T)}}$) for all methods using the Sadlej-pVTZ basis set for the set of organic molecules investigated in this work. (HF) in (a) and (OO) in (b) denote the use of canonical and pCCD-optimized molecular orbitals, respectively.

sectors; splitting the seniority sectors does not improve property values.

Furthermore, our numerical analysis of organic systems shows that XpCCD(OO) already provides reliable dipole and quadrupole moments; the corresponding errors are similar to those of XCCD, XfpCC, and XfpLCC. Specifically, XfpCCD and XfpLCCD are the best (X)pCCD-based methods. On the

other hand, including single excitations in the fpCC or fpLCC models worsens the performance of pCCD-based models. In contrast to the organic-molecule test set, the performance is reversed for the small-molecule test set. Thus, pCCD-based methods are promising for one-electron properties in organic systems. pCCD orbital optimization is also beneficial for reliably modeling the dipole moment surface of the bond-

Table 5. Statistical Analysis of Various CC- and pCCD-Based Methods with Respect to the Quadrupole Moment (Q_{zz} in a.u.) Using the Sadlej-pVTZ Basis Set^a

method	RMSE (a.u.)		ME (a.u.)		MAE (a.u.)		MAPE	
	full set	small	organic	full set	small	organic	full set	small
XpCCD(HF)	0.656	0.716	0.573	0.394	0.349	0.450	0.476	0.491
XpCCD(OO)	0.584	0.737	0.301	0.288	0.407	0.142	0.320	0.424
XCCD(HF)	0.571	0.616	0.510	-0.354	-0.319	-0.400	0.414	0.423
XCCD(OO)	0.486	0.613	0.257	0.238	0.338	0.116	0.269	0.355
XCCSD(HF)	0.170	0.213	0.093	-0.057	-0.061	-0.054	0.103	0.135
XCCSD(OO)	0.165	0.211	0.075	-0.051	-0.064	-0.036	0.097	0.132
XpCCD(HF)	0.571	0.618	0.508	0.353	0.320	0.395	0.416	0.427
XpCCD(OO)	0.481	0.613	0.232	0.223	0.336	0.084	0.259	0.351
XpCCSD(HF)	0.151	0.190	0.082	-0.035	-0.025	-0.047	0.099	0.133
XpCCSD(OO)	0.196	0.209	0.179	-0.073	-0.036	-0.120	0.130	0.131
XpLCCD(HF)	0.470	0.512	0.413	0.277	0.250	0.311	0.343	0.363
XpLCCD(OO)	0.496	0.606	0.312	0.072	0.273	-0.174	0.309	0.366
XpLCCSD(HF)	0.611	0.544	0.684	-0.285	-0.230	-0.353	0.341	0.325
XpLCCSD(OO)	0.541	0.525	0.559	-0.309	-0.211	-0.431	0.377	0.324

^aThe statistical analysis includes the root-mean-square-error (RMSE): $\sqrt{(\sum_i^N (Q_{zz}^{\text{Method},i} - Q_{zz}^{\text{CCSD(T),i}})^2)/N}$, mean-error (ME): $\sum_i^N (Q_{zz}^{\text{Method},i} - Q_{zz}^{\text{CCSD(T),i}})/N$, mean-absolute-error (MAE): $\sum_i^N |Q_{zz}^{\text{Method},i} - Q_{zz}^{\text{CCSD(T),i}}|/N$, and mean-absolute-percentage-error (MAPE): $\sum_i^N |Q_{zz}^{\text{Method},i} - Q_{zz}^{\text{CCSD(T),i}}|/(Q_{zz}^{\text{CCSD(T),i}}) \times 100/N$. The molecular sets are described in Section 3.2. (HF) and (OO) denote the use of canonical and pCCD-optimized orbitals, respectively. CIF has been excluded from the analysis.

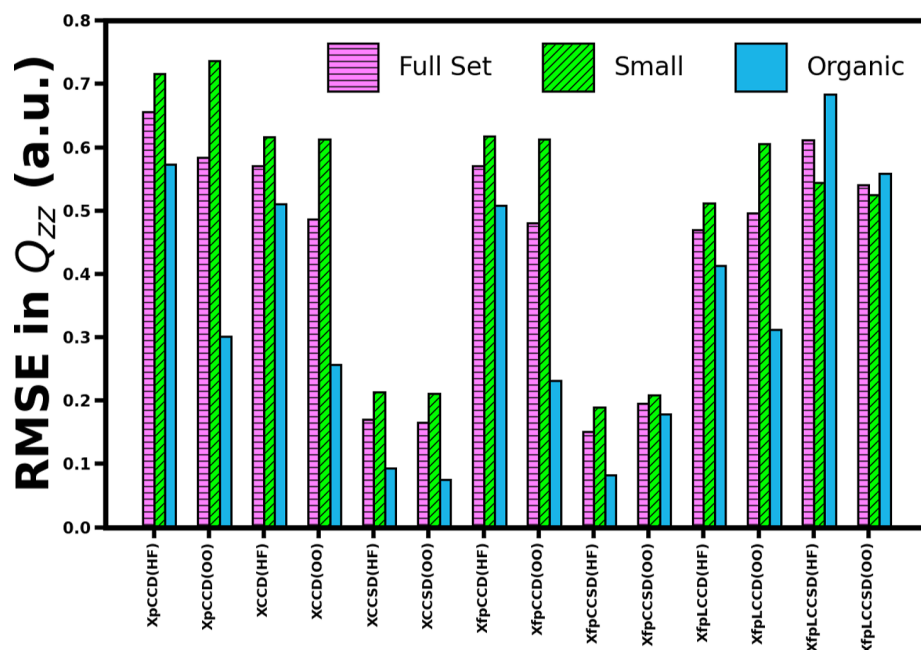


Figure 9. Comparison of root-mean-square-error (RMSE) of quadrupole moment (Q_{zz} in a.u.) values obtained with various methods. The RMSE is calculated as $\sqrt{(\sum_i^N (Q_{zz}^{\text{Method},i} - Q_{zz}^{\text{CCSD(T),i}})^2)/N}$. The sets of molecules have been described in Section 3.2. (HF) and (OO) denote the use of canonical and pCCD-optimized orbitals, respectively. CIF has been excluded from the analysis.

breaking process of the HF molecule. For this bond-breaking example, however, the response DMS resulted in smaller nonparallelity errors than the corresponding expectation value DMS.

To conclude, we demonstrated that pCCD is a computationally cheap method that provides a good starting point to model one-electron properties in organic systems and systems with a strong electron correlation. The good performance of pCCD-based methods in approximating the 1-RDMs within the expectation value approach makes them a computationally attractive alternative to solving the corresponding CC Λ -equations. This feature is particularly important when calculating orbital entanglement and correlation measures^{88,89} (single orbital entropies and orbital-pair mutual information) from the tailored (or frozen-pair) CC methods.⁹⁰

■ ASSOCIATED CONTENT

Data Availability Statement

The data underlying this study are available in the published article and its Supporting Information. The PyBEST code is available on Zenodo at <https://zenodo.org/records/10069179> and on PyPI at <https://pypi.org/project/pybest/>.

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpca.5c03859>.

Bond distances (in Å) of all the “small” molecules and ions and xyz coordinates (in Å) of all organic molecules considered in this work (PDF)

All the dipole and quadrupole moment values obtained with different expectation-value methods, and response CCSD(T) and CCSD using the Sadlej pVTZ basis set (XLSX)

Potential energy surface of HF using the cc-pVDZ basis set with all CC variants; deviation of the XpCCD; response pCCD, XCC, XfpCC, and XfpLCC and

response fpLCC dipole moment surfaces from the reference FCI surface in the cc-pVDZ basis set; dipole moment surface of HF with the XpCCD, XCC, and XfpCC methods using the Sadlej pVTZ basis set; statistical measurements with respect to CCSD response dipole and quadrupole moment values; dipole moment values at all bond distances calculated during the construction of the dipole moment surface of the HF molecule with all the expectation-value methods discussed here in the cc-pVDZ basis set (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Somayeh Ahmadvkhani – Institute of Physics, Faculty of Physics, Astronomy, and Informatics, Nicolaus Copernicus University in Toruń, Toruń 87-100, Poland; orcid.org/0000-0001-6312-632X; Email: so.ahmadvkhani@gmail.com

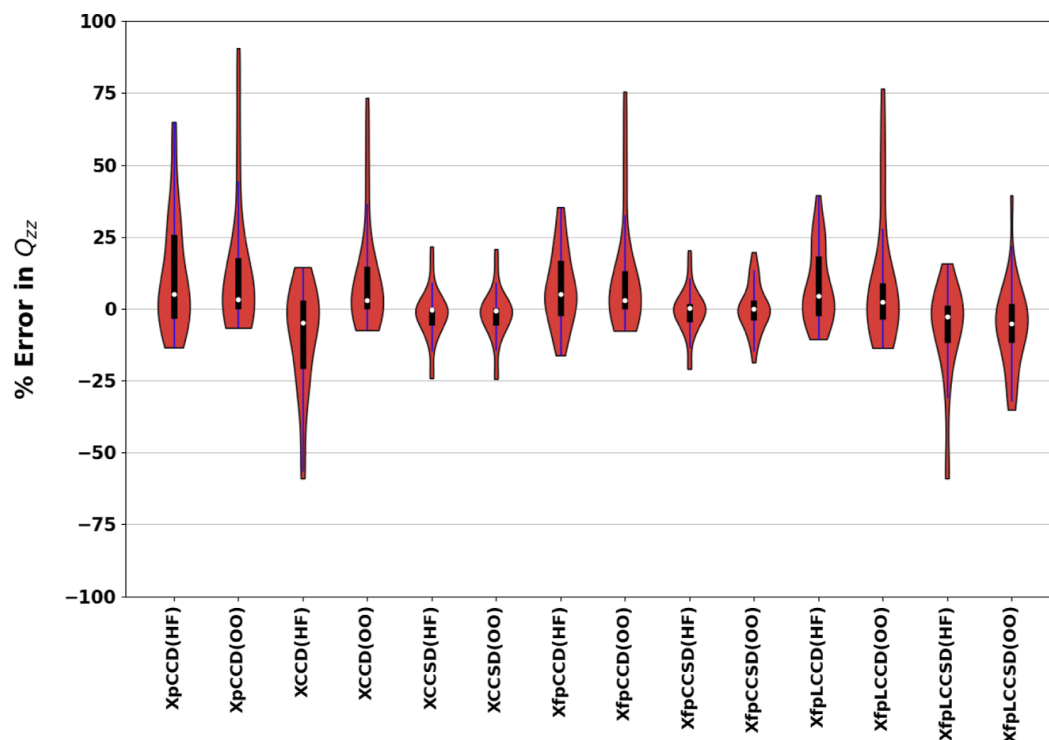
Paweł Tecmer – Institute of Physics, Faculty of Physics, Astronomy, and Informatics, Nicolaus Copernicus University in Toruń, Toruń 87-100, Poland; orcid.org/0000-0001-6347-878X; Email: ptecmer@fizyka.umk.pl

Authors

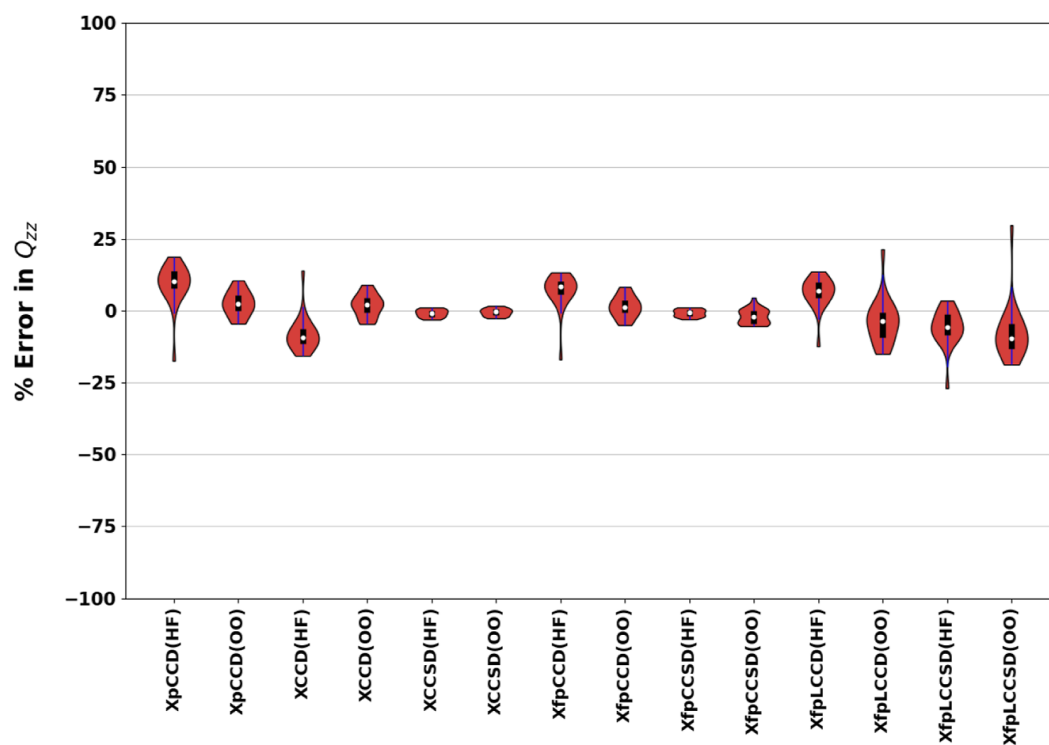
Rahul Chakraborty – Institute of Physics, Faculty of Physics, Astronomy, and Informatics, Nicolaus Copernicus University in Toruń, Toruń 87-100, Poland; orcid.org/0000-0003-4086-6573

Julian Świerczyński – Institute of Physics, Faculty of Physics, Astronomy, and Informatics, Nicolaus Copernicus University in Toruń, Toruń 87-100, Poland

Katharina Boguslawski – Institute of Physics, Faculty of Physics, Astronomy, and Informatics, Nicolaus Copernicus University in Toruń, Toruń 87-100, Poland; orcid.org/0000-0001-7793-1151



(a) Percentage errors in quadrupole moment for the set of small molecules.



(b) Percentage errors in quadrupole moment for the set of organic molecules.

Figure 10. Signed percentage errors with respect to CCSD(T) reference values in the quadrupole moment component ($(Q_{zz}^{\text{Method}} - Q_{zz}^{\text{CCSD(T)}})/Q_{zz}^{\text{CCSD(T)}}$) for all methods using the Sadlej-pVTZ basis set for the set of (a) small and (b) organic molecules investigated in this work. ClF has been excluded. (HF) and (OO) denote the use of canonical and pCCD-optimized molecular orbitals, respectively. The same range for the y-axis has been used in both plots for comparison.

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acs.jpca.5c03859>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

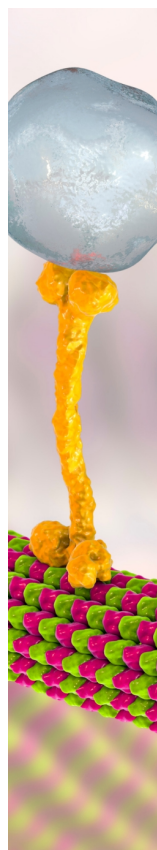
R.C. and P.T. acknowledge financial support from the OPUS research grant from the National Science Centre, Poland (Grant No. 2019/33/B/ST4/02114). S.A. acknowledges financial support from the SONATA BIS research grant from the National Science Centre, Poland (Grant No. 2021/42/E/ST4/00302), funded/co-funded by the European Union (ERC, DRESSED-pCCD, 101077420). Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

REFERENCES

- (1) Čížek, J. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods. *J. Chem. Phys.* **1966**, *45*, 4256–4266.
- (2) Čížek, J. On the use of the cluster expansion and the technique of diagrams in calculations of correlation effects in atoms and molecules. *Adv. Chem. Phys.* **1969**, *14*, 35–89.
- (3) Čížek, J.; Paldus, J. Correlation problems in atomic and molecular systems III. Rederivation of the coupled-pair many-electron theory using the traditional quantum chemical methods. *Int. J. Quantum Chem.* **1971**, *5*, 359–379.
- (4) Bartlett, R. J. Many-Body Perturbation Theory and Coupled Cluster Theory for Electron Correlation in Molecules. *Annu. Rev. Phys. Chem.* **1981**, *32*, 359–401.
- (5) Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291–350.
- (6) Lyakh, D. I.; Musiał, M.; Lotrich, V. F.; Bartlett, J. Multireference Nature of Chemistry: The Coupled-Cluster View. *Chem. Rev.* **2012**, *112*, 182–243.
- (7) Sneskov, K.; Christiansen, O. Excited state coupled cluster methods. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 566–584.
- (8) Zhang, I. Y.; Grüneis, A. Coupled cluster theory in materials science. *Front. Mater.* **2019**, *6*, 123.
- (9) Kowalski, K.; Olson, R. M.; Krishnamoorthy, S.; Tipparaju, V.; Apra, E. Role of Many-Body Effects in Describing Low-Lying Excited States of π -Conjugated Chromophores: High-Level Equation-of-Motion Coupled-Cluster Studies of Fused Porphyrin Systems. *J. Chem. Theory Comput.* **2011**, *7*, 2200–2208.
- (10) Gyevi-Nagy, L.; Kállay, M.; Nagy, P. R. Integral-direct and parallel implementation of the CCSD(T) method: Algorithmic developments and large-scale applications. *J. Chem. Theory Comput.* **2020**, *16*, 366–384.
- (11) Nagy, P. R.; Kállay, M. Approaching the basis set limit of CCSD(T) energies for large molecules with local natural orbital coupled-cluster methods. *J. Chem. Theory Comput.* **2019**, *15*, 5275–5298.
- (12) Szabó, P. B.; Csóka, J.; Kállay, M.; Nagy, P. R. Linear-Scaling Local Natural Orbital CCSD(T) Approach for Open-Shell Systems: Algorithms, Benchmarks, and Large-Scale Applications. *J. Chem. Theory Comput.* **2023**, *19*, 8166–8188.
- (13) Nagy, P. R. State-of-the-art local correlation methods enable affordable gold standard quantum chemistry for up to hundreds of atoms. *Chem. Sci.* **2024**, *15*, 14556–14584.
- (14) Pinski, P.; Riplinger, C.; Valeev, E. F.; Neese, F. Sparse maps—A systematic infrastructure for reduced-scaling electronic structure methods. I. An efficient and simple linear scaling local MP2 method that uses an intermediate basis of pair natural orbitals. *J. Chem. Phys.* **2015**, *143*, 034108.
- (15) Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. Sparse maps—A systematic infrastructure for reduced-scaling electronic structure methods. II. Linear scaling domain based pair natural orbital coupled cluster theory. *J. Chem. Phys.* **2016**, *144*, 024109.
- (16) Guo, Y.; Riplinger, C.; Becker, U.; Liakos, D. G.; Minenkov, Y.; Cavallo, L.; Neese, F. Communication: An improved linear scaling perturbative triples correction for the domain based local pair-natural orbital based singles and doubles coupled cluster method [DLPNO-CCSD(T)]. *J. Chem. Phys.* **2018**, *148*, 011101.
- (17) Datta, D.; Kossmann, S.; Neese, F. Analytic energy derivatives for the calculation of the first-order molecular properties using the domain-based local pair-natural orbital coupled-cluster theory. *J. Chem. Phys.* **2016**, *145*, 114101.
- (18) Bulik, I. W.; Henderson, T. M.; Scuseria, G. E. Can single-reference coupled cluster theory describe static correlation? *J. Chem. Theory Comput.* **2015**, *11*, 3171–3179.
- (19) Tecmer, P.; Boguslawski, K. Geminal-based electronic structure methods in quantum chemistry. Toward geminal model chemistry. *Phys. Chem. Chem. Phys.* **2022**, *24*, 23026–23048.
- (20) Köhn, A.; Hanauer, M.; Mück, L. A.; Jagau, T.-C.; Gauss, J. State-specific multireference coupled-cluster theory. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2013**, *3*, 176–197.
- (21) Waigum, A.; Suchanek, S.; Köhn, A. Simplified Multireference Coupled-Cluster Methods: Hybrid Approaches With Averaged Coupled Pair Theories. *J. Comput. Chem.* **2025**, *46*, e70020.
- (22) Piecuch, P.; Čížek, J.; Paldus, J. Behavior of coupled cluster energy in the strongly correlated limit of the cyclic polyene model. Comparison with the exact results. *Int. J. Quantum Chem.* **1992**, *42*, 165–191.
- (23) Kats, D. C. Communication: The distinguishable cluster approximation. II. The role of orbital relaxation. *J. Chem. Phys.* **2014**, *141*, 061101.
- (24) Limacher, P. A.; Ayers, P. W.; Johnson, P. A.; De Baerdemacker, S.; Van Neck, D.; Bultinck, P. A new mean-field method suitable for strongly correlated electrons: computationally facile antisymmetric products of nonorthogonal geminals. *J. Chem. Theory Comput.* **2013**, *9*, 1394–1401.
- (25) Boguslawski, K.; Tecmer, P.; Ayers, P. W.; Bultinck, P.; De Baerdemacker, S.; Van Neck, D. Efficient description of strongly correlated electrons. *Phys. Rev. B* **2014**, *89*, 201106.
- (26) Stein, T.; Henderson, T. M.; Scuseria, G. E. Seniority zero pair coupled cluster doubles theory. *J. Chem. Phys.* **2014**, *140*, 214113.
- (27) Henderson, T. M.; Bulik, I. W.; Stein, T.; Scuseria, G. E. Seniority-based coupled cluster theory. *J. Chem. Phys.* **2014**, *141*, 244104.
- (28) Bytautas, L.; Henderson, T. M.; Jiménez-Hoyos, C. A.; Ellis, J. K.; Scuseria, G. E. Seniority and orbital symmetry as tools for establishing a full configuration interaction hierarchy. *J. Chem. Phys.* **2011**, *135*, 044119.
- (29) Bytautas, L.; Scuseria, G. E.; Ruedenberg, K. Seniority number description of potential energy surfaces: Symmetric dissociation of water, N_2 , C_2 , and Be_2 . *J. Chem. Phys.* **2015**, *143*, 094105.
- (30) Boguslawski, K.; Tecmer, P.; Limacher, P. A.; Johnson, P. A.; Ayers, P. W.; Bultinck, P.; De Baerdemacker, S.; Van Neck, D. Projected seniority-two orbital optimization of the antisymmetric product of one-reference orbital geminal. *J. Chem. Phys.* **2014**, *140*, 214114.
- (31) Limacher, P. A.; Kim, T. D.; Ayers, P. W.; Johnson, P. A.; De Baerdemacker, S.; Van Neck, D.; Bultinck, P. The influence of orbital rotation on the energy of closed-shell wavefunctions. *Mol. Phys.* **2014**, *112*, 853–862.
- (32) Boguslawski, K.; Tecmer, P.; Ayers, P. W.; Bultinck, P.; De Baerdemacker, S.; Van Neck, D. Non-variational orbital optimization techniques for the AP1roG wave function. *J. Chem. Theory Comput.* **2014**, *10*, 4873–4882.

- (33) Rodriguez-Mayorga, M.; Loos, P.-F.; Bruneval, F.; Visscher, L. Time-reversal symmetry in RDMFT and pCCD with complex-valued orbitals. *J. Chem. Phys.* **2025**, *162*, 054716.
- (34) Leszczyk, A.; Máté, M.; Legeza, Ö.; Boguslawski, K. Assessing the accuracy of tailored coupled cluster methods corrected by electronic wave functions of polynomial cost. *J. Chem. Theory Comput.* **2022**, *18*, 96–117.
- (35) Boguslawski, K.; Ayers, P. W. Linearized Coupled Cluster Correction on the Antisymmetric Product of 1-Reference Orbital Geminals. *J. Chem. Theory Comput.* **2015**, *11*, 5252–5261.
- (36) Tecmer, P.; Boguslawski, K.; Limacher, P. A.; Johnson, P. A.; Chan, M.; Verstraelen, T.; Ayers, P. W. Assessing the accuracy of new geminal-based approaches. *J. Phys. Chem. A* **2014**, *118*, 9058–9068.
- (37) Tecmer, P.; Boguslawski, K.; Ayers, P. W. Singlet ground state actinide chemistry with geminals. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14427–14436.
- (38) Garza, A. J.; Bulik, I. W.; Henderson, T. M.; Scuseria, G. E. Range separated hybrids of pair coupled cluster doubles and density functionals. *Phys. Chem. Chem. Phys.* **2015**, *17*, 22412–22422.
- (39) Boguslawski, K.; Tecmer, P.; Legeza, Ö. Analysis of two-orbital correlations in wavefunctions restricted to electron-pair states. *Phys. Rev. B* **2016**, *94*, 155126.
- (40) Boguslawski, K.; Tecmer, P. Benchmark of dynamic electron correlation models for seniority-zero wavefunctions and their application to thermochemistry. *J. Chem. Theory Comput.* **2017**, *13*, 5966–5983.
- (41) Brzęk, F.; Boguslawski, K.; Tecmer, P.; Żuchowski, P. S. Benchmarking the accuracy of seniority-zero wave function methods for noncovalent interactions. *J. Chem. Theory Comput.* **2019**, *15*, 4021–4035.
- (42) Tecmer, P.; Boguslawski, K.; Borkowski, M.; Żuchowski, P. S.; Kędziera, D. Modeling the electronic structures of the ground and excited states of the ytterbium atom and the ytterbium dimer: A modern quantum chemistry perspective. *Int. J. Quantum Chem.* **2019**, *119*, e25983.
- (43) Nowak, A.; Tecmer, P.; Boguslawski, K. Assessing the accuracy of simplified coupled cluster methods for electronic excited states in f0 actinide compounds. *Phys. Chem. Chem. Phys.* **2019**, *21*, 19039–19053.
- (44) Kossoski, F.; Marie, A.; Scemama, A.; Caffarel, M.; Loos, P.-F. Excited states from state-specific orbital-optimized pair coupled cluster. *J. Chem. Theory Comput.* **2021**, *17*, 4756–4768.
- (45) Marie, A.; Kossoski, F.; Loos, P.-F. Variational coupled cluster for ground and excited states. *J. Chem. Phys.* **2021**, *155*, 104105.
- (46) Tecmer, P.; Gałyńska, M.; Szczuczko, L.; Boguslawski, K. Geminal-based strategies for modeling large building blocks of organic electronic materials. *J. Phys. Chem. Lett.* **2023**, *14*, 9909–9917.
- (47) Ahmadkhani, S.; Boguslawski, K.; Tecmer, P. Linear Response pCCD-Based Methods: LR-pCCD and LR-pCCD+S Approaches for the Efficient and Reliable Modeling of Excited State Properties. *J. Chem. Theory Comput.* **2024**, *20*, 10443–10452.
- (48) Bintrim, S. J.; Carter-Fenk, K. Optimal-Reference Excited State Methods: Static Correlation at Polynomial Cost with Single-Reference Coupled-Cluster Approaches. *J. Chem. Theory Comput.* **2025**, *21*, 4080–4094.
- (49) Salter, E.; Sekino, H.; Bartlett, R. J. Property evaluation and orbital relaxation in coupled cluster methods. *J. Chem. Phys.* **1987**, *87*, 502–509.
- (50) Helgaker, T.; Jørgensen, P. Configuration-interaction energy derivatives in a fully variational formulation. *Theor. Chem. Acc.* **1989**, *75*, 111–127.
- (51) Salter, E.; Trucks, G. W.; Bartlett, R. J. Analytic energy derivatives in many-body methods. I. First derivatives. *J. Chem. Phys.* **1989**, *90*, 1752–1766.
- (52) Koch, H.; Jensen, H. J. A.; Jørgensen, P.; Helgaker, T.; Scuseria, G. E.; Schaefer, I. I. H. F. Coupled cluster energy derivatives. Analytic Hessian for the closed-shell coupled cluster singles and doubles wave function: Theory and applications. *J. Chem. Phys.* **1990**, *92*, 4924–4940.
- (53) Chakraborty, R.; de Moraes, M. M. F.; Boguslawski, K.; Nowak, A.; Świerczyński, J.; Tecmer, P. Toward Reliable Dipole Moments without Single Excitations: The Role of Orbital Rotations and Dynamical Correlation. *J. Chem. Theory Comput.* **2024**, *20*, 4689–4702.
- (54) Jeziorski, B.; Moszyński, R. Explicitly connected expansion for the average value of an observable in the coupled-cluster theory. *Int. J. Quantum Chem.* **1993**, *48*, 161–183.
- (55) Moszyński, R.; Żuchowski, P. S.; Jeziorski, B. Time-independent coupled-cluster theory of the polarization propagator. *Collect. Czechoslov. Chem. Commun.* **2005**, *70*, 1109–1132.
- (56) Korona, T.; Przybytek, M.; Jeziorski, B. Time-independent coupled cluster theory of the polarization propagator. Implementation and application of the singles and doubles model to dynamic polarizabilities and van der Waals constants. *Mol. Phys.* **2006**, *104*, 2303–2316.
- (57) Korona, T.; Jeziorski, B. Dispersion energy from density-fitted density susceptibilities of singles and doubles coupled cluster theory. *J. Chem. Phys.* **2008**, *128*, 144107.
- (58) Korona, T.; Jeziorski, B. One-electron properties and electrostatic interaction energies from the expectation value expression and wave function of singles and doubles coupled cluster theory. *J. Chem. Phys.* **2006**, *125*, 184109.
- (59) Korona, T. The effect of local approximations on first-order properties from expectation-value coupled cluster theory. *Theor. Chem. Acc.* **2011**, *129*, 15–30.
- (60) Boguslawski, K.; Leszczyk, A.; Nowak, A.; Brzęk, F.; Żuchowski, P. S.; Kędziera, D.; Tecmer, P. Pythonic Black-box Electronic Structure Tool (PyBEST). An open-source Python platform for electronic structure calculations at the interface between chemistry and physics. *Comput. Phys. Commun.* **2021**, *264*, 107933.
- (61) Boguslawski, K.; Brzęk, F.; Chakraborty, R.; Cieślak, K.; Jahani, S.; Leszczyk, A.; Nowak, A.; Sujkowski, E.; Świerczyński, J.; Ahmadkhani, S.; et al. PyBEST: improved functionality and enhanced performance. *Comput. Phys. Commun.* **2024**, *297*, 109049.
- (62) Ellis, J. K.; Martin, R. L.; Scuseria, G. E. On Pair Functions for Strong Correlations. *J. Chem. Theory Comput.* **2013**, *9*, 2857–2869.
- (63) Boguslawski, K.; Tecmer, P.; Ayers, P. W.; Bultinck, P.; De Baerdemacker, S.; Van Neck, D. Efficient description of strongly correlated electrons with mean-field cost. *Phys. Rev. B* **2014**, *89*, 201106.
- (64) Oliphant, N.; Adamowicz, L. Multireference coupled-cluster method using a single-reference formalism. *J. Chem. Phys.* **1991**, *94*, 1229–1235.
- (65) Oliphant, N.; Adamowicz, L. The implementation of the multireference coupled-cluster method based on the single-reference formalism. *J. Chem. Phys.* **1992**, *96*, 3739–3744.
- (66) Piecuch, P.; Oliphant, N.; Adamowicz, L. A state-selective multireference coupled-cluster theory employing the single-reference formalism. *J. Chem. Phys.* **1993**, *99*, 1875–1900.
- (67) Piecuch, P.; Adamowicz, L. State-selective multireference coupled-cluster theory employing the single-reference formalism: Implementation and application to the H₈ model system. *J. Chem. Phys.* **1994**, *100*, 5792–5809.
- (68) Adamowicz, L.; Piecuch, P.; Ghose, K. B. The state-selective coupled cluster method for quasi-degenerate electronic states. *Mol. Phys.* **1998**, *94*, 225–234.
- (69) Piecuch, P. Active-space coupled-cluster methods. *Mol. Phys.* **2010**, *108*, 2987–3015.
- (70) Veis, L.; Antalík, A.; Brabec, J.; Neese, F.; Legeza, Ö.; Pittner, J. Coupled cluster method with single and double excitations tailored by matrix product state wave functions. *J. Phys. Chem. Lett.* **2016**, *7*, 4072–4078.
- (71) Faulstich, F. M.; Máté, M.; Laestadius, A.; Csirik, M. A.; Veis, L.; Antalík, A.; Brabec, J.; Schneider, R.; Pittner, J.; Kvaal, S.; et al. Numerical and theoretical aspects of the DMRG-TCC method exemplified by the nitrogen dimer. *J. Chem. Theory Comput.* **2019**, *15*, 2206–2220.

- (72) Mörchen, M.; Freitag, L.; Reiher, M. Tailored coupled cluster theory in varying correlation regimes. *J. Chem. Phys.* **2020**, *153*, 244113.
- (73) Jmol: An Open-Source Java Viewer for Chemical Structures in 3D. <http://www.jmol.org/> (accessed June 12, 2025).
- (74) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular electronic-structure theory*; Wiley: New York, 2000.
- (75) Kriebel, M. H.; Tecmer, P.; Gałyńska, M.; Leszczyk, A.; Katharina, B. Accelerating Pythonic coupled cluster implementations: a comparison between CPUs and GPUs. *J. Chem. Theory Comput.* **2024**, *24*, 1130–1142.
- (76) Sadlej, A. J. Medium-size polarized basis sets for high-level correlated calculations of molecular electric properties. *Collect. Czech. Chem. Commun.* **1988**, *53*, 1995–2016.
- (77) Sadlej, A. J. Medium-size polarized basis sets for high-level-correlated calculations of molecular electric properties: II. Second-row atoms: Si through Cl. *Theor. Chem. Acc.* **1991**, *79*, 123–140.
- (78) Hickey, A. L.; Rowley, C. N. Benchmarking quantum chemical methods for the calculation of molecular dipole moments and polarizabilities. *J. Phys. Chem. A* **2014**, *118*, 3678–3687.
- (79) Pipek, J.; Mezey, P. G. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.* **1989**, *90*, 4916–4926.
- (80) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: A General Purpose Quantum Chemistry Program Package. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 242–253.
- (81) Werner, H.-J.; Knowles, P. J.; Manby, F. R.; Black, J. A.; Doll, K.; Heßelmann, A.; Kats, D.; Köhn, A.; Korona, T.; Kreplin, D. A.; et al. The Molpro quantum chemistry package. *J. Chem. Phys.* **2020**, *152*, 144107.
- (82) Werner, H.-J.; Knowles, P. J.; Celani, P.; Györfy, W.; Hesselmann, A.; Kats, D.; Knizia, G.; Köhn, A.; Korona, T.; Kreplin, D.; et al. MOLPRO, version 2020.2.1, a package of ab initio programs. 2020; see <http://www.molpro.net> (accessed 04 09, 2024).
- (83) Liu, X.; Meijer, G.; Pérez-Ríos, J. A data-driven approach to determine dipole moments of diatomic molecules. *Phys. Chem. Chem. Phys.* **2020**, *22*, 24191–24200.
- (84) Jensen, L.; Åstrand, P.-O.; Osted, A.; Kongsted, J.; Mikkelsen, K. V. Polarizability of molecular clusters as calculated by a dipole interaction model. *J. Chem. Phys.* **2002**, *116*, 4001–4010.
- (85) Scuseria, G. E.; Miller, M. D.; Jensen, F.; Geertsen, J. The dipole moment of carbon monoxide. *J. Chem. Phys.* **1991**, *94*, 6660–6663.
- (86) Samanta, P. K.; Köhn, A. First-order properties from internally contracted multireference coupled-cluster theory with particular focus on hyperfine coupling tensors. *J. Chem. Phys.* **2018**, *149*, 064101.
- (87) Tennyson, J. Vibration–rotation transition dipoles from first principles. *J. Mol. Spectrosc.* **2014**, *298*, 1–6.
- (88) Boguslawski, K.; Tecmer, P. Orbital entanglement in quantum chemistry. *Int. J. Quantum Chem.* **2015**, *115*, 1289–1295.
- (89) Boguslawski, K.; Tecmer, P. Erratum: Orbital entanglement in quantum chemistry. *Int. J. Quantum Chem.* **2017**, *117*, e25455.
- (90) Nowak, A.; Legeza, O.; Boguslawski, K. Orbital entanglement and correlation from pCCD-tailored coupled cluster wave functions. *J. Chem. Phys.* **2021**, *154*, 084111.



CAS BIOFINDER DISCOVERY PLATFORM™

BRIDGE BIOLOGY AND CHEMISTRY FOR FASTER ANSWERS

Analyze target relationships,
compound effects, and disease
pathways

Explore the platform

CAS 
A division of the
American Chemical Society

Expectation Value-pCCD-based Methods for Single-Electron Properties

Rahul Chakraborty,¹ Somayeh Ahmadkhani,^{1,*} Julian Świerczyński,¹
Katharina Boguslawski,¹ and Paweł Tecmer^{1,*}

¹*Institute of Physics, Faculty of Physics, Astronomy and Informatics,
Nicolaus Copernicus University in Torun,
Grudziadzka 5, 87-100 Torun, Poland*

Corresponding authors: so.ahmadkhani@gmail.com, ptecmer@fizyka.umk.pl

Supplementary Information

S1 Potential energy surface of HF molecule for different methods in cc-pVDZ basis set

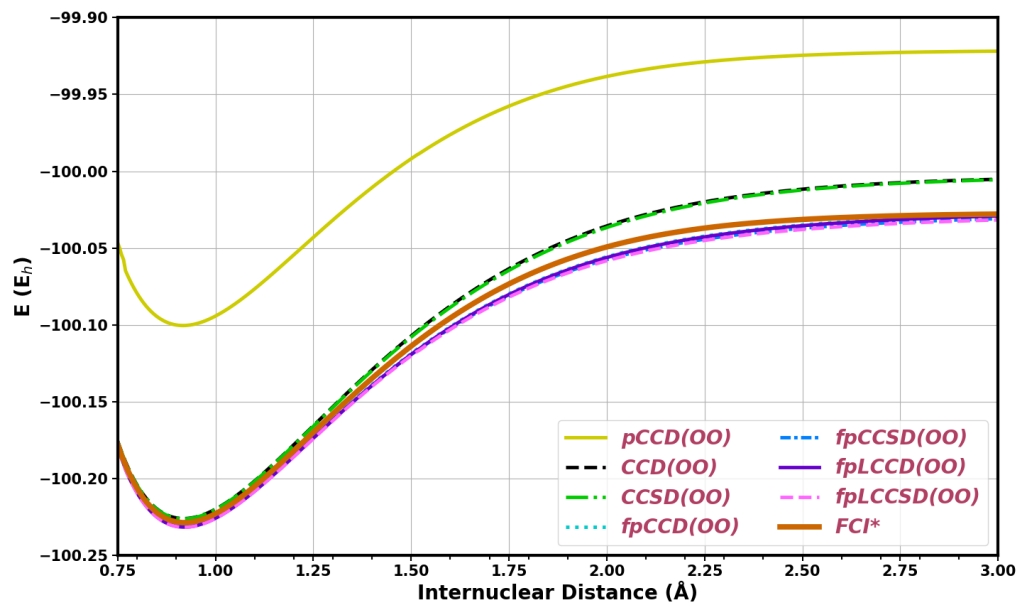


Figure S1: Potential energy surfaces of the HF molecule obtained from different CC flavours investigated in this work using cc-pVDZ atomic basis set. (OO) refers to the use of pCCD-optimized molecular orbitals. The FCI* energies are taken from Samanta et al.[?].

S2 Deviations in dipole moment surfaces of HF molecule for different methods in cc-pVDZ basis set

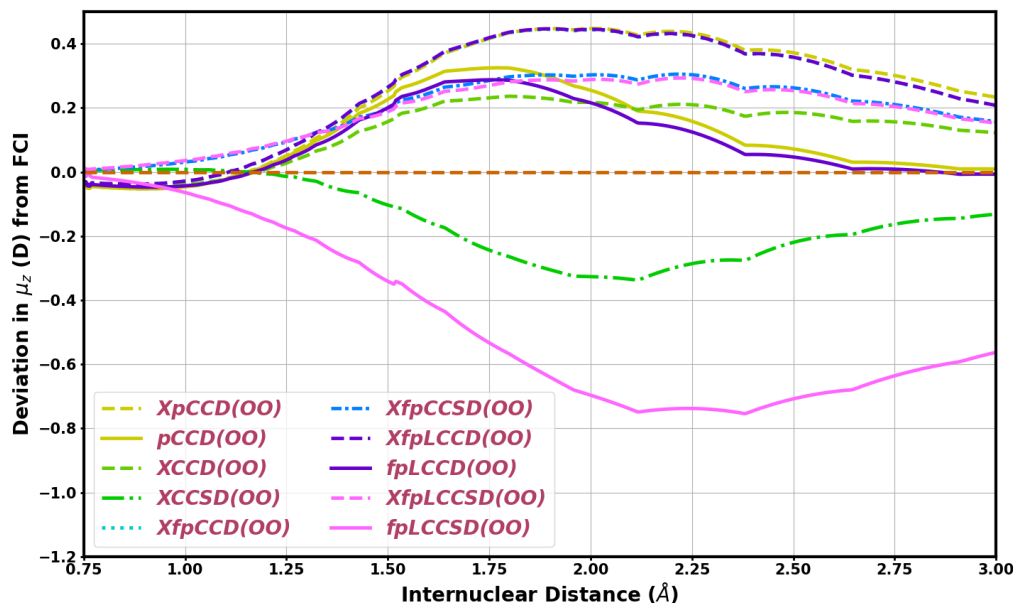


Figure S2: Deviations from the reference FCI/cc-pVDZ curve ($(\mu_{\text{Method}} - \mu_{\text{FCI}})$) for different methods. The horizontal line at 0 D denotes the reference line. (OO) denotes pCCD-optimized molecular orbitals. The FCI* curve is taken from Samanta et al.[?].

Figure S2 shows the deviations for each method from the reference FCI line calculated as $\mu_{\text{Method}} - \mu_{\text{FCI}}$ at each point. It is to be noted that μ_z is negative in our case. Hence, a positive deviation from FCI indicates underestimation of the dipole moment, and negative deviation meaning vice versa. At shorter distances, XCCSD performs very well and lies almost on the horizontal line at error = 0 D. However, from 1.25 Å onwards it starts overestimating the dipole moment, showing higher magnitude of μ_z and negative deviation. XpCCD, XCCD, and XfpCCD start with showing negative deviation and cross the reference FCI line at around 1.15 Å. On the other hand, XfpCCSD underestimates the dipole moment values throughout the bond distances. Response fpLCCD DMS shows better behaviour than its expectation-value counterpart, whereas the opposite is true for fpLCCSD.

S3 Dipole moment surface of HF in Sadlej pVTZ basis set

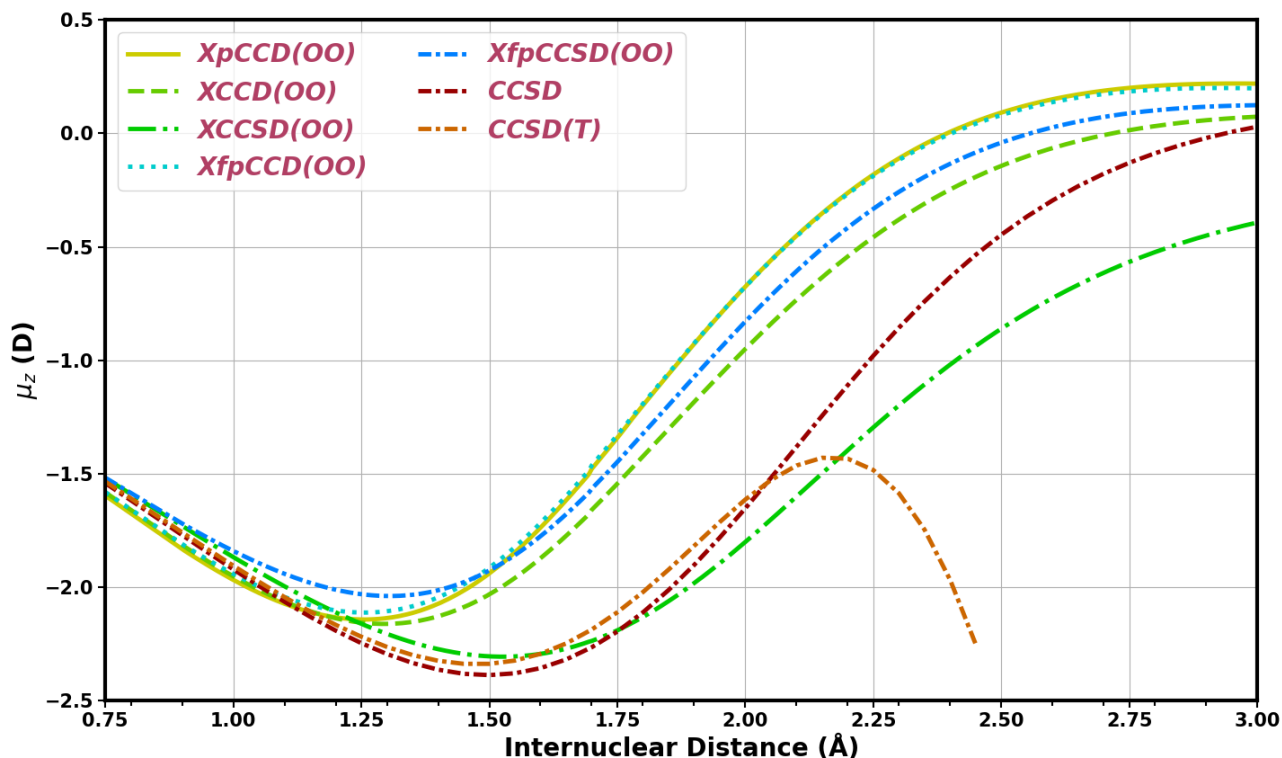


Figure S3: Dipole moment surface of HF in Sadlej-pVTZ atomic basis set calculated with different methods. (OO) denotes pCCD-optimized orbitals. Orbital relaxed CCSD and CCSD(T) response dipole moment curves calculated with canonical HF molecular orbitals are shown for comparison.

The apparent breakdown of CCSD(T) after 2.15 Å is to be noted.

S4 Statistical measurements for dipole moment against CCSD as reference

Table S1: Statistical analysis of performances of various methods with respect to the dipole moment (μ in D) values obtained with them using Sadlej-pVTZ basis set. The three performance measures analyzed here are- root-mean-square-error (RMSE): $\sqrt{(\sum_i^N (\mu_{\text{Method},i} - \mu_{\text{CCSD},i})^2)/N}$, mean-error (ME): $\sum_i^N (\mu_{\text{Method},i} - \mu_{\text{CCSD},i})/N$, mean-absolute-error (MAE): $\sum_i^N |(\mu_{\text{Method},i} - \mu_{\text{CCSD},i})|/N$, and mean-absolute-percentage-error (MAPE): $\sum_i^N [|(\mu_{\text{Method},i} - \mu_{\text{CCSD},i})| / (\mu_{\text{CCSD},i}) * 100] / N$. The three sets of molecules have been described in section *Computational Details* in the main article. (HF) and (OO) denote the use of canonical and pCCD-optimized orbitals, respectively. CO and SiSe have been excluded from the analysis.

Method	RMSE (D)			ME (D)			MAE (D)			MAPE		
	Full set	Small	Organic	Full set	Small	Organic	Full set	Small	Organic	Full set	Small	Organic
XpCCD(HF)	0.288	0.298	0.276	0.179	0.142	0.222	0.236	0.248	0.224	12.0	12.7	11.1
XpCCD(OO)	0.254	0.332	0.107	0.117	0.167	0.059	0.137	0.181	0.087	10.0	10.2	9.9
XCCD(HF)	0.253	0.265	0.237	0.154	0.126	0.186	0.205	0.221	0.188	10.3	11.2	9.3
XCCD(OO)	0.197	0.257	0.086	0.088	0.128	0.043	0.111	0.148	0.068	8.2	8.3	8.1
XCCSD(HF)	0.180	0.225	0.108	-0.080	-0.074	-0.087	0.124	0.154	0.088	6.1	7.6	4.3
XCCSD(OO)	0.173	0.219	0.095	-0.076	-0.074	-0.077	0.116	0.149	0.079	5.7	7.3	4.0
XfpCCD(HF)	0.247	0.256	0.236	0.155	0.128	0.186	0.201	0.214	0.187	10.2	11.0	9.3
XfpCCD(OO)	0.200	0.264	0.080	0.077	0.126	0.021	0.110	0.151	0.063	8.4	8.7	8.0
XfpCCSD(HF)	0.174	0.219	0.100	-0.066	-0.053	-0.081	0.121	0.154	0.083	5.8	7.3	4.0
XfpCCSD(OO)	0.229	0.270	0.171	-0.097	-0.067	-0.132	0.157	0.179	0.132	7.2	8.3	6.0
XfpLCCD(HF)	0.235	0.243	0.225	0.147	-0.121	0.177	0.192	0.204	0.178	9.8	10.4	9.0
XfpLCCD(OO)	0.194	0.255	0.078	0.073	0.120	0.018	0.108	0.149	0.061	8.2	8.6	7.9
XfpLCCSD(HF)	0.401	0.402	0.401	-0.185	0.139	-0.238	0.256	0.270	0.241	11.9	13.2	10.3
XfpLCCSD(OO)	0.322	0.384	0.233	-0.155	-0.137	-0.176	0.219	0.256	0.178	10.4	12.0	8.6

S5 Statistical measurements for quadrupole moment against CCSD as reference

Table S2: Statistical analysis of performances of various methods with respect to the quadrupole moment (Q_{zz} in a.u.) values obtained with them using Sadlej-pVTZ basis set. The three performance measures analyzed here are- root-mean-square-error (RMSE): $\sqrt{(\sum_i^N (Q_{zz}^{\text{Method},i} - Q_{zz}^{\text{CCSD},i})^2)/N}$, mean-error (ME): $\sum_i^N (Q_{zz}^{\text{Method},i} - Q_{zz}^{\text{CCSD},i})/N$, mean-absolute-error (MAE): $\sum_i^N |Q_{zz}^{\text{Method},i} - Q_{zz}^{\text{CCSD},i}|/N$, and mean-absolute-percentage-error (MAPE): $\sum_i^N [|Q_{zz}^{\text{Method},i} - Q_{zz}^{\text{CCSD},i}|/(Q_{zz}^{\text{CCSD},i})] * 100/N$. The three sets of molecules have been described in section *Computational Details* in the main article. (HF) and (OO) denote the use of canonical and pCCD-optimized orbitals, respectively. ClF has been excluded from the analysis.

Method	RMSE (D)			ME (D)			MAE (D)			MAPE		
	Full set	Small	Organic	Full set	Small	Organic	Full set	Small	Organic	Full set	Small	Organic
XpCCD(HF)	0.552	0.599	0.487	0.325	0.273	0.388	0.412	0.426	0.395	11.5	13.3	9.4
XpCCD(OO)	0.483	0.619	0.224	0.218	0.331	0.080	0.256	0.350	0.141	8.1	12.2	3.2
XCCD(HF)	0.463	0.492	0.425	-0.284	-0.243	-0.335	0.343	0.345	0.340	9.8	11.3	7.9
XCCD(OO)	0.385	0.493	0.180	0.169	0.262	0.054	0.206	0.281	0.113	6.5	9.7	2.5
XCCSD(HF)	0.262	0.318	0.169	-0.127	-0.136	-0.115	0.160	0.195	0.117	4.9	7.0	2.2
XCCSD(OO)	0.255	0.317	0.145	-0.121	-0.140	-0.098	0.150	0.191	0.100	4.7	6.9	1.9
XfpCCD(HF)	0.464	0.496	0.422	0.283	0.244	0.333	0.344	0.349	0.338	8.5	9.6	7.3
XfpCCD(OO)	0.385	0.497	0.164	0.154	0.261	0.022	0.204	0.280	0.112	6.6	9.8	2.7
XfpCCSD(HF)	0.237	0.286	0.157	-0.104	-0.101	-0.108	0.152	0.186	0.111	4.5	6.5	2.1
XfpCCSD(OO)	0.283	0.305	0.252	-0.143	-0.111	-0.182	0.183	0.180	0.185	5.4	6.7	3.8
XfpLCCD(HF)	0.370	0.400	0.329	0.208	0.174	0.250	0.279	0.299	0.254	7.6	8.9	6.0
XfpLCCD(OO)	0.448	0.521	0.339	0.003	0.197	-0.236	0.298	0.327	0.262	10.2	12.4	7.4
XfpLCCSD(HF)	0.713	0.668	0.765	-0.355	-0.306	-0.415	0.415	0.406	0.425	10.8	13.3	7.7
XfpLCCSD(OO)	0.639	0.638	0.641	-0.379	-0.286	-0.492	0.443	0.390	0.508	12.6	13.3	11.7

S6 Dipole moment values of HF at different bond lengths

Table S3: Dipole moments (in D) of the HF molecule with the cc-pVDZ basis set calculated with the different expectation-value methods. OO denotes pCCD-optimized molecular orbitals.

r(Å)	XpCCD(OO)	XCCD(OO)	XCCSD(OO)	XfpCCD(OO)	XfpCCSD(OO)	XfpLCCD(OO)	XfpLCCSD(OO)
0.750	-1.599155	-1.589988	-1.588570	-1.55081	-1.588235	-1.593182	-1.571930
0.755	-1.607283	-1.598066	-1.596569	-1.558356	-1.596313	-1.601414	-1.580468
0.760	-1.624447	-1.614320	-1.613515	-1.568842	-1.613326	-1.617612	-1.583658
0.765	-1.633130	-1.622888	-1.622034	-1.576523	-1.621746	-1.626179	-1.592128
0.770	-1.635039	-1.625248	-1.623971	-1.583046	-1.623724	-1.629050	-1.606332
0.775	-1.643481	-1.633615	-1.632268	-1.590670	-1.631982	-1.637450	-1.614871
0.780	-1.651885	-1.641943	-1.640525	-1.598295	-1.640245	-1.645857	-1.623424
0.785	-1.660284	-1.650281	-1.648784	-1.605897	-1.648487	-1.654246	-1.631996
0.790	-1.668621	-1.658550	-1.656975	-1.613505	-1.656705	-1.662613	-1.640580
0.795	-1.676984	-1.666862	-1.665200	-1.621085	-1.664915	-1.670974	-1.649180
0.800	-1.685291	-1.675115	-1.673366	-1.628664	-1.673091	-1.679305	-1.657789
0.805	-1.693678	-1.683456	-1.681614	-1.636226	-1.681286	-1.687657	-1.666408
0.810	-1.702018	-1.691742	-1.689810	-1.643792	-1.689467	-1.695997	-1.675032
0.815	-1.710296	-1.699982	-1.697950	-1.651323	-1.697593	-1.704285	-1.683669
0.820	-1.718615	-1.708255	-1.706125	-1.658853	-1.705730	-1.712586	-1.692308
0.825	-1.726854	-1.716462	-1.714226	-1.666351	-1.713814	-1.720838	-1.700955
0.830	-1.735080	-1.724645	-1.722306	-1.673849	-1.721882	-1.729075	-1.709605
0.835	-1.743135	-1.732699	-1.730236	-1.681261	-1.729851	-1.737219	-1.718264
0.840	-1.751397	-1.740930	-1.738355	-1.688712	-1.737896	-1.745439	-1.726913
0.845	-1.759415	-1.748935	-1.746238	-1.696106	-1.745817	-1.753539	-1.735572
0.850	-1.767461	-1.756976	-1.754152	-1.703471	-1.753723	-1.761627	-1.744225
0.855	-1.775472	-1.764980	-1.762026	-1.710815	-1.761600	-1.769689	-1.752876
0.860	-1.783494	-1.772997	-1.769911	-1.718132	-1.769454	-1.777731	-1.761520
0.865	-1.791366	-1.780881	-1.777654	-1.725397	-1.777212	-1.785680	-1.770165
0.870	-1.799252	-1.788776	-1.785406	-1.732638	-1.784952	-1.793614	-1.778800
0.875	-1.807183	-1.796716	-1.793201	-1.739854	-1.792679	-1.801538	-1.787424
0.880	-1.814928	-1.804478	-1.800812	-1.747021	-1.800275	-1.809335	-1.796047
0.885	-1.822610	-1.812195	-1.808367	-1.754127	-1.807854	-1.817118	-1.804656
0.890	-1.830262	-1.819879	-1.815889	-1.761205	-1.815389	-1.824859	-1.813251
0.895	-1.837862	-1.827514	-1.823357	-1.768239	-1.822852	-1.832532	-1.821835
0.900	-1.845378	-1.835076	-1.830745	-1.775222	-1.830239	-1.840132	-1.830408
0.905	-1.852871	-1.842615	-1.838108	-1.782164	-1.837584	-1.847694	-1.838961
0.910	-1.860232	-1.850035	-1.845343	-1.789042	-1.844830	-1.855161	-1.847501
0.915	-1.867791	-1.857635	-1.852764	-1.795913	-1.852117	-1.862671	-1.856018
0.920	-1.874942	-1.864874	-1.859799	-1.802660	-1.859192	-1.869973	-1.864523
0.925	-1.882025	-1.872033	-1.866757	-1.809374	-1.866198	-1.877208	-1.873005
0.930	-1.889113	-1.879208	-1.873724	-1.816028	-1.873156	-1.884400	-1.881461
0.935	-1.896124	-1.886314	-1.880614	-1.822623	-1.880035	-1.891515	-1.889892
0.940	-1.903055	-1.893350	-1.887427	-1.829152	-1.886832	-1.898552	-1.898296
0.945	-1.909941	-1.900346	-1.894195	-1.835658	-1.893584	-1.905547	-1.906708
0.950	-1.916659	-1.907192	-1.900801	-1.842045	-1.900190	-1.912399	-1.915058
0.955	-1.923290	-1.913959	-1.907322	-1.848363	-1.906710	-1.919167	-1.923377
0.960	-1.929876	-1.920687	-1.913797	-1.854618	-1.913152	-1.925861	-1.931662
0.965	-1.936307	-1.927276	-1.920123	-1.860787	-1.919476	-1.932441	-1.939915
0.970	-1.942648	-1.933784	-1.926361	-1.866881	-1.925707	-1.938930	-1.948133

Continued on next page

Table S3 – Continued from previous page

r(Å)	pCCD(OO)	CCD(OO)	CCSD(OO)	fpCCD(OO)	fpCCSD(OO)	fpLCCD(OO)	fpLCCSD(OO)
0.975	-1.948899	-1.940212	-1.932511	-1.872900	-1.931843	-1.945328	-1.956312
0.980	-1.955004	-1.946511	-1.938521	-1.878828	-1.937858	-1.951607	-1.964456
0.985	-1.961032	-1.952739	-1.944454	-1.884681	-1.943778	-1.957795	-1.972559
0.990	-1.966924	-1.958845	-1.950254	-1.890444	-1.949581	-1.963869	-1.980620
0.995	-1.972728	-1.964874	-1.955969	-1.896125	-1.955283	-1.969845	-1.988638
1.000	-1.978447	-1.970827	-1.961601	-1.901721	-1.960883	-1.975723	-1.996610
1.005	-1.983972	-1.976604	-1.967044	-1.907216	-1.966339	-1.981460	-2.004540
1.010	-1.989419	-1.982312	-1.972410	-1.912623	-1.971693	-1.987097	-2.012420
1.015	-1.994744	-1.987912	-1.977658	-1.917936	-1.976927	-1.992619	-2.020250
1.020	-1.999933	-1.993389	-1.982773	-1.923151	-1.982033	-1.998015	-2.028030
1.025	-2.004999	-1.998755	-1.987768	-1.928268	-1.987016	-2.003292	-2.035757
1.030	-2.010172	-2.004222	-1.992862	-1.933323	-1.992038	-2.008612	-2.043420
1.035	-2.014747	-2.009145	-1.997384	-1.938200	-1.996606	-2.013478	-2.051048
1.040	-2.019417	-2.014157	-2.001993	-1.943011	-2.001204	-2.018379	-2.058609
1.045	-2.023951	-2.019047	-2.006469	-1.947715	-2.005668	-2.023149	-2.066110
1.050	-2.028343	-2.023808	-2.010806	-1.952313	-2.009998	-2.027788	-2.073551
1.055	-2.032621	-2.028469	-2.015033	-1.956803	-2.014205	-2.032307	-2.080928
1.060	-2.037024	-2.033244	-2.019374	-1.961230	-2.018466	-2.036886	-2.088229
1.065	-2.040707	-2.037367	-2.023028	-1.965445	-2.022178	-2.040912	-2.095489
1.070	-2.044508	-2.041599	-2.026790	-1.969591	-2.025936	-2.044990	-2.102671
1.075	-2.048187	-2.045721	-2.030432	-1.973623	-2.029563	-2.048941	-2.109781
1.080	-2.051822	-2.049809	-2.034031	-1.977553	-2.033123	-2.052829	-2.116816
1.085	-2.055094	-2.053562	-2.037277	-1.981333	-2.036378	-2.056413	-2.123786
1.090	-2.058544	-2.057480	-2.040689	-1.985047	-2.039728	-2.060098	-2.130670
1.095	-2.061386	-2.060851	-2.043522	-1.988558	-2.042600	-2.063302	-2.137492
1.100	-2.064291	-2.064280	-2.046409	-1.991984	-2.045482	-2.066521	-2.144228
1.105	-2.067050	-2.067577	-2.049154	-1.995285	-2.048212	-2.069592	-2.150883
1.110	-2.069636	-2.070719	-2.051730	-1.998458	-2.050784	-2.072506	-2.157456
1.115	-2.072125	-2.073778	-2.054212	-2.001509	-2.053230	-2.075299	-2.163940
1.120	-2.074375	-2.076617	-2.056459	-2.004422	-2.055474	-2.077891	-2.170343
1.125	-2.076542	-2.079388	-2.058628	-2.007210	-2.057618	-2.080385	-2.176654
1.130	-2.078524	-2.081989	-2.060613	-2.009868	-2.059583	-2.082703	-2.182872
1.135	-2.080178	-2.084291	-2.062281	-2.012366	-2.061271	-2.084744	-2.189007
1.140	-2.081986	-2.086741	-2.064094	-2.014777	-2.063015	-2.086848	-2.195032
1.145	-2.083235	-2.088681	-2.065370	-2.016991	-2.064336	-2.088525	-2.200977
1.150	-2.084551	-2.090690	-2.066709	-2.019099	-2.065648	-2.090197	-2.206812
1.155	-2.085606	-2.092458	-2.067792	-2.021061	-2.066728	-2.091637	-2.212547
1.160	-2.086536	-2.094116	-2.068752	-2.022887	-2.067674	-2.092946	-2.218181
1.165	-2.087299	-2.095623	-2.069549	-2.024576	-2.068471	-2.094105	-2.223707
1.170	-2.087894	-2.096982	-2.070184	-2.026122	-2.069094	-2.095092	-2.229122
1.175	-2.088376	-2.098245	-2.070709	-2.027528	-2.069581	-2.095945	-2.234424
1.180	-2.088583	-2.099250	-2.070962	-2.028783	-2.069826	-2.096554	-2.239616
1.185	-2.088618	-2.100100	-2.071047	-2.029896	-2.069933	-2.097025	-2.244694
1.190	-2.088580	-2.100893	-2.071063	-2.030870	-2.069894	-2.097353	-2.249646
1.195	-2.088335	-2.101498	-2.070877	-2.031696	-2.069687	-2.097513	-2.254486
1.200	-2.087734	-2.101770	-2.070341	-2.032353	-2.069186	-2.097374	-2.259208
1.205	-2.087105	-2.102025	-2.069777	-2.032879	-2.068618	-2.097173	-2.263806
1.210	-2.086301	-2.102123	-2.069042	-2.033253	-2.067859	-2.096781	-2.268274
1.215	-2.085334	-2.102071	-2.068145	-2.033480	-2.066935	-2.096223	-2.272618
1.220	-2.084140	-2.101816	-2.067030	-2.033546	-2.065817	-2.095473	-2.276839

Continued on next page

Table S3 – Continued from previous page

r(Å)	pCCD(OO)	CCD(OO)	CCSD(OO)	fpCCD(OO)	fpCCSD(OO)	fpLCCD(OO)	fpLCCSD(OO)
1.225	-2.082862	-2.101484	-2.065827	-2.033473	-2.064584	-2.094611	-2.280929
1.230	-2.081362	-2.100948	-2.064404	-2.033240	-2.063135	-2.093532	-2.284890
1.235	-2.079435	-2.100014	-2.062565	-2.032819	-2.061351	-2.092117	-2.288734
1.240	-2.077522	-2.099102	-2.060738	-2.032262	-2.059504	-2.090646	-2.292439
1.245	-2.075413	-2.098008	-2.058714	-2.031545	-2.057470	-2.088991	-2.296016
1.250	-2.073104	-2.096732	-2.056495	-2.030665	-2.055236	-2.087138	-2.299460
1.255	-2.070597	-2.095275	-2.054080	-2.029620	-2.052816	-2.085104	-2.302776
1.260	-2.067885	-2.093630	-2.051461	-2.028407	-2.050186	-2.082863	-2.305956
1.265	-2.064985	-2.091813	-2.048656	-2.027027	-2.047381	-2.080452	-2.309007
1.270	-2.061865	-2.089794	-2.045635	-2.025476	-2.044354	-2.077824	-2.311919
1.275	-2.058745	-2.087775	-2.042604	-2.023787	-2.041259	-2.075142	-2.314703
1.280	-2.055297	-2.085448	-2.039249	-2.021912	-2.037891	-2.072186	-2.317358
1.285	-2.051568	-2.082870	-2.035623	-2.019841	-2.034244	-2.068957	-2.319861
1.290	-2.047303	-2.079802	-2.031478	-2.017529	-2.030172	-2.065295	-2.322198
1.295	-2.043183	-2.076860	-2.027459	-2.015110	-2.026138	-2.061694	-2.324431
1.300	-2.038795	-2.073672	-2.023175	-2.012497	-2.021835	-2.057830	-2.326513
1.305	-2.034362	-2.070451	-2.018845	-2.009729	-2.017460	-2.053908	-2.328480
1.310	-2.029265	-2.066600	-2.013863	-2.006701	-2.012516	-2.049410	-2.330256
1.315	-2.024180	-2.062768	-2.008887	-2.003530	-2.007531	-2.044891	-2.331938
1.320	-2.018868	-2.058727	-2.003686	-2.000174	-2.002327	-2.040165	-2.333502
1.325	-2.013336	-2.054487	-1.998268	-1.996630	-1.996886	-2.035214	-2.334941
1.330	-2.007559	-2.050016	-1.992604	-1.992908	-1.991228	-2.030058	-2.336276
1.335	-2.001557	-2.045345	-1.986722	-1.989000	-1.985333	-2.024679	-2.337508
1.340	-1.995363	-2.040500	-1.980651	-1.984923	-1.979251	-2.019126	-2.338633
1.345	-1.988983	-2.035489	-1.974398	-1.980678	-1.972984	-2.013403	-2.339656
1.350	-1.982398	-2.030293	-1.967943	-1.976261	-1.966515	-2.007491	-2.340575
1.355	-1.975783	-2.025072	-1.961453	-1.971722	-1.960011	-2.001560	-2.341404
1.360	-1.968660	-2.019376	-1.954467	-1.966938	-1.953056	-1.995180	-2.342097
1.365	-1.961704	-2.013856	-1.947647	-1.962068	-1.946192	-1.988917	-2.342704
1.370	-1.954425	-2.008033	-1.940508	-1.957006	-1.939051	-1.982384	-2.343185
1.375	-1.946959	-2.002040	-1.933185	-1.951777	-1.931728	-1.975681	-2.343547
1.380	-1.939600	-1.996159	-1.925963	-1.946439	-1.924419	-1.969009	-2.343784
1.385	-1.931986	-1.990039	-1.918489	-1.940931	-1.916915	-1.962152	-2.343902
1.390	-1.923716	-1.983302	-1.910377	-1.935166	-1.908891	-1.954781	-2.343897
1.395	-1.915852	-1.976952	-1.902650	-1.929363	-1.901132	-1.947693	-2.343759
1.400	-1.907727	-1.970365	-1.894669	-1.923380	-1.893141	-1.940387	-2.343491
1.405	-1.899576	-1.963762	-1.886663	-1.917263	-1.885083	-1.933028	-2.343086
1.410	-1.891175	-1.956922	-1.878408	-1.910977	-1.876818	-1.925472	-2.342552
1.415	-1.882426	-1.949756	-1.869810	-1.904530	-1.868311	-1.917677	-2.341935
1.420	-1.873844	-1.942753	-1.861369	-1.897963	-1.859834	-1.909935	-2.341132
1.425	-1.864928	-1.935441	-1.852605	-1.891201	-1.851081	-1.901927	-2.340198
1.430	-1.856115	-1.928220	-1.843928	-1.884350	-1.842375	-1.893982	-2.339124
1.435	-1.847010	-1.920737	-1.834972	-1.877310	-1.833400	-1.885779	-2.337912
1.440	-1.837778	-1.913133	-1.825886	-1.870126	-1.824308	-1.877473	-2.336572
1.445	-1.828469	-1.905456	-1.816720	-1.862812	-1.815137	-1.869101	-2.335101
1.450	-1.818982	-1.897617	-1.807379	-1.855336	-1.805771	-1.860548	-2.333489
1.455	-1.809414	-1.889695	-1.797950	-1.847737	-1.796336	-1.851940	-2.331751
1.460	-1.799484	-1.881449	-1.788177	-1.839931	-1.786601	-1.843042	-2.329881
1.465	-1.789641	-1.873277	-1.778478	-1.832038	-1.776868	-1.834167	-2.327872
1.470	-1.779658	-1.864976	-1.768640	-1.824002	-1.767001	-1.825170	-2.325727

Continued on next page

Table S3 – Continued from previous page

r(Å)	pCCD(OO)	CCD(OO)	CCSD(OO)	fpCCD(OO)	fpCCSD(OO)	fpLCCD(OO)	fpLCCSD(OO)
1.475	-1.769519	-1.856528	-1.758646	-1.815819	-1.756990	-1.816043	-2.323455
1.480	-1.759001	-1.847729	-1.748286	-1.807439	-1.746713	-1.806663	-2.321079
1.485	-1.748725	-1.839153	-1.738149	-1.799004	-1.736520	-1.797383	-2.318538
1.490	-1.738202	-1.830350	-1.727777	-1.790405	-1.726112	-1.787909	-2.315874
1.495	-1.727686	-1.821540	-1.717396	-1.781708	-1.715686	-1.778429	-2.313083
1.500	-1.716716	-1.812314	-1.706583	-1.772808	-1.704953	-1.768651	-2.310172
1.505	-1.705796	-1.803127	-1.695807	-1.763809	-1.694192	-1.758864	-2.307137
1.510	-1.694820	-1.793883	-1.684970	-1.754695	-1.683360	-1.749021	-2.303973
1.515	-1.683615	-1.784446	-1.673926	-1.745414	-1.672270	-1.738945	-2.300669
1.520	-1.663418	-1.770725	-1.656488	-1.729009	-1.655081	-1.720319	-2.285734
1.525	-1.652692	-1.761732	-1.645901	-1.719637	-1.644185	-1.710484	-2.282135
1.530	-1.641006	-1.751803	-1.634362	-1.709975	-1.632806	-1.700150	-2.278442
1.535	-1.629668	-1.742211	-1.623159	-1.700271	-1.621493	-1.689926	-2.274625
1.540	-1.617777	-1.732103	-1.611422	-1.690346	-1.609804	-1.679328	-2.270704
1.545	-1.605962	-1.722060	-1.599753	-1.680350	-1.598165	-1.668804	-2.266674
1.550	-1.594195	-1.712049	-1.588116	-1.670270	-1.586510	-1.658280	-2.262539
1.555	-1.582185	-1.701828	-1.576261	-1.660036	-1.574592	-1.647521	-2.258273
1.560	-1.570147	-1.691550	-1.564353	-1.649723	-1.562694	-1.636791	-2.253916
1.565	-1.557810	-1.681009	-1.552168	-1.639234	-1.550505	-1.625786	-2.249450
1.570	-1.545849	-1.670812	-1.540339	-1.628755	-1.538556	-1.615056	-2.244867
1.575	-1.533390	-1.660122	-1.528010	-1.618089	-1.526321	-1.604030	-2.240201
1.580	-1.520935	-1.649449	-1.515696	-1.607330	-1.513977	-1.592931	-2.235416
1.585	-1.507981	-1.638326	-1.502914	-1.596348	-1.501190	-1.581409	-2.230548
1.590	-1.495601	-1.627682	-1.490639	-1.585463	-1.488944	-1.570434	-2.225579
1.595	-1.482455	-1.616351	-1.477650	-1.574283	-1.476020	-1.558798	-2.220532
1.600	-1.469616	-1.605328	-1.464976	-1.563099	-1.463238	-1.547344	-2.215367
1.605	-1.456701	-1.594182	-1.452191	-1.551849	-1.450452	-1.535898	-2.210069
1.610	-1.443741	-1.582995	-1.439368	-1.540515	-1.437626	-1.524424	-2.204696
1.615	-1.430654	-1.571663	-1.426406	-1.529101	-1.424786	-1.512943	-2.199243
1.620	-1.417513	-1.560289	-1.413399	-1.517586	-1.411813	-1.501351	-2.193687
1.625	-1.404440	-1.548984	-1.400469	-1.506016	-1.398755	-1.489709	-2.188003
1.630	-1.391253	-1.537563	-1.387427	-1.494365	-1.385674	-1.478058	-2.182235
1.635	-1.378027	-1.526072	-1.374327	-1.482658	-1.372607	-1.466425	-2.176384
1.640	-1.364831	-1.514616	-1.361268	-1.470897	-1.359500	-1.454781	-2.170419
1.645	-1.351459	-1.502962	-1.348019	-1.459055	-1.346395	-1.443137	-2.164400
1.650	-1.338216	-1.491455	-1.334922	-1.447161	-1.333144	-1.431394	-2.158239
1.655	-1.324875	-1.479824	-1.321712	-1.435210	-1.319932	-1.419691	-2.152013
1.660	-1.311506	-1.468156	-1.308474	-1.423203	-1.306690	-1.407976	-2.145700
1.665	-1.298115	-1.456455	-1.295213	-1.411145	-1.293426	-1.396256	-2.139302
1.670	-1.284705	-1.444725	-1.281932	-1.399037	-1.280142	-1.384532	-2.132819
1.675	-1.271277	-1.432966	-1.268633	-1.386883	-1.266839	-1.372807	-2.126253
1.680	-1.257758	-1.421085	-1.255226	-1.374683	-1.253588	-1.361132	-2.119635
1.685	-1.244363	-1.409354	-1.241974	-1.362438	-1.240179	-1.349349	-2.112884
1.690	-1.230911	-1.397530	-1.228646	-1.350162	-1.226848	-1.337643	-2.106081
1.695	-1.217362	-1.385582	-1.215206	-1.337844	-1.213567	-1.325987	-2.099233
1.700	-1.203930	-1.373757	-1.201901	-1.325503	-1.200216	-1.314301	-2.092270
1.705	-1.190408	-1.361822	-1.188498	-1.313112	-1.186859	-1.302610	-2.085259
1.710	-1.176932	-1.349921	-1.175143	-1.300699	-1.173503	-1.290942	-2.078167
1.715	-1.163444	-1.338032	-1.161806	-1.288240	-1.160038	-1.279206	-2.070977
1.720	-1.150049	-1.326169	-1.148521	-1.275792	-1.146702	-1.267585	-2.063745

Continued on next page

Table S3 – Continued from previous page

r(Å)	pCCD(OO)	CCD(OO)	CCSD(OO)	fpCCD(OO)	fpCCSD(OO)	fpLCCD(OO)	fpLCCSD(OO)
1.725	-1.136534	-1.314164	-1.135105	-1.263304	-1.133430	-1.256024	-2.056482
1.730	-1.123253	-1.302431	-1.121975	-1.250827	-1.120016	-1.244375	-2.049097
1.735	-1.109812	-1.290484	-1.108651	-1.238314	-1.106654	-1.232770	-2.041698
1.740	-1.096399	-1.278555	-1.095359	-1.225783	-1.093378	-1.221287	-2.034214
1.745	-1.082931	-1.266495	-1.081960	-1.213249	-1.080280	-1.209945	-2.026657
1.750	-1.069731	-1.254770	-1.068900	-1.200722	-1.066992	-1.198499	-2.019019
1.755	-1.056195	-1.242613	-1.055440	-1.188151	-1.053824	-1.187093	-2.011401
1.760	-1.042908	-1.230800	-1.042324	-1.175571	-1.040412	-1.175551	-2.003690
1.765	-1.029751	-1.219011	-1.029259	-1.163044	-1.027394	-1.164328	-1.995890
1.770	-1.016348	-1.206936	-1.015936	-1.150479	-1.014255	-1.153051	-1.988093
1.775	-1.003281	-1.195214	-1.002974	-1.137958	-1.001222	-1.141900	-1.980209
1.780	-0.990011	-1.183325	-0.989863	-1.125377	-0.988001	-1.130534	-1.972299
1.785	-0.976710	-1.171293	-0.976638	-1.112828	-0.974952	-1.119371	-1.964322
1.790	-0.963665	-1.159531	-0.963695	-1.100314	-0.961976	-1.108299	-1.956305
1.795	-0.950592	-1.147726	-0.950730	-1.087785	-0.948949	-1.097175	-1.948247
1.800	-0.937596	-1.135964	-0.937826	-1.075286	-0.936041	-1.086181	-1.940160
1.805	-0.924593	-1.124187	-0.924926	-1.062799	-0.923179	-1.075247	-1.932036
1.810	-0.911669	-1.112475	-0.912106	-1.050313	-0.910301	-1.064304	-1.923868
1.815	-0.898845	-1.100842	-0.899384	-1.037859	-0.897494	-1.053458	-1.915658
1.820	-0.885754	-1.088892	-0.886379	-1.025386	-0.884688	-1.042558	-1.907464
1.825	-0.873016	-1.077335	-0.873767	-1.012958	-0.871889	-1.031750	-1.899183
1.830	-0.860035	-1.065479	-0.860894	-1.000525	-0.859122	-1.020935	-1.890899
1.835	-0.847207	-1.053730	-0.848146	-0.988135	-0.846455	-1.010236	-1.882599
1.840	-0.834503	-1.042112	-0.835544	-0.975762	-0.833793	-0.999531	-1.874254
1.845	-0.821741	-1.030424	-0.822893	-0.963418	-0.821199	-0.988918	-1.865908
1.850	-0.809076	-1.018810	-0.810337	-0.951101	-0.808624	-0.978334	-1.857525
1.855	-0.796516	-1.007273	-0.797877	-0.938818	-0.796100	-0.967807	-1.849116
1.860	-0.783963	-0.995722	-0.785426	-0.926566	-0.783639	-0.957343	-1.840693
1.865	-0.771460	-0.984197	-0.773023	-0.914349	-0.771233	-0.946935	-1.832264
1.870	-0.759011	-0.972703	-0.760673	-0.902168	-0.758882	-0.936585	-1.823805
1.875	-0.746553	-0.961167	-0.748307	-0.890024	-0.746603	-0.926302	-1.815341
1.880	-0.734296	-0.949823	-0.736143	-0.877931	-0.734411	-0.916121	-1.806859
1.885	-0.722018	-0.938438	-0.723966	-0.865868	-0.722225	-0.905945	-1.798361
1.890	-0.709801	-0.927092	-0.711851	-0.853848	-0.710097	-0.895829	-1.789851
1.895	-0.697646	-0.915786	-0.699797	-0.841873	-0.698030	-0.885776	-1.781332
1.900	-0.685523	-0.904497	-0.687786	-0.829936	-0.686000	-0.875743	-1.772825
1.905	-0.673485	-0.893258	-0.675845	-0.818055	-0.674062	-0.865814	-1.764296
1.910	-0.661467	-0.882017	-0.663928	-0.806225	-0.662190	-0.855960	-1.755762
1.915	-0.649742	-0.871033	-0.652277	-0.794456	-0.650419	-0.846210	-1.747226
1.920	-0.637767	-0.859760	-0.640383	-0.782729	-0.638692	-0.836472	-1.738682
1.925	-0.626174	-0.848924	-0.628918	-0.771060	-0.627083	-0.826891	-1.730144
1.930	-0.614358	-0.837732	-0.617151	-0.759445	-0.615495	-0.817257	-1.721592
1.935	-0.602814	-0.826904	-0.605746	-0.747881	-0.603920	-0.807716	-1.713040
1.940	-0.591097	-0.815820	-0.594133	-0.736372	-0.592429	-0.798189	-1.704524
1.945	-0.579622	-0.804932	-0.582728	-0.724935	-0.581090	-0.788825	-1.695991
1.950	-0.568373	-0.794327	-0.571607	-0.713553	-0.569809	-0.779558	-1.687457
1.955	-0.556985	-0.783462	-0.560282	-0.702234	-0.558555	-0.770246	-1.678934
1.960	-0.545732	-0.772784	-0.549145	-0.690981	-0.547396	-0.761070	-1.670412
1.965	-0.534519	-0.762078	-0.538017	-0.679797	-0.536348	-0.751967	-1.661909
1.970	-0.523502	-0.751538	-0.527064	-0.668688	-0.525414	-0.743018	-1.653392

Continued on next page

Table S3 – Continued from previous page

r(Å)	pCCD(OO)	CCD(OO)	CCSD(OO)	fpCCD(OO)	fpCCSD(OO)	fpLCCD(OO)	fpLCCSD(OO)
1.975	-0.512431	-0.740957	-0.516104	-0.657628	-0.514475	-0.734010	-1.644907
1.980	-0.501440	-0.730476	-0.505266	-0.646624	-0.503577	-0.725019	-1.636450
1.985	-0.490764	-0.720240	-0.494669	-0.635717	-0.492922	-0.716338	-1.627958
1.990	-0.479993	-0.709882	-0.483995	-0.624870	-0.482257	-0.707591	-1.619501
1.995	-0.469192	-0.699454	-0.473291	-0.614092	-0.471583	-0.698788	-1.611068
2.000	-0.458647	-0.689253	-0.462813	-0.603406	-0.461185	-0.690294	-1.602649
2.005	-0.448068	-0.679043	-0.452361	-0.592763	-0.450675	-0.681658	-1.594224
2.010	-0.437634	-0.668938	-0.442030	-0.582208	-0.440353	-0.673208	-1.585834
2.015	-0.427287	-0.658888	-0.431777	-0.571733	-0.430106	-0.664823	-1.577455
2.020	-0.417026	-0.648894	-0.421603	-0.561336	-0.419958	-0.656533	-1.569096
2.025	-0.406850	-0.638954	-0.411508	-0.551019	-0.409925	-0.648322	-1.560728
2.030	-0.396792	-0.629123	-0.401541	-0.540774	-0.399937	-0.640153	-1.552419
2.035	-0.386802	-0.619348	-0.391649	-0.530608	-0.390042	-0.632067	-1.544109
2.040	-0.376928	-0.609633	-0.381835	-0.520538	-0.380262	-0.624087	-1.535858
2.045	-0.367010	-0.599928	-0.372071	-0.510516	-0.370500	-0.616096	-1.527595
2.050	-0.357308	-0.590418	-0.362501	-0.500582	-0.360892	-0.608245	-1.519361
2.055	-0.347718	-0.580936	-0.352984	-0.490749	-0.351375	-0.600510	-1.511160
2.060	-0.338163	-0.571481	-0.343523	-0.480993	-0.341934	-0.592793	-1.502972
2.065	-0.328724	-0.562070	-0.334130	-0.471329	-0.332651	-0.585193	-1.494808
2.070	-0.319275	-0.552771	-0.324867	-0.461721	-0.323284	-0.577542	-1.486699
2.075	-0.310057	-0.543530	-0.315682	-0.452227	-0.314175	-0.570109	-1.478591
2.080	-0.300858	-0.534349	-0.306578	-0.442806	-0.305090	-0.562684	-1.470505
2.085	-0.291766	-0.525302	-0.297626	-0.433459	-0.296084	-0.555308	-1.462469
2.090	-0.282814	-0.516358	-0.288788	-0.424198	-0.287201	-0.548068	-1.454459
2.095	-0.273915	-0.507424	-0.279989	-0.415032	-0.278410	-0.540879	-1.446472
2.100	-0.265111	-0.498569	-0.271288	-0.405952	-0.269716	-0.533768	-1.438512
2.105	-0.256402	-0.489790	-0.262682	-0.396958	-0.261114	-0.526729	-1.430584
2.110	-0.247757	-0.480982	-0.254079	-0.388078	-0.252632	-0.519778	-1.422685
2.115	-0.239297	-0.472468	-0.245758	-0.379238	-0.244233	-0.512901	-1.414831
2.120	-0.230879	-0.463918	-0.237438	-0.370507	-0.235913	-0.506083	-1.406998
2.125	-0.222527	-0.455418	-0.229190	-0.361868	-0.227683	-0.499330	-1.399200
2.130	-0.214247	-0.446912	-0.220970	-0.353336	-0.219529	-0.492631	-1.391423
2.135	-0.206117	-0.438609	-0.212950	-0.344862	-0.211500	-0.486018	-1.383696
2.140	-0.198048	-0.430311	-0.204961	-0.336503	-0.203556	-0.479517	-1.376003
2.145	-0.189994	-0.422097	-0.197076	-0.328210	-0.195682	-0.473013	-1.368332
2.150	-0.182192	-0.414045	-0.189355	-0.320007	-0.187952	-0.466652	-1.360707
2.155	-0.174435	-0.406025	-0.181693	-0.311899	-0.180295	-0.460335	-1.353118
2.160	-0.166764	-0.398074	-0.174117	-0.303882	-0.172737	-0.454093	-1.345549
2.165	-0.159185	-0.390186	-0.166624	-0.295959	-0.165250	-0.447918	-1.338027
2.170	-0.151698	-0.382385	-0.159232	-0.288122	-0.157871	-0.441817	-1.330537
2.175	-0.144274	-0.374717	-0.151984	-0.280351	-0.150576	-0.435747	-1.323116
2.180	-0.136995	-0.367025	-0.144738	-0.272711	-0.143395	-0.429812	-1.315680
2.185	-0.129758	-0.359514	-0.137680	-0.265117	-0.136292	-0.423884	-1.308319
2.190	-0.122663	-0.352007	-0.130645	-0.257642	-0.129285	-0.418063	-1.300978
2.195	-0.115638	-0.344609	-0.123733	-0.250245	-0.122358	-0.412291	-1.293688
2.200	-0.108705	-0.337275	-0.116903	-0.242942	-0.115544	-0.406599	-1.286422
2.205	-0.101867	-0.330016	-0.110164	-0.235730	-0.108829	-0.400979	-1.279202
2.210	-0.095121	-0.322836	-0.103518	-0.228605	-0.102204	-0.395424	-1.272028
2.215	-0.088470	-0.315710	-0.096944	-0.221581	-0.095657	-0.389941	-1.264868
2.220	-0.081839	-0.308681	-0.090483	-0.214627	-0.089183	-0.384482	-1.257783

Continued on next page

Table S3 – Continued from previous page

r(Å)	pCCD(OO)	CCD(OO)	CCSD(OO)	fpCCD(OO)	fpCCSD(OO)	fpLCCD(OO)	fpLCCSD(OO)
2.225	-0.075441	-0.301779	-0.084146	-0.207757	-0.082852	-0.379131	-1.250712
2.230	-0.069060	-0.294917	-0.077873	-0.200982	-0.076580	-0.373830	-1.243689
2.235	-0.062767	-0.288127	-0.071685	-0.194297	-0.070404	-0.368598	-1.236704
2.240	-0.056563	-0.281386	-0.065561	-0.187716	-0.064323	-0.363468	-1.229740
2.245	-0.050424	-0.274742	-0.059550	-0.181205	-0.058332	-0.358350	-1.222855
2.250	-0.044402	-0.268225	-0.053674	-0.174765	-0.052406	-0.353283	-1.215990
2.255	-0.038473	-0.261727	-0.047825	-0.168431	-0.046605	-0.348313	-1.209168
2.260	-0.032628	-0.255317	-0.042080	-0.162188	-0.040899	-0.343425	-1.202390
2.265	-0.026892	-0.249044	-0.036474	-0.156000	-0.035207	-0.338536	-1.195622
2.270	-0.021207	-0.242727	-0.030857	-0.149951	-0.029681	-0.333774	-1.188927
2.275	-0.015612	-0.236581	-0.025415	-0.144036	-0.024179	-0.328978	-1.182270
2.280	-0.010122	-0.230445	-0.019988	-0.138145	-0.018834	-0.324363	-1.175646
2.285	-0.004691	-0.224447	-0.014717	-0.132309	-0.013530	-0.319716	-1.169073
2.290	0.000620	-0.218464	-0.009465	-0.126587	-0.008332	-0.315190	-1.162530
2.295	0.005887	-0.212579	-0.004333	-0.120939	-0.003204	-0.310661	-1.156047
2.300	0.011005	-0.206879	0.000640	-0.115319	0.001831	-0.306239	-1.149591
2.305	0.016092	-0.201058	0.005679	-0.109876	0.006782	-0.301870	-1.143163
2.310	0.021082	-0.195503	0.010480	-0.104421	0.011692	-0.297521	-1.136756
2.315	0.025990	-0.189823	0.015357	-0.099148	0.016439	-0.293276	-1.130459
2.320	0.030797	-0.184354	0.020050	-0.093879	0.021114	-0.289078	-1.124169
2.325	0.035567	-0.178964	0.024641	-0.088702	0.025761	-0.284898	-1.117939
2.330	0.040229	-0.173576	0.029205	-0.083638	0.030301	-0.280784	-1.111731
2.335	0.044781	-0.168295	0.033687	-0.078629	0.034753	-0.276773	-1.105553
2.340	0.049316	-0.163078	0.038070	-0.073706	0.039164	-0.272748	-1.099429
2.345	0.053670	-0.157931	0.042397	-0.068867	0.043446	-0.268860	-1.093314
2.350	0.058073	-0.152816	0.046650	-0.064125	0.047689	-0.264920	-1.087289
2.355	0.062324	-0.147868	0.050763	-0.059408	0.051836	-0.261081	-1.081288
2.360	0.066531	-0.142884	0.054876	-0.054827	0.055867	-0.257310	-1.075341
2.365	0.070619	-0.138090	0.058823	-0.050258	0.059892	-0.253547	-1.069423
2.370	0.074679	-0.133271	0.062766	-0.045812	0.063804	-0.249880	-1.063526
2.375	0.078639	-0.128535	0.066620	-0.041441	0.067633	-0.246245	-1.057702
2.380	0.082502	-0.123868	0.070405	-0.037137	0.071382	-0.242667	-1.051914
2.385	0.086321	-0.119260	0.074116	-0.032924	0.075093	-0.239121	-1.046149
2.390	0.090067	-0.114785	0.077704	-0.028742	0.078699	-0.235637	-1.040444
2.395	0.093661	-0.110375	0.081255	-0.024612	0.082206	-0.232273	-1.034744
2.400	0.097331	-0.105927	0.084756	-0.020652	0.085712	-0.228821	-1.029133
2.405	0.100824	-0.101625	0.088167	-0.016691	0.089079	-0.225520	-1.023547
2.410	0.104250	-0.097395	0.091468	-0.012821	0.092454	-0.222158	-1.017986
2.415	0.107657	-0.093248	0.094718	-0.008983	0.095675	-0.218959	-1.012474
2.420	0.110915	-0.089112	0.097944	-0.005252	0.098852	-0.215787	-1.006990
2.425	0.114191	-0.084963	0.101111	-0.001679	0.102000	-0.212565	-1.001575
2.430	0.117364	-0.081003	0.104142	0.001936	0.105038	-0.209438	-0.996184
2.435	0.120489	-0.077090	0.107132	0.005479	0.108038	-0.206377	-0.990824
2.440	0.123494	-0.073278	0.110015	0.008971	0.110948	-0.203355	-0.985525
2.445	0.126489	-0.069439	0.112901	0.012335	0.113813	-0.200368	-0.980249
2.450	0.129391	-0.065669	0.115714	0.015638	0.116563	-0.197441	-0.975025
2.455	0.132234	-0.062004	0.118430	0.018905	0.119304	-0.194523	-0.969821
2.460	0.135018	-0.058401	0.121082	0.022111	0.121963	-0.191650	-0.964671
2.465	0.137724	-0.054825	0.123698	0.025229	0.124546	-0.188857	-0.959558
2.470	0.140376	-0.051342	0.126224	0.028304	0.127090	-0.186066	-0.954480

Continued on next page

Table S3 – Continued from previous page

r(Å)	pCCD(OO)	CCD(OO)	CCSD(OO)	fpCCD(OO)	fpCCSD(OO)	fpLCCD(OO)	fpLCCSD(OO)
2.475	0.142971	-0.047912	0.128694	0.031314	0.129553	-0.183322	-0.949444
2.480	0.145497	-0.044531	0.131111	0.034255	0.131965	-0.180628	-0.944438
2.485	0.147961	-0.041207	0.133468	0.037133	0.134310	-0.177977	-0.939486
2.490	0.150359	-0.037934	0.135772	0.039946	0.136586	-0.175367	-0.934565
2.495	0.152721	-0.034748	0.137995	0.042723	0.138815	-0.172788	-0.929687
2.500	0.155008	-0.031586	0.140187	0.045419	0.140987	-0.170256	-0.924833
2.505	0.157255	-0.028463	0.142331	0.048048	0.143104	-0.167726	-0.920025
2.510	0.159394	-0.025428	0.144404	0.050631	0.145152	-0.165312	-0.915236
2.515	0.161512	-0.022492	0.146376	0.053195	0.147147	-0.162866	-0.910518
2.520	0.163565	-0.019537	0.148346	0.055639	0.149101	-0.160469	-0.905807
2.525	0.165588	-0.016697	0.150221	0.058084	0.150985	-0.158109	-0.901155
2.530	0.167543	-0.013871	0.152066	0.060435	0.152814	-0.155783	-0.896543
2.535	0.169446	-0.011109	0.153853	0.062744	0.154600	-0.153492	-0.891951
2.540	0.171296	-0.008395	0.155590	0.064995	0.156329	-0.151238	-0.887402
2.545	0.173096	-0.005733	0.157278	0.067197	0.158004	-0.149021	-0.882889
2.550	0.174843	-0.003122	0.158914	0.069345	0.159629	-0.146836	-0.878413
2.555	0.176538	-0.000562	0.160495	0.071436	0.161202	-0.144680	-0.873975
2.560	0.178177	0.001952	0.162022	0.073461	0.162722	-0.142548	-0.869580
2.565	0.179793	0.004430	0.163507	0.075420	0.164189	-0.140454	-0.865216
2.570	0.181338	0.006856	0.164953	0.077345	0.165613	-0.138394	-0.860881
2.575	0.182835	0.009220	0.166339	0.079234	0.166992	-0.136365	-0.856584
2.580	0.184286	0.011529	0.167674	0.081078	0.168319	-0.134371	-0.852334
2.585	0.185690	0.013790	0.168961	0.082874	0.169599	-0.132408	-0.848105
2.590	0.187048	0.016008	0.170205	0.084619	0.170840	-0.130470	-0.843914
2.595	0.188358	0.018177	0.171402	0.086319	0.172035	-0.128560	-0.839757
2.600	0.189624	0.020304	0.172556	0.087969	0.173183	-0.126680	-0.835635
2.605	0.190854	0.022386	0.173670	0.089576	0.174290	-0.124831	-0.831541
2.610	0.192033	0.024423	0.174736	0.091134	0.175351	-0.123009	-0.827487
2.615	0.193164	0.026412	0.175755	0.092648	0.176370	-0.121213	-0.823467
2.620	0.194256	0.028353	0.176735	0.094133	0.177353	-0.119448	-0.819474
2.625	0.195299	0.030236	0.177668	0.095597	0.178289	-0.117728	-0.815521
2.630	0.196351	0.032112	0.178592	0.096995	0.179190	-0.116025	-0.811578
2.635	0.197332	0.033914	0.179439	0.098360	0.180070	-0.114329	-0.807695
2.640	0.198219	0.035721	0.180286	0.099641	0.180850	-0.112686	-0.803849
2.645	0.199086	0.037423	0.181023	0.100896	0.181618	-0.111058	-0.800032
2.650	0.200007	0.039132	0.181794	0.102154	0.182363	-0.109460	-0.796238
2.655	0.200786	0.040818	0.182515	0.103296	0.183080	-0.107871	-0.792463
2.660	0.201598	0.042419	0.183181	0.104478	0.183747	-0.106319	-0.788739
2.665	0.202348	0.044014	0.183838	0.105593	0.184361	-0.104812	-0.785044
2.670	0.203033	0.045545	0.184420	0.106654	0.184956	-0.103288	-0.781391
2.675	0.203695	0.047049	0.184971	0.107661	0.185501	-0.101806	-0.777759
2.680	0.204361	0.048495	0.185491	0.108694	0.186039	-0.100349	-0.774153
2.685	0.204949	0.049920	0.186003	0.109614	0.186504	-0.098926	-0.770586
2.690	0.205533	0.051274	0.186457	0.110614	0.186995	-0.097506	-0.767027
2.695	0.206041	0.052662	0.186921	0.111493	0.187403	-0.096147	-0.763524
2.700	0.206552	0.053975	0.187296	0.112327	0.187794	-0.094770	-0.760055
2.705	0.207012	0.055269	0.187664	0.113133	0.188131	-0.093428	-0.756605
2.710	0.207477	0.056484	0.187967	0.113953	0.188492	-0.092072	-0.753172
2.715	0.207915	0.057733	0.188297	0.114683	0.188756	-0.090814	-0.749804
2.720	0.208241	0.058911	0.188567	0.115413	0.189020	-0.089528	-0.746440

Continued on next page

Table S3 – Continued from previous page

r(Å)	pCCD(OO)	CCD(OO)	CCSD(OO)	fpCCD(OO)	fpCCSD(OO)	fpLCCD(OO)	fpLCCSD(OO)
2.725	0.208579	0.060054	0.188796	0.116097	0.189258	-0.088262	-0.743128
2.730	0.208938	0.061164	0.189000	0.116762	0.189468	-0.087017	-0.739822
2.735	0.209200	0.062253	0.189197	0.117406	0.189645	-0.085811	-0.736549
2.740	0.209489	0.063295	0.189341	0.118012	0.189800	-0.084607	-0.733308
2.745	0.209735	0.064321	0.189474	0.118581	0.189923	-0.083430	-0.730094
2.750	0.209945	0.065311	0.189575	0.119127	0.190020	-0.082273	-0.726909
2.755	0.210128	0.066271	0.189651	0.119644	0.190090	-0.081135	-0.723752
2.760	0.210285	0.067202	0.189700	0.120133	0.190135	-0.080016	-0.720624
2.765	0.210416	0.068103	0.189724	0.120595	0.190154	-0.078915	-0.717523
2.770	0.210508	0.068961	0.189707	0.121028	0.190155	-0.077819	-0.714448
2.775	0.210619	0.069826	0.189707	0.121445	0.190119	-0.076779	-0.711393
2.780	0.210686	0.070630	0.189650	0.121843	0.190081	-0.075724	-0.708381
2.785	0.210700	0.071425	0.189581	0.122187	0.189991	-0.074697	-0.705392
2.790	0.210723	0.072186	0.189501	0.122553	0.189905	-0.073690	-0.702414
2.795	0.210662	0.072909	0.189354	0.122818	0.189766	-0.072680	-0.699473
2.800	0.210670	0.073621	0.189253	0.123187	0.189665	-0.071716	-0.696564
2.805	0.210614	0.074304	0.189091	0.123447	0.189500	-0.070755	-0.693679
2.810	0.210540	0.074962	0.188909	0.123687	0.189315	-0.069811	-0.690817
2.815	0.210445	0.075595	0.188707	0.123904	0.189109	-0.068883	-0.687982
2.820	0.210329	0.076204	0.188485	0.124099	0.188883	-0.067972	-0.685173
2.825	0.210193	0.076788	0.188242	0.124271	0.188638	-0.067075	-0.682388
2.830	0.210038	0.077348	0.187980	0.124423	0.188373	-0.066194	-0.679629
2.835	0.209856	0.077913	0.187701	0.124482	0.188067	-0.065332	-0.676920
2.840	0.209643	0.078411	0.187392	0.124613	0.187766	-0.064488	-0.674206
2.845	0.209467	0.078887	0.187078	0.124758	0.187460	-0.063666	-0.671520
2.850	0.209216	0.079379	0.186744	0.124772	0.187093	-0.062833	-0.668863
2.855	0.208970	0.079835	0.186401	0.124817	0.186744	-0.062029	-0.666234
2.860	0.208704	0.080265	0.186039	0.124856	0.186377	-0.061238	-0.663608
2.865	0.208410	0.080681	0.185661	0.124854	0.185981	-0.060464	-0.661042
2.870	0.208130	0.081057	0.185255	0.124867	0.185578	-0.059700	-0.658470
2.875	0.207813	0.081383	0.184815	0.124906	0.185192	-0.058916	-0.655932
2.880	0.207460	0.081755	0.184381	0.124777	0.184713	-0.058197	-0.653451
2.885	0.207182	0.082058	0.183933	0.124794	0.184275	-0.057481	-0.650950
2.890	0.206767	0.082377	0.183468	0.124673	0.183812	-0.056758	-0.648479
2.895	0.206392	0.082682	0.183009	0.124574	0.183330	-0.056064	-0.646048
2.900	0.206024	0.082966	0.182549	0.124502	0.182839	-0.055403	-0.643604
2.905	0.205614	0.083199	0.182033	0.124395	0.182347	-0.054707	-0.641211
2.910	0.205194	0.083428	0.181510	0.124257	0.181831	-0.054041	-0.638849
2.915	0.204831	0.083645	0.180978	0.124113	0.181298	-0.053405	-0.636499
2.920	0.204344	0.083849	0.180430	0.123911	0.180739	-0.052764	-0.634158
2.925	0.203871	0.084014	0.179862	0.123750	0.180195	-0.052125	-0.631844
2.930	0.203426	0.084201	0.179317	0.123546	0.179615	-0.051529	-0.629565
2.935	0.202957	0.084353	0.178747	0.123352	0.179041	-0.050926	-0.627293
2.940	0.202451	0.084495	0.178162	0.123109	0.178446	-0.050342	-0.625062
2.945	0.201998	0.084661	0.177645	0.122942	0.177910	-0.049759	-0.622789
2.950	0.201492	0.084732	0.177001	0.122714	0.177265	-0.049193	-0.620626
2.955	0.200969	0.084856	0.176417	0.122435	0.176641	-0.048651	-0.618451
2.960	0.200500	0.084888	0.175761	0.122243	0.176036	-0.048093	-0.616256
2.965	0.199881	0.084948	0.175083	0.121852	0.175360	-0.047535	-0.614147
2.970	0.199342	0.084995	0.174444	0.121583	0.174706	-0.047019	-0.612033

Continued on next page

Table S3 – Continued from previous page

$r(\text{\AA})$	pCCD(OO)	CCD(OO)	CCSD(OO)	fpCCD(OO)	fpCCSD(OO)	fpLCCD(OO)	fpLCCSD(OO)
2.975	0.198763	0.085008	0.173765	0.121282	0.174056	-0.046484	-0.609927
2.980	0.198207	0.085009	0.173084	0.120985	0.173379	-0.045989	-0.607853
2.985	0.197615	0.085058	0.172454	0.120642	0.172723	-0.045480	-0.605801
2.990	0.197092	0.085023	0.171763	0.120384	0.172056	-0.045002	-0.603736
2.995	0.196455	0.085037	0.171092	0.119979	0.171348	-0.044518	-0.601734
3.000	0.195878	0.085009	0.170396	0.119623	0.170633	-0.044065	-0.599746

Frozen density embedding with pCCD electron densities

Rahul Chakraborty¹ and Paweł Tecmer¹

Institute of Physics, Faculty of Physics, Astronomy, and Informatics, Nicolaus Copernicus University in Toruń, Grudziądzka 5, 87-100 Toruń, Poland.

(*Electronic mail: rahuldev.chakraborty91@gmail.com)

(Dated: 17 April 2026)

The pair-coupled-cluster doubles (pCCD) method has emerged as a viable approach for quantum-chemical studies of strongly correlated systems. Despite its lower formal scaling ($\mathcal{O}(N^4)$) compared to other versions of coupled cluster (CC) theory, applications to large chemical structures are still expensive. Fragmentation and embedding strategies offer a viable approach in such cases. In this work, we present a simple and efficient density-embedding scheme based on pCCD electron densities. The main computational benefit arises from the fact that pCCD response Λ -equations are much cheaper to compute than those of standard CC methods, providing easy access to one-electron properties. The pCCD densities of the individual subsystems are used to generate static embedding potentials that capture the environment's effect on the embedded system. The individual fragment energies are then iteratively converged in a self-consistent fashion. We demonstrate the reliable performance of this scheme with the estimation of dipole moments of the weakly bound $\text{CO}_2 \cdots \text{Rg}$ [Rg = He, Ne, Ar, and Kr] complexes and with the modeling vertical excitations of some microsolvated molecules.

I. INTRODUCTION

Quantum chemical studies of molecules and complexes mainly rely on an accurate description of the interplay of interactions between atoms or molecular fragments. However, even the most efficient electronic structure algorithms available today cannot handle large structures due to computational demands. In the condensed phase, much of the chemistry occurs in a small part of the structure (the so-called active region or active subsystem), and the environment plays a key role in the energetics through its interactions with the active part. Keeping this in mind, the problem's complexity can be reduced by using a fragmentation approach.^{1–4} Quantum embedding theories have made significant strides in this direction. One class of such embedding theories is where both active and environment fragments are treated with the same or different quantum-mechanical (QM) theories, i.e., QM/QM embedding.⁵ In this aspect, Cortona introduced the frozen density embedding (FDE) scheme for solid states,⁶ and Wośowski and Warshel for solvated molecules.⁷ In its original formulation, FDE is a Density Functional Theory (DFT)-based approach⁸ where the total density of the system is partitioned into smaller, coupled subsystems,

$$\rho_{\text{tot}} = \rho_I + \rho_{II}. \quad (1)$$

Here ρ_{tot} is the supramolecular density; ρ_I is the density of the active fragment/subsystem of our interest where the chemistry is happening, and ρ_{II} represents the density of the 'frozen' environment. The supramolecular energy can then be expressed as a sum of fragment energies plus an additional coupling term— $E_{\text{int}}[\rho_I, \rho_{II}]$,

$$E_{\text{tot}}[\rho_I + \rho_{II}] = E_I[\rho_I] + E_{II}[\rho_{II}] + E_{\text{int}}[\rho_I, \rho_{II}]. \quad (2)$$

The individual subsystem energy can be expressed as

$$E_i[\rho_i] = \int \rho_i(\mathbf{r}) v_{\text{nuc}}^i(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \iint \frac{\rho_i(\mathbf{r}) \rho_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + V_{\text{nuc,nuc}}^i + E_{\text{xc}}[\rho_i] + T_s[\rho_i] \quad (3)$$

where v_{nuc}^i in the first term is the nuclear charge, the second term denotes the Coulomb potential, and $V_{\text{nuc,nuc}}^i$ is the nuclear repulsion energy of subsystem i ($i = I, II$). In Eq. (3), E_{xc} and T_s denote the exchange–correlation (xc) energy and kinetic energy of the individual subsystem(s), respectively. The $E_{\text{int}}[\rho_I, \rho_{II}]$ coupling term in Eq. (2) represents the interaction between the subsystems I and II, and depends on their Kohn–Sham electron densities and nuclear charges,

$$E_{\text{int}}[\rho_I, \rho_{II}] = \int \rho_I(\mathbf{r}) v_{\text{nuc}}^{II}(\mathbf{r}) d\mathbf{r} + \int \rho_{II}(\mathbf{r}) v_{\text{nuc}}^I(\mathbf{r}) d\mathbf{r} + \iint \frac{\rho_I(\mathbf{r}_1) \rho_{II}(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 + V_{\text{nuc,nuc}}^{I-II} + E_{\text{xc}}^{\text{nadd}}[\rho_I, \rho_{II}] + T_s^{\text{nadd}}[\rho_I, \rho_{II}]. \quad (4)$$

The non-additive energy contributions coming from the xc and kinetic terms are defined as

$$E_{\text{xc}}^{\text{nadd}}[\rho_I, \rho_{II}] = E_{\text{xc}}[\rho_I + \rho_{II}] - E_{\text{xc}}[\rho_I] - E_{\text{xc}}[\rho_{II}] \quad (5)$$

and

$$T_s^{\text{nadd}}[\rho_I, \rho_{II}] = T_s[\rho_I + \rho_{II}] - T_s[\rho_I] - T_s[\rho_{II}], \quad (6)$$

respectively. The quantities in Eqs. (5) and (6) are estimated using approximate density functionals. The computational advantage of expressing the energy of the whole system according to Eq. (2) arises from the fact that only one subsystem is optimized self-consistently at a time, keeping the electron density of the other subsystem 'frozen'. The derivative of the interaction energy functional that describes the

inter-system dependence is the embedding potential, which can be decomposed into Coulomb interactions with the environment (nuclei and frozen electron density) plus derivatives of the non-additive parts of the xc and kinetic energy functionals. The performance of the FDE approach heavily relies on the non-additive xc and kinetic terms in the employed density functionals.^{9,10} Numerical results show that the latter term is the more dominating one, especially for strongly interacting (e.g., via covalent bonds) subsystems, and acts as a major source of error in FDE calculations.^{11–13} Substantial work on excitation energy and response property calculations using time-dependent density functional theorem (TD-DFT) within the FDE framework has been done by different groups, which have expanded the applicability of FDE.^{14–16} Projection-based embedding techniques^{17–20} and their Huzinaga equation-based versions^{21–23} constitute a path of development parallel to density embedding methods.

Due to the flexibility in the nature of the density used in FDE, wave-function theory (WFT)-based approaches can be easily integrated into this framework. The WFT-in-DFT version of the embedding approach employs various WFT methods for the ‘active’ subsystem while the environment is treated at the DFT level. The formal correctness of such an approach, for exact density functional along with full configuration interaction (FCI), was shown by Wesolowski.²⁴ Carter and coworkers have done extensive work using variational WFT methods, such as configuration interaction,²⁵ complete active space self-consistent field (CASSCF),²⁶ to study localized excitations in solids or surfaces. Coupled cluster (CC) methods have also been used in this context. Höfener et al. implemented a CC-in-DFT method based on the response formalism.^{27,28} A simplified approach was explored by Gomes and coworkers, in which a DFT-in-DFT FDE calculation is performed to obtain converged subsystem densities and the resulting embedding potential, which is then used to represent the environment in the WFT-in-DFT calculation.^{13,29,30} Recent works have also explored WFT-in-WFT frozen density embedding, thereby alleviating dependence on DFT and its associated problems.^{31,32}

pCCD-based methods have attracted significant attention as efficient and viable WFT computational models in recent years. Though tracing its origin to geminal theories,³³ pCCD can also be considered as a simplification of the CC ansatz,³⁴ which includes only paired excitations from the same spatial orbitals. pCCD has demonstrated success in treating complex electronic structures at a mean-field computational cost.^{35–40} The method has been utilized to model electronic structures of a range of chemical species, such as actinides,⁴¹ organic molecules,^{42,43} studying catalytic action of molybdenum co-factor,⁴⁴ etc. A plethora of recent works have used pCCD to model excited states.^{43,45–48} However, simulating larger structures, especially in condensed phases, remains a challenge. This is further corroborated by the need for post-pCCD dynamic electron-correlation corrections to achieve chemical accuracy, which steeply increases the computational demand. This motivates the development of pCCD-based embedding techniques.

In our previous work, we implemented a pCCD-in-DFT ap-

proach,⁴⁹ where the static embedding potential was generated from DFT-in-DFT calculations in external software. In this work, we focus on the implementation and analysis of the WFT-in-WFT computational scheme based solely on electron densities obtained from pCCD. The key advantage over conventional CC-in-CC embedding approaches is that the pCCD one-particle density matrix can be obtained straightforwardly by solving the associated Λ -equations, which scale as $\mathcal{O}(N^4)$. This results in a simple, fully flexible, and computationally efficient pCCD-based embedding framework that introduces virtually no additional cost beyond the underlying pCCD calculation. The performance of the method is assessed on two test cases: (i) the dipole moments of weakly bound $\text{CO}_2 \cdots \text{Rg}$ [$\text{Rg} = \text{He}, \text{Ne}, \text{Ar}, \text{and Kr}$] complexes and (ii) the vertical excitation energies of selected microsolvated systems.

This work is structured as follows: Section II offers a concise overview of the theoretical models under investigation, followed by a detailed description of the computational methods in Section III. Section IV summarizes the numerical results. Concluding remarks are presented in Section V.

II. THEORY

A. Ground and excited states with pCCD-based methods

The pCCD ansatz was inspired by the concept of electron-pair wave functions (geminals³³) and originally developed as the antisymmetric product of a one-reference geminal model. It can be expressed as a simple form of the CC ansatz³⁴

$$\Psi_{\text{pCCD}} = e^{\hat{T}_{\text{p}}} |\Phi_0\rangle, \quad (7)$$

where

$$\hat{T}_{\text{p}} = \sum_i^{n_{\text{occ}}} \sum_a^{n_{\text{virt}}} t_{ii}^{aa} \hat{a}_a^\dagger \hat{a}_i^\dagger \hat{a}_i \hat{a}_a \quad (8)$$

is a cluster operator that excites electron pairs from the same spatial orbital. Here $\hat{a}_p^\dagger (\hat{a}_{\bar{p}}^\dagger)$ are the electron creation and $\hat{a}_p (\hat{a}_{\bar{p}})$ are the electron annihilation operators for the alpha (p) and beta ($\bar{p}$) electrons, respectively. t_{ii}^{aa} are the pCCD cluster amplitudes, and $|\Phi_0\rangle$ is the reference wave function. When augmented with a variational orbital optimization (oo) protocol, pCCD becomes size-consistent.^{35,50} Several studies have shown that oo-pCCD provides a reliable description of quasi-degenerate states and allows for the modeling of potential energy surfaces.^{35,38,51–54}

However, a major drawback of this model is the lack of description of dynamic electron correlation, as pCCD specifically focuses on the seniority-zero part of the full electronic Hamiltonian. A remedy for this is to use the tailored CC (tCC).^{55,56} In tCC, the excitation space is partitioned^{57–59} using the cluster operators $\hat{T}_{\text{int}}$ and $\hat{T}_{\text{ext}}$.

$$|\Psi_{\text{tCC}}\rangle = e^{\hat{T}_{\text{ext}}} e^{\hat{T}_{\text{int}}} |\Phi_0\rangle. \quad (9)$$

First, $\hat{T}_{\text{int}}$, describing the strong correlation effect, is applied to the reference wave function $|\Phi_0\rangle$. Subsequently, $\hat{T}_{\text{ext}}$ is applied to this intermediate wave function, with the t_{int} amplitudes kept constant to introduce additional excitations and incorporate dynamic correlation in the final wave function. In a similar fashion, the dynamic correlation excluded in the pCCD/oo-pCCD wave function can be accounted for by *a posteriori* corrections, where single and unpaired double excitations are added to the wave function. Such an approach is termed as fpCC, ⁶⁰ which expands the pCCD reference wave function by adding unpaired doubles, along with single excitations (fpCCSD) to $|\Psi_{\text{pCCD}}\rangle$, ⁶¹

$$|\Psi_{\text{fpCCSD}}\rangle = e^{\hat{T}_{\text{ext}}} e^{\hat{T}_{\text{p}}} |\Phi_0\rangle. \quad (10)$$

In Eq. (10), $\hat{T}_{\text{ext}}$ is the cluster operator which acts upon the pCCD wave function and generates non-pair doubles (also known as broken pairs) and singles excitation. The amplitudes of $\hat{T}_{\text{p}}$, which acts as $\hat{T}_{\text{int}}$ in this case, are kept constant (and hence the name ‘frozen pair’). Specifically, in the spin-free CC formulation, where all excitation operators are written in terms of singlet excitation operators $\hat{E}_{ai}$, $\hat{T}_{\text{ext}}$ reads as

$$\begin{aligned} \hat{T}_{\text{ext}} = \hat{T}_1 + (\hat{T}_2 - \hat{T}_{\text{p}}) = & \sum_i \sum_a^{n_{\text{occ}}} \sum_{\text{virt}} t_i^a \hat{E}_{ai} + \frac{1}{2} \sum_{i,j}^{n_{\text{occ}}} \sum_{a,b}^{n_{\text{virt}}} t_{ij}^{ab} \hat{E}_{ai} \hat{E}_{bj} \\ & - \frac{1}{2} \sum_i^{n_{\text{occ}}} \sum_a^{n_{\text{virt}}} t_{ii}^{aa} \hat{E}_{ai} \hat{E}_{ai}, \end{aligned} \quad (11)$$

where

$$\hat{E}_{ai} = \hat{a}_a^\dagger \hat{a}_i + \hat{a}_a^\dagger \hat{a}_{\bar{i}} \quad (12)$$

and $\hat{E}_{bj}$ is defined analogously. One simplified form of fpCCSD is the frozen pair linearized coupled cluster singles and doubles (fpLCCSD) ansatz, ^{61,62}

$$|\Psi_{\text{fpLCCSD}}\rangle \approx (1 + \hat{T}_{\text{ext}}) |\Psi_{\text{pCCD}}\rangle. \quad (13)$$

fpLCC can be understood as a linearized pCCD-tCC approach, where only the linear terms involving $\hat{T}_{\text{ext}}$ are included in the amplitude equations. In that case, the energy is calculated by terminating the Baker–Campbell–Hausdorff expansion at the second term in $\hat{T}_{\text{ext}}$,

$$(\hat{H}_N + [\hat{H}_N, \hat{T}_{\text{ext}}]) |\Psi_{\text{pCCD}}\rangle = E |\Psi_{\text{pCCD}}\rangle. \quad (14)$$

Here, $\hat{H}_N = \hat{H} - \langle \Phi_0 | \hat{H} | \Phi_0 \rangle$ is the normal ordered form of the electronic Hamiltonian $\hat{H}$. This ansatz can be understood as a linearized, pCCD-tailored CC approach, in which the full frozen-pair equations (including all nonlinear and disconnected terms involving $\hat{T}_{\text{p}}$ vertices) are included in the LCCSD equations. Here we would like to point out to the readers that fpLCCSD was also denoted previously as pCCD-LCCSD in the literature. ⁶²

Excited states can be modeled with pCCD using the equation-of-motion (EOM) formalism. ^{63–65} The simplest excited state extension for pCCD is the EOM-pCCD model, where the cluster operator $\hat{T}$ and the CI-type operator $\hat{R}$ are restricted to pair-double excitations only,

$$\hat{R}_{\text{p}} = c_0 \hat{t}_0 + \sum_{ia} c_{ii}^{aa} \hat{t}_{ii}^{aa}. \quad (15)$$

Here, $\hat{t}_{ii}^{aa} = \frac{1}{2} \hat{E}_{ai} \hat{E}_{ai}$ denotes pair excitations, and $\hat{t}_0$ is the identity operator which describes the contribution of the ground state in the particular excited state. The corresponding EOM-pCCD equation for a given excited state k can be expressed as

$$[\hat{\mathcal{H}}_N^{(\text{p})}, \hat{R}_{\text{p}}] |\Phi_0\rangle = \omega_k \hat{R}_{\text{p}} |\Phi_0\rangle, \quad (16)$$

where

$$\begin{aligned} \hat{\mathcal{H}}_N^{(\text{p})} &= e^{-\hat{T}_{\text{p}}} \hat{H}_N e^{\hat{T}_{\text{p}}} \\ &= \hat{H}_N + [\hat{H}_N, \hat{T}_{\text{p}}] + \frac{1}{2} [[\hat{H}_N, \hat{T}_{\text{p}}], \hat{T}_{\text{p}}] \end{aligned} \quad (17)$$

is the similarity-transformed pCCD Hamiltonian, and ω_k is the excitation energy for the particular state. For accessing general (non-pair) doubly excited and singly excited states, $\hat{T}_2'$ and $\hat{T}_1$ are included in the cluster operator. ⁶⁶ In that case, the ground state wave function is fpLCCSD, and the approach is called EOM-fpLCCSD (also denoted as EOM-pCCD-LCCSD in previous articles). ⁶⁷ $\hat{R}$, including the singles and doubles contributions, bears the form

$$\hat{R}_{\text{sd}} = c_0 \hat{t}_0 + \sum_{ia} c_i^a \hat{t}_i^a + \frac{1}{2} \sum_{ia} \sum_{jb} c_{ij}^{ab} \hat{t}_{ij}^{ab}. \quad (18)$$

Here $\hat{t}_i^a = \hat{E}_{ai}$ is the (spin-free) single excitation operator and $\hat{t}_{ij}^{ab} = \hat{E}_{ai} \hat{E}_{bj}$ is the (spin-free) doubles excitation operator. In this case, the EOM-fpLCCSD equation, again for a given k -th excited state, becomes

$$[\hat{\mathcal{H}}_N^{(\text{sd})}, \hat{R}_{\text{sd}}] |\Phi_0\rangle = \omega_k \hat{R}_{\text{sd}} |\Phi_0\rangle. \quad (19)$$

It is important to note here that the similarity transformed Hamiltonian $\hat{\mathcal{H}}_N^{(\text{sd})}$ has different expressions for pair and non-pair excitations. For non-pair excitations, it has the following form

$$\begin{aligned} \hat{\mathcal{H}}_N^{(\text{np})} &= e^{-(\hat{T}_{\text{p}} - \hat{T}_{\text{np}})} \hat{H}_N e^{(\hat{T}_{\text{p}} + \hat{T}_{\text{np}})} \\ &\approx \hat{H}_N + [\hat{H}_N, \hat{T}_{\text{np}}] + [[\hat{H}_N, \hat{T}_{\text{np}}], \hat{T}_{\text{p}}] \\ &\quad + [\hat{H}_N, \hat{T}_{\text{p}}] + \frac{1}{2} [[\hat{H}_N, \hat{T}_{\text{p}}], \hat{T}_{\text{p}}], \end{aligned} \quad (20)$$

because of the exclusion of terms non-linear in $\hat{T}_{\text{np}}$ in the linearized cluster operator, while the pair excitations are still obtained with Eq. (17).

B. FDE with pCCD

In our pCCD-in-pCCD FDE scheme, we obtain subsystem electron densities from the pCCD response one-particle reduced density matrices (1-RDMs)^{37,68} of each subsystem defined as

$$\text{pCCD}\gamma_q^p = \langle \Psi | (1 + \hat{\Lambda}_p) e^{-\hat{T}_p} \hat{a}_p^\dagger \hat{a}_q e^{\hat{T}_p} | \Psi \rangle, \quad (21)$$

where $\hat{\Lambda}_p = \sum_{ia} \lambda_{ia}^{ii} \hat{a}_i^\dagger \hat{a}_i^\dagger \hat{a}_a \hat{a}_a$ is the electron-pair de-excitation operator. It is to be noted that pCCD 1-RDM is diagonal and depicts the occupation numbers in the natural orbital basis. The 1-RDM is already obtained during the pCCD orbital-optimization step and does not involve any additional computation.

The derivative of the coupling energy term defined in Eq. (4) with respect to the subsystem density gives the embedding potential

$$v_{\text{emb}}^I(\mathbf{r}) = v_{\text{nuc}}^I(\mathbf{r}) + \int \frac{\rho_{II}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{\text{xc}}^{\text{nadd}}[\rho_I, \rho_{II}]}{\delta \rho(\mathbf{r})} \bigg|_{\rho_I} + \frac{\delta T_s^{\text{nadd}}[\rho_I, \rho_{II}]}{\delta \rho(\mathbf{r})} \bigg|_{\rho_I}, \quad (22)$$

where ρ_i terms denote the electron density of the i -th subsystem obtained from the corresponding pCCD calculation. There is no exact method to describe the non-additive terms other than using approximate functionals. In this work, we used a simple Thomas-Fermi model,^{69,70} which corresponds to the local density approximation (LDA) in DFT, to describe the non-additive kinetic potential

$$v_{\text{kin}}^{\text{nadd}}(\mathbf{r}) = \frac{\delta T_{\text{TF}}^{\text{nadd}}[\rho_I, \rho_{II}]}{\delta \rho_I(\mathbf{r})} = \frac{1}{2} (3\pi^2)^{2/3} \times [(\rho_I + \rho_{II})^{2/3} - \rho_I^{2/3}]. \quad (23)$$

The non-additive xc term in the embedding potential of Eq. (22) is estimated using the Slater $X\alpha$ potential⁷¹

$$v_{\text{xc}}^{\text{nadd}}(\mathbf{r}) = \frac{\delta E_{X\alpha}^{\text{nadd}}[\rho_I, \rho_{II}]}{\delta \rho_I(\mathbf{r})} = -\alpha \frac{3}{2} \left(\frac{3}{\pi} \right)^{1/3} \times [(\rho_I + \rho_{II})^{1/3} - \rho_I^{1/3}]. \quad (24)$$

The value of the Slater exchange parameter α is actually atom-dependent and also varies for the property under question. For example, Zope et al. estimated the optimal α values for atomization energies of many different atoms.⁷² For simplicity, in this work, we used $\alpha=0.66$ as originally proposed by Kohn and Sham.⁷³ The LDA approximation for the non-additive kinetic potential term has been used for weakly interacting systems and has given reliable results; however, for strongly overlapping subsystems, LDA performs rather

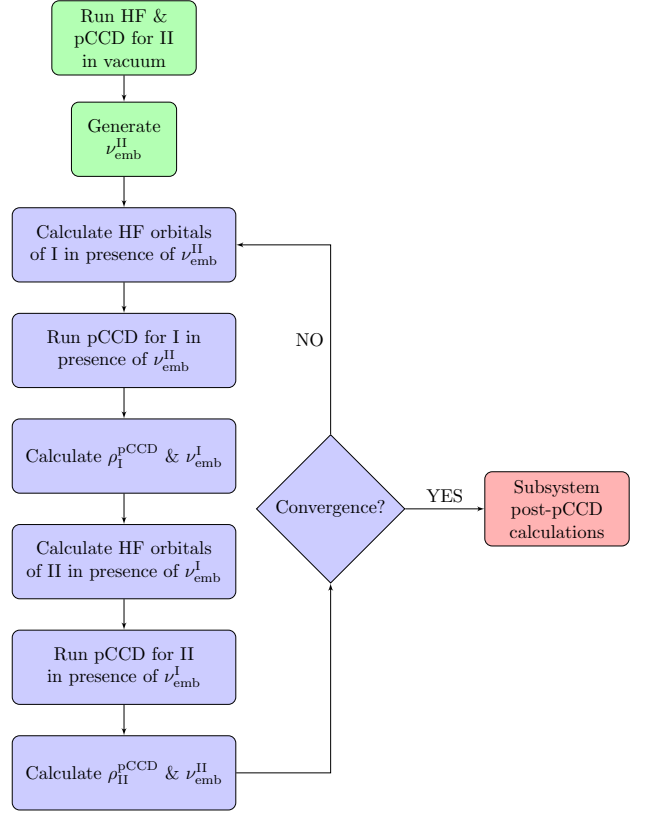


FIG. 1: Depiction of FDE freeze-and-thaw cycles with pCCD densities. The green blocks show steps performed with the isolated environment subsystem. The red block shows post-pCCD calculations (fpCCSD, fpLCCSD) with the converged pCCD wave function of each subsystem. EOM-fpLCCSD calculations are also done at this stage after fpLCCSD.

poorly.⁷⁴ The Slater exchange potential, on the other hand, does not show proper long-range Coulomb decay. And of course, using only the Slater exchange potential in the non-additive xc term means the correlation part is missing from the embedding potential. These simple models definitely leave room for improvement in the pCCD-based FDE framework.

An important aspect to note here is that a completely frozen ρ_{II} can not represent the polarization effect of the environment on the active subsystem and vice versa. This is solved by employing the so-called ‘freeze-and-thaw’ cycles.⁷⁵ Here, one subsystem density is calculated at a time, keeping the other subsystem frozen. Thereafter, the roles of the subsystems are reversed. The previously frozen subsystem is now thawed, and its density is calculated by freezing the other subsystem’s density (obtained in the previous step). This is done till self-consistency. Figure 1 describes the freeze-and-thaw workflow followed in this work. At the beginning, the so-called environment subsystem (marked as II) is considered in isolation, i.e., in the absence of the other subsystem (but at the supramolecular geometry). HF and subsequent pCCD calculations are run

for this fragment. This works as a guess density to initiate the freeze-and-thaw cycles. The embedding potential representing the isolated environment subsystem is then evaluated on a numerical grid. From the next cycle, the one-electron matrix elements of this potential are calculated as

$$v_{ij}^I = \langle \phi_i | v_{\text{emb}}^I | \phi_j \rangle \approx \sum_k w_k v_{\text{emb}}^I(r_k) \phi_i(r_k) \phi_j(r_k), \quad (25)$$

where $\phi_i(r_k)$ is the numerical value of the orbital ϕ_i at grid point r_k and w_k is the integration weight of this point. This term is then plugged into the Fock matrix of the current active subsystem (I) at the HF level, and the SCF iterations are run. This brings, to a certain extent, the relaxation effect of the orbitals of subsystem I due to the presence of subsystem II. The resultant perturbed canonical MOs are used for pCCD calculations in the presence of the static embedding potential, which represents the environment/frozen subsystem (II). From the updated pCCD electron density of subsystem I obtained at this stage, its embedding potential is generated again. After that, the roles of the active and environment subsystems are swapped, and the same steps are repeated. This completes one freeze-and-thaw macroiteration. These steps are followed until convergence of the pCCD energy of each subsystem.

Post-convergence, fpCCSD/fpLCCSD calculations are run for each fragment using its converged pCCD wave function. In the case of excitation energy, EOM-fpLCCSD calculations are performed after the fpLCCSD step to get the vertical excitation energies. A limitation of any static embedding approach, such as ours, is the neglect of the change in electron densities in response to electronic excitations. Consequently, we always consider a fixed embedding potential based on the ground-state subsystem density. This can only produce reliable results as long as the non-additive kinetic and xc potentials are not very different between the ground and excited states. For the weakly bound systems investigated in this work, we assume this to be true. The effect of electronic excitations in different subsystems can be incorporated following Neugebauer et al.,¹⁵ which has also been implemented for CC2 by Höfener et al.³²

III. COMPUTATIONAL DETAILS

A. Structures

Bond parameters of the $\text{CO}_2 \cdots \text{Rg}$ complexes were taken from Maroulis et al.⁷⁶ Coordinates of $\text{NH}_3 \cdots \text{H}_2\text{O}$ complex was taken from our previous work.⁴⁹ The solvated uracil structure was taken from Etinski and Marian.⁷⁷ The xyz structures of all investigated molecules are listed in the SI.

B. Generation of embedding potential

The ‘Grid’ package was used and interfaced with PyBEST to generate numerical integration grids for embedding potentials.⁷⁸ All embedding calculations involve supramolecu-

lar grids. The grids were constructed using the Becke radial transformation in combination with Lebedev–Laikov angular grids with 131 points for all heavy atoms.⁷⁹ Becke partitioning weights were used in the numerical integration.

pCCD-based freeze-and-thaw (FT) cycles were run as described in section II B. Convergence was accepted when the change in subsystem pCCD energies of both subsystems went below $10^{-8} E_h$ between two consecutive cycles. For the weakly interacting systems investigated in this work, it takes at most 3–4 FT cycles to fully converge the pCCD energies of each subsystem. Also, for any subsystem, beyond the step in the FDE workflow (see Figure 1) where the effect of the appropriate embedding potential is introduced in the calculation for the first time, the energy change and changes in dipole moment are minimal. Taking this into account, we also explored another simplified strategy. The pCCD densities and embedding potentials of the environment subsystem were generated in vacuum; these potentials were then used to embed the active subsystem in a single-step calculation. Such an approach is referred to as ‘non-FT’ throughout this work.

C. Coupled cluster calculations in PyBEST

All pCCD and post-pCCD calculations have been performed in a developer version of the PyBEST v2.1.0dev0 software package.^{80–82} In this work, we used monomer-centered basis functions, specifically the aug-cc-pVDZ basis set.⁸³ Pipek–Mezey⁸⁴ localized orbitals were used in all cases. However, pCCD orbital optimization was not performed. The major reason is to avoid computationally expensive orbital rotations during each FT macro-iteration. As far as the ground state molecular properties are concerned, our previous works have shown that the use of optimized orbitals actually worsens the performance of pCCD-based methods to a certain extent.^{85,86} This is also true for a certain type of excited states.⁴⁶ In addition, the excitation energies of EOM-fpLCCSD do not differ much between the canonical HF and natural pCCD orbitals. Still, excited-state identification is simpler and easier in the canonical picture.⁶⁷ Furthermore, the behavior of the pCCD orbital optimization during FT cycles needs to be investigated in detail, which is beyond the scope of this work. Since we are not performing orbital optimization, the reference wave function $|\Phi_0\rangle$ in Eq. (7) refers to the HF Slater determinant. All electronic integrals were generated with the Libint package.⁸⁷ Cholesky decomposition^{88,89} was used with a tight threshold of 10^{-8} . The frozen-core approximation was used in all correlated calculations as implemented in PyBEST.

Expectation-value-based (denoted by XpCCD, XfpCCSD, XfpLCCSD, XCCSD) dipole moments were calculated using the cluster operator amplitudes and $\hat{S}$ -operator based approximations,⁹⁰ as described in our previous publication.⁸⁶ The response fpLCCSD dipole moments (marked without X), on the other hand, are calculated by solving the associated Λ -equations.⁸⁵ All fpCCSD and fpLCCSD calculations are performed after convergence of the freeze-and-thaw cycles in the FT framework. Though the CCSD calculation is not a post-pCCD step in a theoretical sense, we use the perturbed HF

MOs of the subsystems from the final FT iteration, and hence these were also included as a part of our ‘post-pCCD’ step in Fig. 1. CCSD(T) response dipole moments were obtained from the Molpro package.^{91,92}

IV. RESULTS AND DISCUSSION

A. Dipole moments of $\text{CO}_2 \cdots \text{Rg}$ complexes

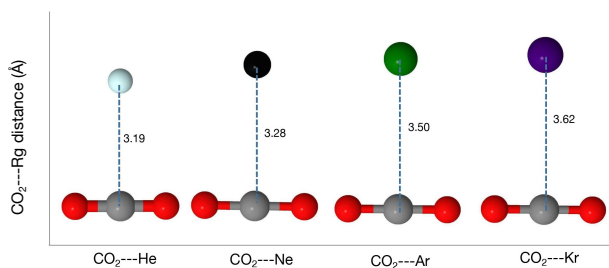


FIG. 2: Distances of the Rg atoms and CO_2 moiety in the $\text{CO}_2 \cdots \text{Rg}$ complexes. The C=O distance is kept constant at 1.16 Å.

The dipole moment of the van der Waals complexes of $\text{CO}_2 \cdots \text{Rg}$ ($\text{Rg} = \text{He}, \text{Ne}, \text{Ar}, \text{and Kr}$) is mainly governed by electrostatic and Pauli interactions between the CO_2 moiety and the inert gas atom. The weak interactions make the dipole moments of such complexes highly sensitive to changes in electron density, providing an excellent testing ground for embedding approaches.^{85,93,94} All four complexes adopt a T-shaped equilibrium (as shown in Figure 2), with the CO_2 molecule lying along the x -axis (carbon atom at the origin) and the rare-gas atom located along the positive z -axis. Consequently, the z -component of the dipole moment is the dominant contribution and serves as the most relevant quantity for assessing the quality of the embedding scheme.

Supramolecular dipole moments, obtained from the response density matrices of pCCD and fpLCCSD, are compared with CCSD(T) reference values in Table I. Overall, pCCD shows good agreement with the CCSD(T) benchmarks. For the $\text{CO}_2 \cdots \text{He}$ and $\text{CO}_2 \cdots \text{Ne}$ complexes, pCCD slightly overestimates the dipole moments by approximately 13%. In the $\text{CO}_2 \cdots \text{Kr}$ complex, the agreement improves markedly, with pCCD underestimating the dipole moment by only 2.54%. In contrast, fpLCCSD substantially overestimates the dipole moments across all complexes, with relative errors ranging from 35% to 50%. This poor performance of fpLCCSD is not unexpected. As shown in our previous work,^{85,86} the linearized correction model has difficulty accurately reproducing dipole moments for non-ionic systems, particularly small carbon-based molecules. To further investigate this behavior, we performed the same calculations using expectation-value density matrices. The corresponding results

are summarized in Table II. A comparison of the supramolecular dipole moments in Tables I and II reveals a clear improvement in the fpLCCSD results when the expectation-value-based approach is employed (denoted as XpCCD, XCCSD, and XfpCC(S)D). However, a detailed analysis of this lies beyond the scope of the present work, since our primary focus is the development and assessment of the pCCD-based embedding schemes.

We next examine the embedding results presented in Tables I and II. Given that the investigated complexes are only weakly bound, self-consistent polarization is expected to have only a minor influence on the molecular properties. As shown in both tables, the non-FT and FT dipole moments are very similar at each level of theory (pCCD or fpLCCSD). The FT cycles provide a slight improvement in the embedded dipole moments for both (X)pCCD and (X)fpLCC for the $\text{CO}_2 \cdots \text{He}$ and $\text{CO}_2 \cdots \text{Ne}$ complexes. For the heavier $\text{CO}_2 \cdots \text{Ar}$ and $\text{CO}_2 \cdots \text{Kr}$ complexes, the non-FT approach yields more reliable results. Overall, the (X)fpLCCSD-based embedding schemes produce dipole moments in closer agreement with the supramolecular CCSD(T) reference than the (X)pCCD-based schemes. In these weakly interacting complexes, the dipole moment resides predominantly on the inert gas atom, since the interaction is too weak to significantly polarize the CO_2 unit. Consequently, the dynamic correlation correction primarily affects the dipole moment of the inert gas atom. Here, (X)fpLCC(S)D captures electron correlation in a more balanced manner than in the supramolecular calculation, where it tends to overestimate the induced polarization (distortion) of the CO_2 electron cloud. As a result, the (X)fpLCCSD-in-pCCD embedding approaches consistently improve upon the supramolecular (X)fpLCCSD dipole moments. Finally, among the embedding schemes considered, XfpLCCSD-in-pCCD(FT) provides the most reliable results for the $\text{CO}_2 \cdots \text{He}$ and $\text{CO}_2 \cdots \text{Ne}$ complexes (with dipole moment errors below 9% and 12%, respectively), while XfpLCCSD-in-pCCD(non-FT) yields the best agreement for the $\text{CO}_2 \cdots \text{Ar}$ and $\text{CO}_2 \cdots \text{Kr}$ complexes (with errors below 1%). It is to be noted here that for these expectation-value-based calculations, only the dipole moments were calculated using the 1-RDM obtained from cluster amplitudes. The pCCD electron densities, for the embedding potentials, were still obtained with the associated response Λ -equations.

B. Vertical excitation energies of the $\text{H}_2\text{O} \cdots \text{NH}_3$ complex

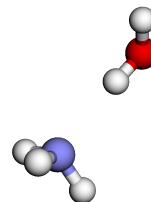


FIG. 3: Schematic structure of the $\text{H}_2\text{O} \cdots \text{NH}_3$ complex.

TABLE I: z -component of the response dipole moments (in D) of the $\text{CO}_2 \cdots \text{Rg}$ complexes with percentage error (%) relative to the response CCSD(T) values in parentheses. FT denotes the fully converged freeze-and-thaw self-consistent embedding scheme, whereas non-FT corresponds to calculations using a static embedding potential obtained from the isolated subsystems.

Method / aug-cc-pVDZ	$\text{CO}_2 \cdots \text{He}$	$\text{CO}_2 \cdots \text{Ne}$	$\text{CO}_2 \cdots \text{Ar}$	$\text{CO}_2 \cdots \text{Kr}$
pCCD-in-pCCD(non-FT)	0.0144 (-18.18)	0.0196 (-26.59)	0.0720 (-2.83)	0.0920 (-2.54)
fpLCCSD-in-pCCD(non-FT)	0.0150 (-14.77)	0.0228 (-14.61)	0.0747 (0.81)	0.0939 (-0.53)
pCCD-in-pCCD(FT)	0.0156 (-11.36)	0.0217 (-18.73)	0.0772 (4.18)	0.0978 (3.60)
fpLCCSD-in-pCCD(FT)	0.0160 (-9.09)	0.0237 (-11.24)	0.0802 (8.23)	0.1001 (6.04)
pCCD(supra.)	0.0198 (12.50)	0.0312 (16.85)	0.0742 (0.13)	0.0920 (-2.54)
fpLCCSD(supra.)	0.0264 (50.00)	0.0373 (39.70)	0.0997 (34.55)	0.1327 (40.57)
CCSD(T)	0.0176	0.0267	0.0741	0.0944

TABLE II: z -component of the expectation value (X) dipole moments (in D) of the $\text{CO}_2 \cdots \text{Rg}$ complexes with percentage error (%) relative to the response CCSD(T) values in parentheses. FT denotes the fully converged freeze-and-thaw self-consistent embedding scheme, whereas non-FT corresponds to calculations using a static embedding potential obtained from the isolated subsystems.

Method / aug-cc-pVDZ	$\text{CO}_2 \cdots \text{He}$	$\text{CO}_2 \cdots \text{Ne}$	$\text{CO}_2 \cdots \text{Ar}$	$\text{CO}_2 \cdots \text{Kr}$
XpCCD-in-pCCD(non-FT)	0.0144 (-18.18)	0.0196 (7.49)	0.0720 (-2.83)	0.0920 (-2.54)
XfpCCSD-in-pCCD(non-FT)	0.0149 (-15.34)	0.0220 (-17.60)	0.0744 (0.40)	0.0938 (-0.64)
XfpLCCSD-in-pCCD(non-FT)	0.0149 (-15.34)	0.0225 (-15.73)	0.0747 (0.81)	0.0940 (-0.42)
XCCSD-in-pCCD (non-FT)	0.0151 (-14.20)	0.0226 (-15.36)	0.0741 (0.00)	0.0934 (-1.06)
XpCCD-in-pCCD(FT)	0.0156 (-11.36)	0.0217 (-18.73)	0.0772 (4.18)	0.0978 (3.60)
XfpCCSD-in-pCCD(FT)	0.0160 (-9.09)	0.0234 (-12.36)	0.0797 (7.56)	0.0998 (5.72)
XfpLCCSD-in-pCCD(FT)	0.0160 (-9.09)	0.0236 (-11.61)	0.0801 (8.10)	0.1001 (6.04)
XCCSD-in-pCCD (FT)	0.0161 (-8.52)	0.0236 (-11.61)	0.0797 (7.56)	0.0940 (-0.42)
XpCCD(supra.)	0.0198 (12.50)	0.0312 (16.85)	0.0788 (6.34)	0.0920 (-2.54)
XfpCCSD(supra.)	0.0202 (14.77)	0.0330 (23.60)	0.0771 (4.05)	0.1108 (17.37)
XfpLCCSD(supra.)	0.0200 (13.64)	0.0324 (21.35)	0.0746 (0.67)	0.1094 (15.90)
XCCSD(supra.)	0.0180 (2.27)	0.0279 (4.49)	0.0757 (2.16)	0.0966 (2.33)
CCSD(T)	0.0176	0.0267	0.0741	0.0944

TABLE III: Vertical excitation energies (in eV) computed with EOM-fpLCCSD/aug-cc-pVDZ for the $\text{H}_2\text{O} \cdots \text{NH}_3$ model system. Results are shown for the isolated subsystems at the complex geometry (Iso. subsystem), the EOM-fpLCCSD-in-pCCD/aug-cc-pVDZ embedding calculation, and the supramolecular EOM-fpLCCSD/aug-cc-pVDZ reference.

Source subsystem	Iso. subsystem	EOM-fpLCCSD-in-pCCD		Supramolecule
		non-FT	FT	
NH_3	6.43	6.88	6.99	6.90
	7.97	8.45	8.58	8.47
	7.97	8.50	8.62	8.52
H_2O	7.58	7.47	7.44	7.49

Table III reports the three lowest-lying vertical singlet-

singlet excitation energies (VEEs) originating from the NH_3 moiety and one excitation from the H_2O moiety in the $\text{H}_2\text{O} \cdots \text{NH}_3$ complex (structure shown in Figure 3). Supramolecular EOM-fpLCCSD calculations confirm that these excitations are entirely localized on their respective parent fragments. The remaining excitations from both fragments exhibit charge-transfer character and, therefore, cannot be described by our embedding model. The non-FT embedding results (NH_3 embedded in H_2O) are already in excellent agreement with the supramolecular reference, with deviations of only 0.02 eV. In contrast, the FT cycles lead to a non-negligible overestimation of the VEEs for the NH_3 -localized excitations (by 0.12–0.16 eV). These numerical results suggest that the FT procedure overestimates the polarization of the NH_3 ground state by the H_2O environment. For comparison, the bare NH_3 fragment at the complex geometry yields VEEs of 6.43 eV and 7.97 eV. The hydrogen bond with H_2O induces a blue shift of these transitions, which is overestimated in the FT calculations due to excessive polarization of

the NH_3 fragment by the environment. These findings are consistent with previous CC-in-CC freeze-thaw (FT) calculations on the same complex.³¹

The overpolarization effect is less pronounced when H_2O is embedded in NH_3 . The first excitation of the bare H_2O fragment at the complex geometry is located at 7.58 eV. The influence of the NH_3 environment on this transition is significantly smaller than in the reverse case. Both the non-FT and FT embedding schemes reproduce this excitation accurately, with deviations from the supramolecular reference of only 0.02 eV and 0.05 eV, respectively.

C. Vertical excitation energies of microsolvated Uracil

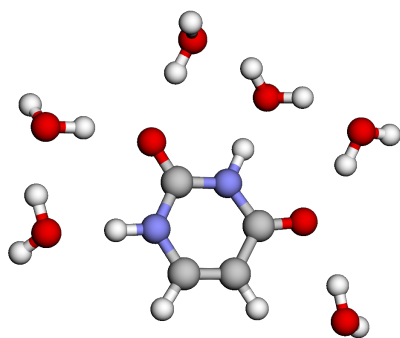


FIG. 4: Schematic structure of uracil ($\text{C}_4\text{H}_4\text{N}_2\text{O}_2$) surrounded by six H_2O molecules.

TABLE IV: EOM-fpLCCSD/aug-cc-pVDZ vertical excitation energies (in eV) for the $\text{C}_4\text{H}_4\text{N}_2\text{O}_2 \cdot (\text{H}_2\text{O})_6$ structure. Comparison between the isolated uracil (at the complex geometry) and EOM-fpLCCSD-in-pCCD/aug-cc-pVDZ embedding results.

Iso. uracil	EOM-fpLCCSD-in-pCCD		Supramolecule
	non-FT	FT	
5.62	5.59	5.75	5.73
5.16	5.70	5.97	5.89

Table IV compares the vertical excitation energies (VEEs) of uracil surrounded by six water molecules (Figure 4) calculated at the supramolecular EOM-fpLCCSD level with those obtained using the embedding approach. At the supramolecular level, the lowest-lying transition is of $\pi \rightarrow \pi^*$ character at 5.73 eV, followed by an $n_{\text{O}} \rightarrow \pi^*$ [n_{O} denoting the non-bonded electron pair of O] transition at 5.89 eV. For isolated uracil (at the complex geometry), the order of these states is reversed: the $n_{\text{O}} \rightarrow \pi^*$ excitation appears at 5.16 eV and the $\pi \rightarrow \pi^*$ excitation at 5.62 eV. Consequently, the water environment induces a substantial blue shift of +0.73 eV for the $n_{\text{O}} \rightarrow \pi^*$ transition and a much smaller shift of +0.11 eV for the $\pi \rightarrow \pi^*$ transition. In the embedding calculations,

all six water molecules are treated collectively as the environment, with uracil as the active subsystem. Compared to the supramolecular reference, the non-FT embedding results deviate by 0.14 eV and 0.19 eV for the two transitions, respectively. The FT cycles significantly improve the accuracy, reducing the errors to 0.02 eV and 0.08 eV. The presence of several H_2O molecules in this case causes higher environmental polarization effect. This is reflected in the fact that, particularly for this system, the FT calculations took more (9, precisely) steps to converge, which is higher than the other weakly interacting complexes discussed above. Hence, it is not surprising to see the improvement in the VEEs brought in by the FT cycles compared to the non-FT calculations.

V. CONCLUSION

In this work, we have implemented a new WFT-in-WFT density embedding framework in the PyBEST software package. The approach integrates the pCCD method with a simple frozen-density embedding (FDE) scheme. Its efficiency stems from the low computational scaling of pCCD and the straightforward access to its response and expectation-value 1-RDMs, which enable inexpensive and accurate evaluation of the electron density in PyBEST. The pCCD-based embedding scheme is particularly advantageous when DFT fails to reliably describe the ground-state density, such as in strongly correlated systems. It accurately reproduces the dipole moments of weakly interacting $\text{CO}_2 \cdots \text{Rg}$ complexes and the vertical excitation energies (VEEs) of microsolvated uracil. Despite the very small magnitude of the dipole moments in the $\text{CO}_2 \cdots \text{Rg}$ systems, our embedding approach recovers them within 5–10% of the supramolecular CCSD(T) reference, outperforming even supramolecular pCCD-based calculations. For excited states, the method successfully describes localized excitations in the $\text{H}_2\text{O} \cdots \text{NH}_3$ complex. The EOM-fpLCCSD-in-pCCD results yield VEEs in close agreement with supramolecular EOM-fpLCCSD calculations while correctly preserving the character and ordering of the excitations in this hydrogen-bonded system.

Overall, these results demonstrate that the proposed embedding framework can reliably treat a wide range of intermolecular interactions. The accuracy of excited-state calculations can be further enhanced by incorporating subsystem couplings via linear-response methods.^{15,32,46} The promising performance on small systems motivates extensions to more challenging cases, particularly actinide complexes embedded in large organic ligands, where pCCD has already shown reliable results for both ground and excited states.^{45,49,95} Future developments will focus on these directions. In addition, we plan to accelerate the embedding potential evaluation on modern GPU architectures,^{82,96} thereby extending the applicability of the method to significantly larger systems.

VI. ACKNOWLEDGMENT

R.C. and P.T. acknowledge financial support from the OPUS research grant from the National Science Centre, Poland (Grant No. 2019/33/B/ST4/02114). We gratefully acknowledge the Polish high-performance computing infrastructure PLGrid (HPC Centers: ACK Cyfronet AGH and WCSS) for providing computer facilities and support within the computational grant no. PLG/2025/018840.

- ¹M. S. Gordon, D. G. Fedorov, S. R. Pruitt, and L. V. Slipchenko, *Chem. Rev.* **112**, 632 (2012).
- ²A. S. P. Gomes and C. R. Jacob, *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **108**, 222 (2012).
- ³L. O. Jones, M. A. Mosquera, G. C. Schatz, and M. A. Ratner, *J. Am. Chem. Soc.* **142**, 3281 (2020).
- ⁴J. Liu and X. He, *WIREs Comput. Mol. Sci.* **13**, e1650 (2023).
- ⁵Q. Sun and G. K.-L. Chan, *Acc. Chem. Res.* **49**, 2705 (2016).
- ⁶P. Cortona, *Phys. Rev. B* **46**, 2008 (1992).
- ⁷T. A. Wesolowski and A. Warshel, *J. Phys. Chem.* **97**, 8050 (1993).
- ⁸T. A. Wesolowski, S. Shedge, and X. Zhou, *Chem. Rev.* **115**, 5891 (2015).
- ⁹C. R. Jacob, S. M. Beyhan, and L. Visscher, *J. Chem. Phys.* **126**, 234116 (2007).
- ¹⁰S. Fux, C. R. Jacob, J. Neugebauer, L. Visscher, and M. Reiher, *J. Chem. Phys.* **132**, 164101 (2010).
- ¹¹A. W. Gotz, S. M. Beyhan, and L. Visscher, *J. Chem. Theory Comput.* **5**, 3161 (2009).
- ¹²J. D. Goodpaster, N. Ananth, F. R. Manby, and T. F. Miller, *J. Chem. Phys.* **133**, 084103 (2010).
- ¹³P. Tecmer, H. van Lingem, A. S. P. Gomes, and L. Visscher, *J. Chem. Phys.* **137**, 084308 (2012).
- ¹⁴M. E. Casida and T. A. Wesolowski, *Int. J. Quantum Chem.* **96**, 577 (2004).
- ¹⁵J. Neugebauer, *J. Chem. Phys.* **126**, 134116 (2007).
- ¹⁶J. Neugebauer, *J. Chem. Phys.* **131**, 084104 (2009).
- ¹⁷F. R. Manby, M. Stella, J. D. Goodpaster, and T. F. Miller III, *J. Chem. Theory Comput.* **8**, 2564 (2012).
- ¹⁸S. J. Lee, M. Welborn, F. R. Manby, and T. F. Miller III, *Acc. Chem. Res.* **52**, 1359 (2019).
- ¹⁹P. Beran, K. Pernal, F. Pavosevic, and L. Veis, *J. Phys. Chem. Lett.* **14**, 716 (2023).
- ²⁰E. Monino, D. Drwal, P. Beran, M. Hapka, L. Veis, and K. Pernal, *J. Chem. Theory Comput.* **22**, 1780 (2026).
- ²¹B. Hégely, P. R. Nagy, G. G. Ferenczy, and M. Kállay, *J. Chem. Phys.* **145**, 064107 (2016).
- ²²D. S. Graham, X. Wen, D. V. Chulhai, and J. D. Goodpaster, *J. Chem. Theory Comput.* **16**, 2284 (2020).
- ²³J. Csóka, B. Hégely, P. R. Nagy, and M. Kállay, *J. Chem. Phys.* **160**, 124113 (2024).
- ²⁴T. A. Wesolowski, *Phys. Rev. A* **77**, 012504 (2008).
- ²⁵N. Govind, Y. Wang, A. Da Silva, and E. Carter, *Chem. Phys. Lett.* **295**, 129 (1998).
- ²⁶N. Govind, Y. A. Wang, and E. A. Carter, *J. Chem. Phys.* **110**, 7677 (1999).
- ²⁷S. Höfener, A. Severo Pereira Gomes, and L. Visscher, *J. Chem. Phys.* **136**, 044104 (2012).
- ²⁸S. Höfener, A. S. P. Gomes, and L. Visscher, *J. Chem. Phys.* **139**, 104106 (2013).
- ²⁹A. S. P. Gomes, C. R. Jacob, and L. Visscher, *Phys. Chem. Chem. Phys.* **10**, 5353 (2008).
- ³⁰A. S. P. Gomes, C. R. Jacob, F. Réal, L. Visscher, and V. Vallet, *Phys. Chem. Chem. Phys.* **15**, 15153 (2013).
- ³¹S. Höfener and L. Visscher, *J. Chem. Phys.* **137**, 204120 (2012).
- ³²S. Hofener and L. Visscher, *J. Chem. Theory Comput.* **12**, 549 (2016).
- ³³P. A. Limacher, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck, and P. Bultinck, *J. Chem. Theory Comput.* **9**, 1394 (2013).
- ³⁴T. Stein, T. M. Henderson, and G. E. Scuseria, *J. Chem. Phys.* **140**, 214113 (2014).
- ³⁵K. Boguslawski, P. Tecmer, P. W. Ayers, P. Bultinck, S. De Baerdemacker, and D. Van Neck, *Phys. Rev. B* **89**, 201106(R) (2014).
- ³⁶K. Boguslawski, P. Tecmer, P. A. Limacher, P. A. Johnson, P. W. Ayers, P. Bultinck, S. De Baerdemacker, and D. Van Neck, *J. Chem. Phys.* **140**, 214114 (2014).
- ³⁷T. M. Henderson, I. W. Bulik, T. Stein, and G. E. Scuseria, *J. Chem. Phys.* **141**, 244104 (2014).
- ³⁸K. Boguslawski, P. Tecmer, P. W. Ayers, P. Bultinck, S. De Baerdemacker, and D. Van Neck, *J. Chem. Theory Comput.* **10**, 4873 (2014).
- ³⁹K. Boguslawski, P. Tecmer, and Ö. Legeza, *Phys. Rev. B* **94**, 155126 (2016).
- ⁴⁰P. Tecmer and K. Boguslawski, *Phys. Chem. Chem. Phys.* **24**, 23026 (2022).
- ⁴¹P. Tecmer, K. Boguslawski, and P. W. Ayers, *Phys. Chem. Chem. Phys.* **17**, 14427 (2015).
- ⁴²P. Tecmer, M. Gałyńska, L. Szczuczko, and K. Boguslawski, *J. Phys. Chem. Lett.* **14**, 9909 (2023).
- ⁴³R. D. Pandey, M. Gałyńska, K. Boguslawski, and P. Tecmer, *J. Phys. Chem. A* **130**, 446 (2026).
- ⁴⁴M. Gałyńska, M. M. F. de Moraes, P. Tecmer, and K. Boguslawski, *Phys. Chem. Chem. Phys.* **26**, 18918 (2024).
- ⁴⁵A. Nowak, P. Tecmer, and K. Boguslawski, *Phys. Chem. Chem. Phys.* **21**, 19039 (2019).
- ⁴⁶S. Ahmadvani, K. Boguslawski, and P. Tecmer, *J. Chem. Theory Comput.* **20**, 10443 (2024).
- ⁴⁷M. Gałyńska and K. Boguslawski, *J. Chem. Theory Comput.* **20**, 4182 (2024).
- ⁴⁸R. D. Pandey, M. M. F. de Moraes, K. Boguslawski, and P. Tecmer, *J. Chem. Theory Comput.* **21**, 5049 (2025).
- ⁴⁹R. Chakraborty, K. Boguslawski, and P. Tecmer, *Phys. Chem. Chem. Phys.* **25**, 25377 (2023).
- ⁵⁰P. Tecmer, K. Boguslawski, P. A. Limacher, P. A. Johnson, M. Chan, T. Verstraelen, and P. W. Ayers, *J. Phys. Chem. A* **118**, 9058 (2014).
- ⁵¹P. A. Limacher, T. D. Kim, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck, and P. Bultinck, *Mol. Phys.* **112**, 853 (2014).
- ⁵²P. Tecmer, K. Boguslawski, M. Borkowski, P. S. Żuchowski, and D. Kędziera, *Int. J. Quantum Chem.* **119**, e25983 (2019).
- ⁵³F. Brzęk, K. Boguslawski, P. Tecmer, and P. S. Żuchowski, *J. Chem. Theory Comput.* **15**, 4021 (2019).
- ⁵⁴P. Tecmer and K. Boguslawski, *Phys. Chem. Chem. Phys.* **24**, 23026 (2022).
- ⁵⁵T. Kinoshita, O. Hino, and R. J. Bartlett, *J. Chem. Phys.* **123**, 074106 (2005).
- ⁵⁶M. Mörchen, L. Freitag, and M. Reiher, *J. Chem. Phys.* **153**, 244113 (2020).
- ⁵⁷N. Oliphant and L. Adamowicz, *J. Chem. Phys.* **96**, 3739 (1992).
- ⁵⁸P. Piecuch, N. Oliphant, and L. Adamowicz, *J. Chem. Phys.* **99**, 1875 (1993).
- ⁵⁹P. Piecuch and L. Adamowicz, *J. Chem. Phys.* **100**, 5792 (1994).
- ⁶⁰A. Leszczyk, M. Máté, Ö. Legeza, and K. Boguslawski, *J. Chem. Theory Comput.* **18**, 96 (2022).
- ⁶¹T. M. Henderson, I. W. Bulik, T. Stein, and G. E. Scuseria, *J. Chem. Phys.* **141**, 244104 (2014).
- ⁶²K. Boguslawski and P. W. Ayers, *J. Chem. Theory Comput.* **11**, 5252 (2015).
- ⁶³D. J. Rowe, *Rev. Mod. Phys.* **40**, 153 (1968).
- ⁶⁴J. Geersten, M. Rittby, and R. J. Bartlett, *Chem. Phys. Lett.* **164**, 57 (1989).
- ⁶⁵J. F. Stanton and R. J. Bartlett, *J. Chem. Phys.* **98**, 7029 (1993).
- ⁶⁶K. Boguslawski and P. Tecmer, *arXiv v1*, arXiv:2604.13009 (2026).
- ⁶⁷K. Boguslawski, *J. Chem. Theory Comput.* **15**, 18 (2019).
- ⁶⁸K. Boguslawski, P. Tecmer, and Ö. Legeza, *Phys. Rev. B* **94**, 155126 (2016).
- ⁶⁹L. H. Thomas, *Math. Proc. Camb. Phil. Soc.* **23**, 542 (1927).
- ⁷⁰E. H. Lieb, *Rev. Mod. Phys.* **53**, 603 (1981).
- ⁷¹I. Slater, *Phys. Rev.* **81**, 385 (1951).
- ⁷²R. R. Zope and B. I. Dunlap, *J. Chem. Theory Comput.* **1**, 1193 (2005).
- ⁷³W. Kohn and L. J. Sham, *Physical review* **140**, A1133 (1965).
- ⁷⁴T. A. Wesolowski and F. Tran, *J. Chem. Phys.* **118**, 2072 (2003).
- ⁷⁵T. A. Wesolowski and J. Weber, *Chem. Phys. Lett.* **248**, 71 (1996).
- ⁷⁶G. Maroulis and A. Haskopoulos, *Chem. Phys. Lett.* **349**, 335 (2001).
- ⁷⁷M. Etinski and C. M. Marian, *Phys. Chem. Chem. Phys.* **12**, 4915 (2010).
- ⁷⁸A. Tehrani, X. D. Yang, M. Martínez-González, L. Pujal, R. Hernández-Esparza, M. Chan, E. Vöhringer-Martinez, T. Verstraelen, P. W. Ayers, and F. Heidar-Zadeh, *J. Chem. Phys.* **160**, 172503 (2024).

- ⁷⁹V. I. Lebedev and D. N. Laikov, *Doklady Akademii nauk SSSR* **366**, 741 (1999).
- ⁸⁰K. Boguslawski, A. Leszczyk, A. Nowak, F. Brzęk, P. S. Żuchowski, D. Kędziera, and P. Tecmer, *Comput. Phys. Commun.* **264**, 107933 (2021).
- ⁸¹K. Boguslawski, F. Brzęk, R. Chakraborty, K. Cieślak, S. Jahani, A. Leszczyk, A. Nowak, E. Sujkowski, J. Świerczyński, S. Ahmadkhani, D. Kędziera, M. H. Kriebel, P. S. Żuchowski, and P. Tecmer, *Comput. Phys. Commun.* **297**, 109049 (2024).
- ⁸²M. H. Kriebel, P. Tecmer, M. Gałyńska, A. Leszczyk, and K. Boguslawski, *J. Chem. Theory Comput.* **24**, 1130 (2024).
- ⁸³R. A. Kendall, T. H. Dunning, and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
- ⁸⁴J. Pipek and P. G. Mezey, *J. Chem. Phys.* **90**, 4916 (1989).
- ⁸⁵R. Chakraborty, M. M. F. de Moraes, K. Boguslawski, A. Nowak, J. Świerczyński, and P. Tecmer, *J. Chem. Theory Comput.* **20**, 4689 (2024).
- ⁸⁶R. Chakraborty, S. Ahmadkhani, J. Świerczyński, K. Boguslawski, and P. Tecmer, *J. Phys. Chem. A* **129**, 6713 (2025).
- ⁸⁷E. F. Valeev, "Libint: A library for the evaluation of molecular integrals of many-body operators over gaussian functions," <http://libint.valeev.net/> (2024), version 2.9.0.
- ⁸⁸H. Koch, A. Sánchez de Merás, and T. B. Pedersen, *J. Chem. Phys.* **118**, 9481 (2003).
- ⁸⁹F. Aquilante, L. Boman, J. Boström, H. Koch, R. Lindh, A. S. de Merás, and T. B. Pedersen, *Linear-Scaling Techniques in Computational Chemistry and Physics: Methods and Applications*, 301 (2011).
- ⁹⁰T. Korona, *Theor. Chem. Acc.* **129**, 15 (2011).
- ⁹¹H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2**, 242 (2012).
- ⁹²H.-J. Werner, P. J. Knowles, F. R. Manby, J. A. Black, K. Doll, A. Heßelmann, D. Kats, A. Köhn, T. Korona, D. A. Kreplin, and et al., *J. Chem. Phys.* **152**, 144107 (2020).
- ⁹³T. Wesołowski, Y. Ellinger, and J. Weber, *J. Chem. Phys.* **108**, 6078 (1998).
- ⁹⁴T. A. Wesołowski, P.-Y. Morgantini, and J. Weber, *J. Chem. Phys.* **116**, 6411 (2002).
- ⁹⁵P. Tecmer, K. Boguslawski, and P. W. Ayers, *Phys. Chem. Chem. Phys.* **17**, 14427 (2015).
- ⁹⁶A. Dobrowolska, J. Świerczyński, P. Tecmer, E. Sujkowski, S. Ahmadkhani, G. Mazur, K. Noga, J. Hammond, and K. Boguslawski, *arXiv* **v1**, arXiv:2603.20912 (2026).

Frozen density embedding with pCCD electron densities

Rahul Chakraborty¹ and Paweł Tecmer¹

Institute of Physics, Faculty of Physics, Astronomy, and Informatics, Nicolaus
Copernicus University in Toruń, Grudziądzka 5,
87-100 Toruń, Poland.

(*Electronic mail: rahuldev.chakraborty91@gmail.com)

Supplementary Information

Section I. Coordinates of the complexes studied in this work

1. CO₂-He

C	0.000000000000	0.000000000000	0.000000000000
O	1.159956529808	0.000000000000	0.000000000000
O	-1.159956529808	0.000000000000	0.000000000000
He	0.000000000000	0.000000000000	3.185117861731

2. CO₂-Ne

C	0.000000000000	0.000000000000	0.000000000000
O	1.159956529808	0.000000000000	0.000000000000
O	-1.159956529808	0.000000000000	0.000000000000
Ne	0.000000000000	0.000000000000	3.284059780000

3. CO₂-Ar

C	0.000000000000	0.000000000000	0.000000000000
O	1.159956529808	0.000000000000	0.000000000000
O	-1.159956529808	0.000000000000	0.000000000000
Ar	0.000000000000	0.000000000000	3.495215470000

4. CO₂-kr

C	0.000000000000	0.000000000000	0.000000000000
O	1.159956529808	0.000000000000	0.000000000000
O	-1.159956529808	0.000000000000	0.000000000000
Kr	0.000000000000	0.000000000000	3.622747190000

5. H₂O-NH₃

N	-0.3247549	-1.0163481	0.2284577
H	0.3853716	-1.6985302	0.3923495
H	-0.9158405	-0.9874769	1.0326156
H	-0.8749286	-1.3247542	-0.5458969
O	0.3171256	1.9052994	-0.5156991
H	1.2106186	2.1227180	-0.3263054
H	0.2024080	0.9990920	-0.2655214

6. Uracil-6H₂O

C	0.109941	0.038928	-0.098341
C	1.458110	0.057796	-0.104838
C	2.137130	1.328745	0.003333
N	1.309292	2.443598	0.082128
C	-0.058514	2.422204	0.119470
N	-0.634478	1.181781	0.010750
O	3.366554	1.466099	0.030820
O	-0.734251	3.452420	0.245192
H	-0.463231	-0.875800	-0.170295
H	2.054103	-0.839020	-0.185547
H	-1.664936	1.118742	0.010448
H	1.768158	3.393847	0.093400
O	-3.508738	3.680667	0.428178
H	-2.534027	3.726915	0.351564
H	-3.861048	4.426543	-0.066725
O	4.766427	3.701986	1.086355
H	4.399641	2.870666	0.727409
H	5.682794	3.741131	0.796423
O	2.501757	4.893902	0.074878
H	1.954082	5.509746	0.591314
H	3.371457	4.787807	0.505441
O	4.505193	-1.140661	-0.341887
H	4.306192	-0.198231	-0.200356
H	5.134614	-1.165518	-1.068982
O	-3.442345	1.001156	0.043743
H	-3.715413	1.938150	0.167843
H	-4.039097	0.618838	-0.606622
O	0.214088	5.975344	1.229986
H	-0.045001	6.047716	2.154035
H	-0.193096	5.152781	0.900644