SSM
Scuola Superiore Meridionale

SCUOLA NORMALE SUPERIORE

UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

Doctoral Thesis in
Molecular Sciences for Earth and Space

XXXVII CYCLE

Towards Accurate Thermochemical Properties: Computational Methodologies for Molecular Systems of Increasing Dimensions

---

Author:
Silvia Di Grande

Supervisor:
Prof. Nicola Tasinato
Co-Supervisors:
Prof. Vincenzo Barone
Prof. Mihály Kállay

A.A. 2024/2025

Coordinator: Prof. Nadia Rega

SCUOLA SUPERIORE MERIDIONALE
SCUOLA NORMALE SUPERIORE
UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

Abstract

Doctoral thesis in
Molecular Sciences for Earth and Space
XXXVII Cycle

Towards Accurate Thermochemical Properties: Computational Methodologies for Molecular Systems of Increasing Dimensions

Silvia Di Grande

The quest for ever-increasing accuracy remains one of the central challenges in quantum chemistry, motivating the development of methods that aim to achieve a one-to-one correspondence between experiments and first-principles models. However, the most accurate approaches are generally feasible only for small molecules, while methods applicable to large systems rely on approximations that limit their predictive power. This thesis tackles this limitation by developing new composite approaches that combine accuracy and scalability, designed to bridge the gap between the size of the molecular systems investigated and the level of reliability of simulated properties.

This objective is pursued by developing three closely related research lines. First, the junChS-F12 model is improved, leading to a composite scheme (CS) capable of providing accurate thermochemical and rotational–vibrational spectroscopic properties for medium-size molecules containing elements from the first three rows of the periodic table. This advancement has improved the predictive power of the existing explicitly correlated ChS framework and extended its applicability beyond energetics to equilibrium geometries and harmonic vibrational frequencies. The results have established the junChS-F12 as a robust and consistent CS, forming a solid foundation for subsequent methodological developments.

Building on these achievements, in a second step, the Pisa Composite Scheme (PCS) family is introduced a new and modular generation of composite protocols offering higher accuracy than the ChS while retaining similar computational demands. Within this context, the core protocol of the PCS series has been designed for molecular systems containing approximately 5–15 atoms, with the energetic and structural schemes rooted in explicitly correlated CCSD(T) methods, which have served as the workhorses for this molecular size regime. The approach is then been extended to larger systems, up to about 50 atoms, while maintaining the same methodological philosophy. In this case, energetic contributions have been evaluated using the frozen natural orbital (FNO) approximation within the Coupled Cluster (CC) formalism, whereas geometries have been determined through DFT-based

schemes. In both methodological variants, the computational level has been progressively tailored to the computational cost of each target property and to the characteristics of the examined molecular system.

Finally, with the aim of extending accurate composite methodologies to larger molecular systems at reduced computational cost, post-CCSD(T) correlation effects are accounted for through the FNO approximation. This approach has allowed the computation of high-order correlation contributions for molecular systems beyond the reach of canonical CC methods, significantly reducing computational requirements while maintaining accuracy. This methodology has thus opened the way toward spectroscopic accuracy for molecular sizes that were previously unattainable due to the prohibitive cost of higher-order CC excitations.

Overall, this research demonstrates that the integration of state-of-the-art wave function-based techniques within carefully designed composite frameworks enables the development of predictive quantum chemical models that are both accurate and computationally scalable, effectively reducing the gap between benchmark-level methods and molecular systems spanning different size regimes, with applications ranging from astrochemical to biochemical and technological contexts.

Acknowledgements

These four years of my PhD have certainly been intense and, at times, demanding, but they have also been filled with discovery and wonder. As I wrote in my motivation letter for admission to this doctoral program, I have always been fascinated, since childhood, by the *how* and *why* of things. This field has allowed me to follow that curiosity, and the PhD in *Molecular Sciences for Earth and Space* could not have been a better fit for this purpose.

I would like to express my heartfelt gratitude to the entire MOSES scientific area, to its coordinator Professor Nadia Rega, as well as to the institutions that make this doctoral program possible: the University of Naples Federico II, the Scuola Superiore Meridionale, and the Scuola Normale Superiore di Pisa. I would like to give special thanks to the latter, where I have spent these years of study and growth, which has truly become my second home.

I am also deeply grateful to the STARK research group and all its members for allowing me to carry out this research, which has sparked my interest and passion for the world of benchmarking.

My sincere thanks go to my supervisors, Professor Vincenzo Barone, Professor Nicola Tasinato, and Professor Mihály Kállay.

I would like to thank Professor Barone for the countless conversations in which I often felt small compared to his vast knowledge, and for passing on to me his innate chemical intuition and passion for the subject.

I am deeply grateful to Professor Tasinato for taking on the responsibility of supervising my thesis, for his continuous guidance, and for accompanying me through this important final stage of my doctoral journey. I truly thank him for answering my endless questions and for doing everything possible to support the completion of this work.

Finally, I wish to thank Professor Kállay for the time I spent in Budapest under his supervision, which I truly appreciated and will always treasure. I had the opportunity to learn so much, exploring different research topics, and I am deeply grateful for the support he offered me with his patience and kindness.

For the acknowledgements addressed to my family and friends, I choose to use my native language, Italian, as it allows me to convey my thoughts with greater depth and authenticity.

Un immenso *grazie* alla mia famiglia, che mi ha accompagnato e sostenuto in questo percorso. A mia mamma, la mia fonte di ispirazione, la mia eroina, sempre pronta ad abbracciarmi ed ascoltarmi, spronandomi in ogni momento a non desistere. *Grazie* per essere il mio esempio di umanità, coraggio e caparbietà. A mio papà, che cerchi sempre di farmi vedere il bicchiere mezzo pieno e non mezzo vuoto, che mi insegni a non focalizzarmi solo sui dettagli, ma avere sempre una visione di insieme ed un pensiero laterale non comune. *Grazie* per ogni consiglio, per essere orgoglioso di me e di ogni traguardo raggiunto. Ai miei fratelli, Elena e Marco, che amo con tutto il cuore. Anche se questo percorso mi ha portato lontano da voi, tremendamente assente nella vostra quotidianità, facendomi perdere parte della vostra

splendida infanzia ed inizio adolescenza. Voglio che sappiate che siete con me in ogni momento, che fate parte della mia forza, mancandomi ogni secondo. Spero che questo possa in qualche modo ispirarvi e farvi credere in voi, perché, fidatevi, non c'è frase più vera di "quando un Di Grande si mette in testa di fare qualcosa, può assolutamente riuscirci". Non pretendo di essere uno dei vostri esempi, ma spero almeno un minimo di potervi supportare e sostenere, e farvi sapere che, anche se a distanza, io ci sono sempre per voi, essendo una dei vostri primi fan, e che sono sempre fiera di voi, sempre. Un *grazie* per volermi bene incondizionatamente.

Un *grazie* a tutte le persone che la Normale mi ha fatto incontrare, in particolare a coloro che oggi posso chiamare amici. Jacopo, mio primo tutor, che mi ha accompagnata e seguita durante il mio primo anno di dottorato. Marco, con cui ho condiviso infinite risate in ufficio e nelle nostre pause di aggiornamento. Matteo e Andrea, compagni di avventure nell'esplorazione del mondo inseguendo conferenze lontane.

Infine, un *grazie* a tre persone speciali, per cui sicuramente il termine "amico" è riduttivo.

Sara, il 3 gennaio 2025 è stato davvero un giorno speciale, perché mi ha regalato te. Un'amica preziosa, sempre pronta ad esserci per me, arricchendo la mia vita e di cui ormai non posso più fare a meno. Un *grazie* per ogni risata, per ogni suggerimento e per ogni "de".

Michela, mia compagna di vita. Ormai da quindici anni sei la mia persona, il mio porto sicuro, la mia metà, con cui basta un solo sguardo per capirci. Anche se in questi anni siamo state fisicamente lontane, siamo sempre state vicine, pronte a rincorrerci al telefono per riuscire a sentirci. Un *grazie* perché ci sei sempre.

Alberto, non avrei potuto non menzionarti in questi ringraziamenti, avendo avuto uno dei ruoli più importanti in questo percorso. Mi hai insegnato tanto, sia scientificamente che "nella vita" (ad esempio come fare un sondaggio WhatsApp...), spronandomi sempre a combattere con tutte le forze. Se oggi sono qui lo devo a te, per avermi supportato, sopportato, ascoltato, consolato e amato. Un *grazie* per ogni volta che ci siamo guardati negli occhi, dicendoci di potercela fare, perché oggi ce l'ho fatta, sperando di averti trasmesso almeno un po' della mia sconfinata passione per il benchmarking.

A mia nonna Maria Rosa e alla signora Anna. . .

Contents

List of Figures

List of Tables

Acronyms

(T) perturbative treatment of Triple excitations.

ae all-electron.

fc frozen core.

AO atomic orbital.

ATcT Active Thermochemical Tables.

ATOMIC *Ab initio* Thermochemistry using Optimal-balance Models with Isodesmic Corrections.

B3 B3LYP-D3(BJ).

BO Born–Oppenheimer.

BSIE basis set incompleteness error.

BSSE basis set superposition error.

CABS complementary auxiliary basis set.

CBS complete basis set.

CBS-*X* Complete Basis Set *X*.

CC Coupled Cluster.

ccCA correlation consistent Composite Approach.

CCD Coupled Cluster with Double excitations.

CCSD Coupled Cluster with Single and Double excitations.

CCSD(F12*) explicitly correlated Coupled Cluster with Single and Double excitations with the (F12*) approximation.

CCSD(F12*)(T+) explicitly correlated Coupled Cluster with Single and Double excitations with the F12* approximation, and a perturbative treatment of Triple excitations with the (T+) approach.

CCSD(T) Coupled Cluster with Single and Double excitations, and a perturbative treatment of Triple excitations.

CCSD(T)-F12b explicitly correlated Coupled Cluster with Single and Double excitations with the F12b approximation, and a perturbative treatment of Triple excitations.

CCSD-F12 explicitly correlated Coupled Cluster with Single and Double excitations.

CCSD-F12b explicitly correlated Coupled Cluster with Single and Double excitations with the F12b approximation.

CCSDT Coupled Cluster with Single, Double, and Triple excitations.

CCSDT(Q) Coupled Cluster with Single, Double, and Triple excitations, and a perturbative treatment of Quadruple excitations.

CCSDTQ Coupled Cluster with Single, Double, Triple, and Quadruple excitations.

CCSDTQ(P) Coupled Cluster with Single, Double, Triple, and Quadruple excitations, and a perturbative treatment of Pentuple excitations.

CCSDTQP Coupled Cluster with Single, Double, Triple, Quadruple, and Pentuple excitations.

CGTO contracted Gaussian-type orbital.

ChS Cheap Scheme.

CO correlated orbital.

CS composite scheme.

CV core-valence.

DBOC diagonal Born–Oppenheimer correction.

DF density fitting.

DFT Density Functional Theory.

DLPNO domain-based local pair natural orbital.

EC electronic correlation.

ERI electron repulsion integral.

F12 explicitly correlated.

FIX FIXed amplitudes approximation.

FNO frozen natural orbital.

FPA Focal Point Analysis.

FPD Feller–Peterson–Dixon.

G*n* Gaussian-*n*.

GGA generalized gradient approximation.

GTO Gaussian-type orbital.

HEAT High-accuracy Extrapolated *Ab initio* Thermochemistry.

HF Hartree-Fock.

junChS jun Cheap Scheme.

junChS-F12 explicitly correlated version of the jun Cheap Scheme.

KAW Knizia–Adler–Werner.

KS Kohn-Sham.

LDA local density approximation.

LNO local natural orbital.

LR linear regression.

LSDA local spin density approximation.

MAD mean absolute deviation.

MAX max unsigned error.

meta-GGA meta-generalized gradient approximation.

MO molecular orbital.

MP Møller-Plesset perturbation theory.

MP1 first-order Møller-Plesset perturbation theory.

MP2 second-order Møller-Plesset perturbation theory.

MP2-F12 explicitly correlated second-order Møller-Plesset perturbation theory.

MP3 third-order Møller-Plesset perturbation theory.

MP4 fourth-order Møller-Plesset perturbation theory.

MUE mean unsigned error.

MW microwave.

NAB natural auxiliary basis.

NAF natural auxiliary function.

NO natural orbital.

OVO optimized virtual orbital.

PAH polycyclic aromatic hydrocarbon.

PCS Pisa Composite Scheme.

PES potential energy surface.

PGTO primitive Gaussian-type orbital.

PNO pair-specific natural orbital.

PPL Particle-Particle Ladder.

QC quantum chemical.

Rel scalar relativistic.

revDSD revDSD-PBEP86-D3(BJ).

RHF restricted Hatree-Fock.

RI resolution of identity.

RMSD root-mean-square deviation.

ROHF restricted open-shell Hatree-Fock.

SCF self-consistent field.

SD Slater determinant.

SE semi-experimental.

SO spin-orbit.

SPT simple perturbation theory.

STO Slater-type orbital.

SVECV-f12 Simple Version of an Extrapolated, Core-Valence correlation corrected, CCSD(T)-f12.

TAE total atomization energy.

TM template molecule.

TS templating synthon.

UHF unrestricted Hatree-Fock.

VPT2 second-order vibrational perturbation theory.

W1-F12 explicitly correlated version of the Weizmann-1.

W2-F12 explicitly correlated version of the Weizmann-2.

W4 Weizmann-4.

W*n* Weizmann-*n*.

ZPVE zero-point vibrational energy.

List of Publications

- S. Di Grande, M. Kállay, V. Barone, *One Basis Set to Rule Them All: Toward an Affordable Platinum Standard via Natural Orbital Thresholds*, Journal of Chemical Theory and Computation, **2025**, 21 (21), 11021–11038.
- F. Lazzari, L. Crisci, S. Di Grande, V. Barone, *Cost-Effective Accuracy in Molecular Structures via Smart Databases, Topological Features, and Random Forests*, Journal of Chemical Theory and Computation, **2025**, 21 (20), 10633–10644.
- V. Barone, F. Lazzari, S. Di Grande, *Accurate Structures and Spectroscopic Parameters of CN-Substituted Polycyclic Hydrocarbons at DFT Cost*, Journal of Physical Chemistry A, **2025**, 129 (12), 2876–2886.
- F. Lazzari, S. Di Grande, L. Crisci, M. Mendolicchio, V. Barone, *Molecular Structures with Spectroscopic Accuracy at DFT Cost by the Templating Synthon Approach and the PCS141 Database*, Journal of Chemical Physics, **2025**, 162 (11), 114310.
- V. Barone, F. Lazzari, S. Di Grande, L. Crisci, *Accurate Structures and Rotational Constants of Nitroaromatic Explosives and their Taggants at DFT Cost by Pisa Composite Schemes and Vibrational Perturbation Theory*, Chemical Physics Letters, **2025**, 866, 141956.
- F. Lazzari, L. Uribe, S. Di Grande, L. Crisci, M. Mendolicchio, V. Barone, *Structures and Rotational Constants of Monocyclic Monoterpenes at DFT Cost by Pisa Composite Schemes and Vibrational Perturbation Theory*, Journal of Physical Chemistry A, **2025**, 129 (2), 503–517.
- L. Uribe, S. Di Grande, M. Mendolicchio, N. Tasinato, V. Barone, *Accurate Structure and Spectroscopic Properties of Azulene and Its Derivatives by Means of Pisa Composite Schemes and Vibrational Perturbation Theory to Second Order*, Journal of Physical Chemistry A, **2024**, 128 (49), 10474-10488.
- S. Di Grande[†], F. Lazzari[†], V. Barone, *Accurate Geometries of Large Molecules at DFT Cost by Semiexperimental and Coupled Cluster Templating Fragments*, Journal of Chemical Theory and Computation, **2024**, 20 (20), 9243-9258. [†]The authors contributed equally to the work.
- L. Uribe, F. Lazzari, S. Di Grande, L. Crisci, M. Mendolicchio, V. Barone, *Accurate Structures and Rotational Constants of Bicyclic Monoterpenes at DFT Cost by Means of the Bond-Corrected Pisa Composite Scheme (BPCS)*, Journal of Chemical Physics, **2024**, 160 (1), 014307.

- S. Di Grande, V. Barone, *Toward Accurate Quantum Chemical Methods for Molecules of Increasing Dimension: the New Family of Pisa Composite Schemes*, Journal of Physical Chemistry A, **2024**, 128 (24), 4886–4900.
- L. Uribe, S. Di Grande, L. Crisci, F. Lazzari, M. Mendolicchio, V. Barone, *Accurate Structures and Rotational Constants of Steroid Hormones at DFT Cost: Androsterone, Testosterone, Estrone, β-Estradiol, and Estriol*, Journal of Physical Chemistry A, **2024**, 128 (13), 2629-2642.
- L. Crisci, S. Di Grande, C. Cavallotti, V. Barone, *Toward an Accurate Black-Box Tool for the Kinetics of Gas-Phase Reactions Involving Barrier-Less Elementary Steps*, Journal of Chemical Theory and Computation, **2023**, 19 (21), 7626-7639.
- S. Di Grande, M. Kállay, V. Barone, *Accurate Thermochemistry at Affordable Cost by Means of an Improved Version of the junChS-F12 Model Chemistry*, Journal of Computational Chemistry, **2023**, 44 (27), 2149-2157.
- V. Barone, L. Crisci, S. Di Grande, *Accurate Thermochemical and Kinetic Parameters at Affordable Cost by Means of the Pisa Composite Scheme (PCS)*, Journal of Chemical Theory and Computation, **2023**, 19 (20), 7273-7286.
- V. Barone, L. Uribe, S. Di Grande, F. Lazzari, M. Mendolicchio, *DFT Meets Wave-Function Methods for Accurate Structures and Rotational Constants of Histidine, Tryptophan, and Proline*, Journal of Physical Chemistry A, **2023**, 127 (36), 7534-7543.
- V. Barone, S. Di Grande, F. Lazzari, M. Mendolicchio, *Accurate Structures and Spectroscopic Parameters of Guanine Tautomers in the Gas Phase by the Pisa Conventional and Explicitly Correlated Composite Schemes (PCS and PCS-F12)*, Journal of Physical Chemistry A, **2023**, 127 (32), 6771-6778.
- V. Barone, S. Di Grande, C. Puzzarini, *Toward Accurate yet Effective Computations of Rotational Spectroscopy Parameters for Biomolecule Building Blocks*, Molecules, **2023**, 28 (2), 913.

Chapter 1

Introduction

Over the past decades, quantum chemistry has become an indispensable component of modern chemistry and physics, offering the possibility of simulating molecular systems at the atomic level, thus providing an in-depth understanding of both their static and dynamic properties. Beyond qualitative insights, quantum chemical (QC) methods have become increasingly capable of delivering predictions that can support, complement, or even challenge experimental observations. Such progress has been driven by methodological breakthroughs, the development of more efficient algorithms, and the increasing computational power of modern hardware.

The accurate prediction of thermochemical properties represents one of the central goals of theoretical and computational chemistry. Typical applications include the evaluation of interaction, reaction, and isomerization energies, as well as barrier heights and conformational distributions. Such quantities are fundamental to a wide range of research areas, including combustion and atmospheric chemistry, as well as astro- and biochemistry. With the continuous improvement of computational methods and resources, QC approaches have become increasingly capable of describing reaction pathways, quantifying energy barriers, and determining product distributions. QC methods are therefore instrumental for elucidating transient species that are experimentally challenging to generate or detect, such as free radicals and short-lived intermediates, for modeling the mechanisms of demanding chemical transformations, and for interpreting experimental observations and trends. For small and medium-sized molecules, highly accurate calculations enable the detailed characterization of thermochemical and kinetic parameters, providing the foundation for understanding reaction networks and supporting experimental investigations. In astrochemistry, QC simulations help to elucidate how complex organic molecules form and evolve under extreme conditions of low temperature and pressure, revealing reactive processes on icy grains and in the gas phase that are often inaccessible to experiments. In atmospheric and combustion chemistry, theoretical data are essential for characterizing reactive intermediates and determining degradation mechanisms and corresponding rate coefficients. Moving to larger systems, computational approaches play a crucial role in biochemical and medicinal chemistry, enabling detailed investigations of enzymatic reactions and molecular recognition, and contributing to the computer-aided design of new therapeutic compounds and enzyme mimetics.

In order to deliver quantitative results, QC methods[1–3] must fulfill strict accuracy requirements. The concept of *chemical accuracy* typically refers to an error of about 4.18 kJ mol^{-1} (1 kcal mol^{-1}) relative to reference data. Although seemingly modest, such a tolerance is highly significant: for example, in reactions with multiple competing pathways, an uncertainty of this extent in barrier heights can shift product distributions by an order of magnitude.[4,5] However, the precise definition of how accuracy is quantified is often ambiguous. It may refer to a mean absolute deviation (MAD), to a root-mean-square deviation (RMSD), or to a 95% confidence interval of 4.18 kJ mol^{-1} relative to reference data. The lack of a uniform metric reduces the clarity of the term. Ruscic has analyzed this issue in detail,[6] recommending the use of 95% confidence intervals as a robust measure of uncertainty. Similarly, Peterson, Feller, and Dixon[7] suggested that, to meet the "experimentalist's" notion of chemical accuracy, twice the MAD should remain below 4.18 kJ mol^{-1}. Nevertheless, for properties that benefit from substantial error cancellation, such as interaction energies in non-covalent complexes, tighter thresholds are more appropriate. Values as low as 0.42 kJ mol^{-1} (0.1 kcal mol^{-1}) have been proposed,[8,9] consistent with high-level benchmarks for non-covalently bonded systems, and conformational isomerizations.[10]

Beyond chemical accuracy, stricter targets have been formulated. One of them is *spectroscopic accuracy*, which refers to deviations within 1 kJ mol^{-1} (0.24 kcal mol^{-1}) from experimental data. This accuracy is also essential for the reliable prediction of molecular properties, including rotational-vibrational spectroscopic parameters, central for atmospheric and astrochemical applications.[11–14]

A closely related yet distinct concept is that of *benchmark accuracy*, which also lies within the sub-chemical-accuracy regime but is defined relative to high-level theoretical references rather than experimental data. In this framework, spectroscopic accuracy refers to comparison with experiment, whereas benchmark accuracy considers the theoretical counterpart; together, they delineate the highest target of predictive quantum chemistry.

Focusing on *ab initio* electronic structure methods, which represent the most systematically improvable class of approaches, the current *gold standard* for high-accuracy calculations is Coupled Cluster with Single and Double excitations, and a perturbative treatment of Triple excitations[15] [CCSD(T)], which remains broadly applicable across chemical space. For systems well described by a single Slater determinant (SD) and when large basis sets are employed, CCSD(T) is capable of delivering highly accurate results. However, the computational cost of CCSD(T) formally scales as the seventh power of the system size. Moreover, the correlation energy converges only slowly with respect to basis set size, with errors decreasing roughly inversely with the number of basis functions. While these issues can be partially alleviated through complete basis set (CBS) extrapolation techniques, CCSD(T)/CBS calculations remain feasible only for small molecules,[16] motivating the development of alternative strategies to extend accuracy to larger systems.

Two complementary directions have been particularly successful. The steep computational scaling with molecular size can be mitigated by exploiting the short-range nature of

dynamical correlation through local approximations. Conversely, the slow basis set convergence can be greatly accelerated by explicitly correlated (F12) approaches, in which the wave function directly depends on the interelectronic distance.

Although these methodological advances and increasing computational power have significantly broadened the applicability of *ab initio* methods,[17–25] the cost of canonical high-level computations remains prohibitive for applications aiming at the highest accuracy. To reconcile accuracy and efficiency, composite schemes (CSs) were developed.

First introduced in the late 1980s by Pople, Raghavachari, Curtiss, and co-workers,[2,26] composite methods combine results from a hierarchy of lower-cost computations. CSs exploit systematic additivity schemes to incorporate basis set extrapolations, higher-level correlation, and relativistic corrections, thereby achieving high accuracy at a fraction of the computational cost of the full high-level calculation. These developments laid the foundation for modern composite approaches to thermochemistry,[27] whose accuracy can be systematically improved by including additional effects such as higher-order excitations and core-valence (CV) correlation.

CSs played therefore a key role in establishing chemical accuracy as a practical target. In subsequent developments, increasingly refined variants have pushed this boundary toward stricter accuracy targets, thereby extending the predictive power of quantum chemistry. Indeed, a wide range of composite strategies has since been developed and optimized, targeting both chemical and spectroscopic accuracy. These methods have been applied successfully to increasingly large molecular systems, even reaching the size of buckminsterfullerene (C_{60}).[28] Today, composite *ab initio* methods are among the most accurate tools for investigating chemical processes at the atomistic level. These approaches enable quantum mechanics not only to rationalize the chemistry of complex molecules but also to provide thermochemical and kinetic properties, such as reaction energies and barrier heights, with accuracies that rival the best available experiments. Their broad applicability has consolidated their status as benchmark approaches in theoretical chemistry, with an impact on both fundamental studies and applied research.[29]

Popular CSs comprise several well-established series, such as the Gaussian-n[2,26,30–36] (Gn), the Complete Basis Set X[1,37–43] (CBS-X), the Weizmann-n[11,44–48] (Wn), as well as the correlation consistent Composite Approach[3,49–56] (ccCA), the Feller–Peterson–Dixon[57,58] (FPD) method, the Focal Point Analysis[59–64] (FPA), and the High-accuracy Extrapolated *Ab initio* Thermochemistry[12,65,66] (HEAT) protocol. More recent formulations include the Cheap Scheme[67] (ChS) family, the Simple Version of an Extrapolated, Core-Valence correlation corrected, CCSD(T)-f12[68] (SVECV-f12) method, and the *Ab initio* Thermochemistry using Optimal-balance Models with Isodesmic Corrections[69–71] (ATOMIC) protocol.

CSs can be broadly divided according to their accuracy targets. The Gn, CBS-X, and ccCA families play a central role in main-group thermochemistry at the level of chemical accuracy; the Wn, HEAT, FPA, and FPD methods represent the state-of-the-art in the pursuit of spectroscopic accuracy. Within transition-metal chemistry, specialized ccCA variants

and the FPD methodology have shown considerable potential, though their validation still depends on further benchmarking against experimental data. For heavier elements, where reliable reference data remain limited, advances in spectroscopy are expected to play a key role in guiding the future development of composite approaches.[72]

Composite methodologies aimed at chemical accuracy are typically computationally less demanding and therefore applicable to larger systems. On the other hand, protocols targeting spectroscopic accuracy are generally restricted to diatomic or very small polyatomic molecules, due to their high computational cost. While the Active Thermochemical Tables[73,74] (ATcT) provide accurate information, this level of precision surpasses that of several widely used thermochemical databases, such as the NIST Chemistry WebBook[75] and the CCCBDB.[76]

The transition from chemical to spectroscopic accuracy is generally associated with extending Coupled Cluster[77] (CC) theory beyond CCSD(T). Post-CCSD(T) methods incorporate higher-order excitations, also including the extrapolation to the CBS limit of these terms,[78] whose contributions become particularly significant in systems with non-negligible multireference character,[11,46,79–82] thereby enabling their quantitative treatment and providing a more robust predictive framework. These benefits, however, come at the expense of a steep increase in computational cost, which currently restricts their applicability to relatively small systems, typically containing no more than about eight non-hydrogen atoms.[79,83–85]

The importance of higher-order correlation effects has been highlighted in benchmark studies on hydrocarbons, which revealed systematic trends in the growth of residual errors with molecular size. For well-behaved single-reference systems, such as saturated n-alkanes, deviations in CCSD(T)/CBS atomization energies increase with the chain length, showing that small post-CCSD(T) contributions can accumulate and become non-negligible for larger molecules.[86] Beyond covalent systems, analyses of non-covalent interactions have shown that CCSD(T) benefits from a fortuitous error compensation, yielding results close to Coupled Cluster with Single, Double, and Triple excitations, and a perturbative treatment of Quadruple excitations[87] [CCSDT(Q)]. This balance, however, deteriorates for particular systems such as in extended aromatic π-stacks, where post-CCSD(T) excitations become essential.[88]

In addition to CSs targeting chemical or spectroscopic accuracy, several methods have been developed to achieve an intermediate level of accuracy. These approaches do not include excitations beyond the perturbative triples of CCSD(T), yet they improve upon the chemical-accuracy limit while remaining applicable to larger molecular systems. Their main goal is to extend the size range of molecules that can be studied with high accuracy, particularly those composed of elements regarded as the molecular "bricks of life", namely, from the first three rows of the periodic table. Examples include the SVECV-f12 method and the ChS family, which enables reliable predictions for medium-sized molecules. Within the ChS family, specific variants have been optimized for non-covalent interactions: jun Cheap Scheme[89] (junChS) to capture intermolecular forces, and its extension, explicitly correlated version of the jun Cheap Scheme[90] (junChS-F12), combines this strategy with explicitly correlated

methods.

The reliability of *ab initio* composite methods is rooted in their rigorous foundation on high-level wave function formalisms, among which CCSD(T) stands as the reference point for accuracy. However, the prohibitive scaling of its computational cost, together with considerable data demands, limits its feasibility for larger systems, despite advances in parallel implementations.[91] When employing medium-sized basis sets, conventional CCSD(T) calculations typically reach 20–30 atoms,[92] a limit that can be approximately doubled by adopting natural orbital-based (NO) cost-reduction strategies, such as the frozen natural orbital (FNO) approximation.[93–95]

As molecular size increases, or when transition metals and heavy elements are involved, methods rooted in Density Functional Theory (DFT) often become the practical choice, even though their accuracy cannot be systematically improved as in wave function–based approaches. A wide variety of density functionals has been developed, but their reliability can only be assessed through comparison with experimental data or, in their absence, with accurate *ab initio* wave function results. To this end, error estimates are typically derived from benchmark studies targeting specific molecular classes or reaction types. Consequently, the construction of high-quality reference datasets has become a central activity in theoretical and computational chemistry, underpinning both the validation of density functionals[96–99] and, more recently, the training of machine-learning models.[100,101] Despite the substantial progress achieved in recent years, significant challenges remain, particularly in extending these benchmarks to larger and more diverse chemical spaces.

For main-group chemistry, Karton, Martin, and co-workers,[83,84] for instance, demonstrated that for nearly one thousand energetic data points, including TAEs, isomerization reactions, bond dissociation energies, nucleophilic substitution reactions, and heavy-atom transfer reactions, different families of functionals generally yield comparable RMSD values relative to experiment, with only a few exceptions; however, errors tend to accumulate with molecular size. Several studies have shown that the accuracy of density functional approximations deteriorates with increasing molecular size, leading to progressively larger deviations in computed thermochemical quantities such as enthalpies of formation. For instance, DeYonker and co-workers[3] and Redfern and co-workers[102] independently reported substantial discrepancies between calculated and experimental values for larger hydrocarbons, with errors reaching several tens of kilocalories per mole. Such trends are not unique to specific functionals but reflect a general limitation of DFT-based thermochemistry. Consequently, reaction hierarchy schemes, such as isodesmic or homodesmotic reactions, are often employed to mitigate these systematic errors.[103]

For transition-metal systems, the situation is considerably more complex. Early benchmark studies reported large deviations between calculated and experimental enthalpies of formation,[104] and subsequent assessments confirmed that DFT performance varies significantly with the metal center, ligand environment, and target property. No density functional has proven universally reliable for this class of compounds, and there is no systematic

improvement across Jacob's ladder. Even under favorable conditions, average errors for first- and second-row transition-metal species typically remain well above the uncertainty of high-quality experimental data.[72]

The description of heavy elements poses an even greater challenge, as reliable experimental reference data are scarce, making the assessment of DFT accuracy particularly difficult. Benchmark studies on lanthanide and actinide systems have revealed large, functional-dependent deviations from experimental and high-level theoretical results. Even when scalar relativistic (Rel) and spin-orbit (SO) effects are included, errors remain well above the limits of chemical accuracy, highlighting the persistent difficulties in modeling heavy-element compounds.[72]

Despite the substantial progress achieved, DFT still faces significant challenges in consistently describing systems of varying electronic complexity, particularly those involving transition metals and heavy elements. Further improvements in functional design, the availability of more accurate experimental data, and the adoption of reaction hierarchy schemes may help alleviate some of these shortcomings. Altogether, these limitations highlight the need for alternative, systematically improvable strategies that can deliver reliable accuracy across chemically diverse systems.

In the frame of *ab initio* methods, the locality of dynamical electron correlation can be further exploited, giving rise to local correlation approaches. The extensive development of these approaches, particularly in conjunction with NO–based basis set compression techniques, has broadened the applicability of methods up to the CCSD(T) level.[16,91]

Approaches that exploit the locality of electron correlation are generally classified into two groups: (i) direct local approaches, in which locality is introduced at the algorithmic level by neglecting near-zero contributions or trivial factors, requiring a localized representation of the virtual space, and (ii) piecewise local approaches, in which the molecule is partitioned into subsystems (e.g., fragments or orbital groups) and correlation energies are evaluated separately for each subsystem and subsequently combined to approximate the total result. The first category combines pair-specific natural orbitals (PNOs) with modern local correlation techniques. This line of development was pioneered by Neese and co-workers in the domain-based local pair natural orbital (DLPNO) family of approaches[105–107] and has since been adopted by several other groups.[108–111] The second category is represented by the local natural orbital (LNO) methods, which construct NOs specifically for each localized orbital. This concept was introduced and developed by Kállay and co-workers.[91,112]

The systematically improvable nature of wave function methods can be preserved in local approaches by employing a hierarchy of local correlation settings alongside the conventional atomic orbital (AO) basis set and CC excitation levels. This allows for reliable extrapolation, definition of CSs, and estimation of the model error.[91,113,114]

With recent advances, the CBS limit of CCSD(T) can be approached using LNO and other local correlation methods, such as DLPNO, for systems comprising hundreds of atoms. Local correlated CCSD(T) schemes have also been extended to explicitly correlated variants

and applied successfully to diverse areas, including thermochemistry, isomerization and conformational energies, non-covalent interactions, and transition-metal reactions.[16,91]

Composite *ab initio* methods have also been extended beyond thermochemistry to the prediction of equilibrium geometries and rotational-vibrational spectroscopic properties that require successive energy derivatives.[29,67] As illustrated in Figure 1.1, the computational effort increases with the derivative order: while total energies can often be evaluated at the highest feasible level of theory, the calculation of geometries, vibrational frequencies, and anharmonic contributions rapidly becomes more demanding. In practice, this hierarchy necessitates the use of progressively lower levels of theory for higher-order derivatives, aiming at a trade-off between computational cost and accuracy.

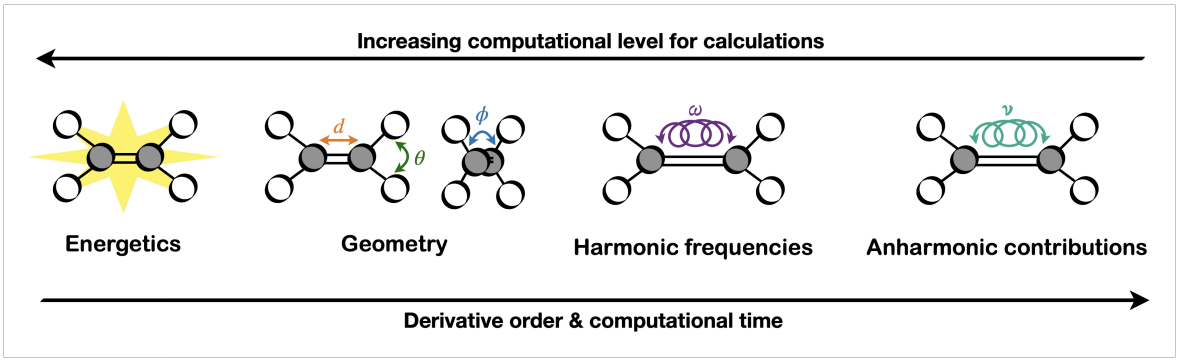


FIGURE 1.1: Hierarchy of molecular properties obtained from successive energy derivatives.

In this PhD thesis, particular attention is devoted to the accurate characterization of equilibrium geometries, which provide the structural foundation for subsequent spectroscopic predictions. Accurate gas phase molecular structures are indeed indispensable for disentangling intrinsic stereoelectronic effects from dynamical and environmental contributions. In this regard, the synergy between high-resolution spectroscopy and accurate QC calculations has proven to be crucial, both for interpreting spectroscopic features in terms of molecular structure[67,115–117] and for enabling the unambiguous identification of chemical species in complex or unknown mixtures.[118,119]

The primary observables in microwave (MW) spectroscopy are the rotational constants, which are intrinsically linked to molecular structure, being inversely proportional to the moments of inertia and therefore determined by atomic masses and geometrical parameters. Since its first demonstration in 1934 by Cleeton and Williams[120] and its rapid development after 1945, MW spectroscopy has enabled systematic studies of gases and liquids with high vapour pressures. More recently, thanks to technological advances in MW spectrometers,[121–124] it has become feasible to investigate solids with low vapor pressures and thermally fragile systems.[125] High-resolution rotational spectra can now be recorded for large and flexible molecules, containing up to about 50 atoms, including hormones,[126–128]

nucleosides,[129] molecular motors,[130] Van der Waals clusters,[131–133] amino acids,[125,134] DNA-scaffold ribosugars,[135] and unstable species.[136–138]

In this framework, QC predictions of accurate rotational constants, and consequently of the underlying geometry, are essential for assigning spectral transitions and hence deriving experimental rotational constants from the analysis of the measured spectra.[139,140] These, in turn, can be exploited to obtain structural information for the investigated species. Within the semi-experimental[141–147] (SE) approach, experimentally derived ground state rotational constants are combined with theoretical vibrational corrections to yield the so-called SE equilibrium rotational constants. A least-squares fit of these constants across a sufficiently large number of isotopologues then provides the corresponding SE equilibrium structure. Although the SE approach affords highly accurate geometrical parameters,[148,149] its applicability remains limited by the need for multiple isotopologues, which are accessible only for a relatively small molecules. The mixed regression method[144] alleviates this limitation by reducing the number of necessary isotopologues, particularly hydrogen substitutions, and has been successfully applied to several medium-sized systems.[150,151] Nevertheless, obtaining accurate SE structures remains challenging for many key molecular fragments found in biologically relevant molecules and pharmaceuticals, such as ring systems in hormones and drugs. In these cases, QC approaches, and particularly CSs, become invaluable, as they offer an optimal balance between accuracy and computational efficiency.

For small semi-rigid molecules, achieving spectroscopic accuracy requires relative errors below 0.1% in rotational constants, corresponding to deviations smaller than 1 mÅ in bond lengths and 0.1° in bond angles, sufficient for unbiased spectroscopic analyses.[152]

To meet these stringent targets, several highly accurate CSs based on the CC ansatz have been developed. Among the most successful are the HEAT[65,153] and W4[11] protocols, together with the series of schemes proposed by Stanton and co-workers,[154,155] available in different variants of increasing theoretical sophistication depending on the desired accuracy. These approaches systematically include basis-set extrapolation, relativistic and CV effects, and post-CCSD(T) correlation contributions, enabling the prediction of equilibrium rotational constants within the 0.1% accuracy threshold. However, the inclusion of all these corrections substantially increases the computational cost, which currently limits the applicability of such CSs to relatively small molecules. To broaden the applicability of composite strategies to medium systems, simplified variants have been designed by reducing the theoretical requirements, balancing accuracy and efficiency while retaining a consistent methodological framework.[67,90,156–163]

For larger or flexible molecules, while experimental data are more difficult to obtain and interpret, computational studies face two main challenges: exploring the potential energy surface (PES) to locate low-energy structures, and accurately characterizing the resulting conformers and rotamers.[117,152,164–167] The steep scaling of accurate methods with system size has motivated the search for reduced-cost alternatives. As previously noted, efficient

strategies based on NOs or local correlation have proven to be effective for electronic energies, but the lack of analytical derivatives limits their applicability to molecular structures. Conversely, standard methods routinely adopted in spectroscopic studies, such as second-order Møller-Plesset perturbation theory[168] (MP2) and hybrid density functionals,[169–172] exhibit systematic errors that preclude fully predictive results without empirical corrections or chemical intuition[173]

The development of the template molecule[145,174] (TM) and linear regression[145] (LR) models, together with their combination (TM-LR),[148,149,175] has allowed computing accurate equilibrium structures for molecules of increasing dimensions. Both approaches are based on the principle that high-quality data obtained for smaller systems or molecular fragments can be used to refine the structures of larger molecules predicted at the DFT level.[176–178] In the TM method, corrections are obtained as differences between accurate and computed geometrical parameters for templating molecules that share similar fragments with the target. In the same way, the LR approach derives corrections from statistical fits to large databases of high-quality reference structures, without requiring explicit fragment assignment. These strategies have been successfully combined in workflows such as nano-LEGO[148,149] and LEGO-brick[179,180] approaches. Recently, a refinement of these models has been introduced through the templating synthon[181] (TS) approach. This method still requires a reference database; however, instead of manually selecting specific fragments, it automatically identifies the most similar analogues for each bond in the target molecule and applies the corresponding corrections. This automation results in a fully black-box protocol, improving usability and extending the applicability to a broader range of chemical systems.

Within this broader context, the overarching goal of this PhD thesis is the accurate characterization of increasingly large molecular systems through the development and refinement of efficient protocols that balance computational cost and accuracy.

The reliable characterization of thermochemical properties requires more than an accurate description of energetics alone. Appropriate levels of theory must be employed at every stage of the computational protocol, encompassing equilibrium geometries as well as harmonic and anharmonic vibrational frequencies. When the objective is benchmark accuracy with respect to theoretical reference data, high-quality energetics may be sufficient, since benchmark protocols generally rely on standardized geometries to avoid biases. Conversely, achieving spectroscopic accuracy relative to experimental observations demands a consistent and accurate treatment of all components of the computational workflow, including energetics, optimized structures, and both harmonic and anharmonic vibrational contributions, since each step cumulatively affects the final observables. Accordingly, this thesis is devoted to the development of predictive computational approaches aiming to approach spectroscopic accuracy, bridging the gap between theoretical predictions and experimental observables. Moving beyond the definition of purely energetic CSs, balanced protocols have been developed that encompass all energetic derivative orders. At each stage, the trade-off between accuracy and computational cost has been carefully optimized, ensuring the highest

possible reliability while maintaining applicability to molecules of increasing size.

The present PhD work builds on this perspective by pursuing two complementary lines of development: (i) the improvement of the existing junChS-F12 approach and (ii) the design of the new Pisa Composite Scheme[176,182,183] (PCS) family. Both efforts are directed toward extending the applicability of composite methods to medium- and large-sized molecules of chemical and biological relevance, while avoiding reliance on local correlation approximations.

A first line of research concerns the refinement of the junChS-F12 protocol,[184] previously developed within the research group. The model is upgraded through the adoption of a recent treatment of triple excitations within explicitly correlated CCSD(T) calculations, which significantly reduces the basis set incompleteness error (BSIE) associated with the perturbative triples contribution and enforces exact size-consistency.[185] Extensive benchmarks confirm that the revised junChS-F12 improves the accuracy of thermochemical predictions for molecules containing elements from the first three rows of the periodic table. The protocol provides two levels of structural accuracy that can be coupled with the energetic scheme. At the first level, geometries are optimized with a double–hybrid density functional, achieving an excellent balance between accuracy and efficiency. At the second level, explicitly correlated CCSD(T) optimizations with a partially augmented basis set are supplemented by explicitly correlated MP2 core-valence corrections. This combination proved to be a reliable strategy for improving structural predictions without requiring CBS extrapolations. Similarly, explicitly correlated CCSD(T) harmonic frequencies with a partially augmented basis set were found to be highly accurate without further corrections. Representative applications to non-covalent interactions, conformational landscapes, and tautomeric equilibria confirmed the robustness and transferability of the improved junChS-F12 scheme.[184,186]

The second line of development concerns the PCS family, introduced in our research group as cost–effective alternatives to traditional high–accuracy protocols such as W*n* or HEAT, whose steep scaling restricts them to very small molecules. In this thesis, PCS variants are designed through a careful selection of models for electronic energies, equilibrium structures, and vibrational frequencies, with the goal of maximizing accuracy at each step while maintaining computational feasibility. For electronic energies, improvements are achieved by implementing reliable CBS extrapolation strategies, refined treatments of core-valence correlation, and optimized basis sets, in particular for third–row atoms. The FNO approximation is introduced and validated, extending the applicability of the energetic scheme to molecules with up to about 50 atoms without resorting to local correlation methods. For equilibrium geometries, accurate predictions are obtained starting from explicitly correlated CCSD(T) optimizations and adding CV contributions evaluated at the MP2 level, avoiding the need for CBS extrapolations. Since analytical gradients are not yet available for the FNO implementation, alternative cost–effective strategies are adopted. In particular, reliable structures for larger molecules are obtained by combining valence contributions from double–hybrid functionals with MP2 core-valence corrections, or by using simplified

one–parameter approximations.[167,187,188] For zero-point vibrational energy (ZPVE) and partition functions, either double–hybrid or hybrid functionals are employed, depending on molecular size. After extensive validation on small benchmark systems, the PCS variants are applied to a representative panel of "biochemical building blocks", as well as systems of current chemical and technological relevance. The results confirm the effectiveness of the PCS family in providing accurate yet affordable predictions for both thermochemical and rotational-vibrational spectroscopic properties.[177,186,189–193]

The final part of this research is devoted to bridging the gap between chemical and spectroscopic accuracy within the PCS family. As discussed above, post-CCSD(T) excitations are often required to reduce the intrinsic error of composite methods below 1 kJ mol^{-1}. However, the computational requirements of these contributions are generally prohibitive, particularly for schemes targeting medium- and large-sized molecules. To address this challenge, this thesis explores the use of the FNO approximation for the evaluation of post-CCSD(T) contributions. Particular attention is given to assessing the impact of (i) the number of correlated orbitals, (ii) the density used for NOs construction, and (iii) the employed basis sets. Extensive benchmarks, including perturbative excitations up to the quintuple level, demonstrate that by combining the FNO approximation with higher–order CC corrections, high accuracy can be retained while substantially reducing computational cost. This strategy paves the way for the reliable characterization of increasingly large molecular systems, which would otherwise remain out of reach for available high–quality CSs.

The thesis is organized as follows.

- **Chapter 1** provides a general introduction to the field, outlining the state-of-the-art in CSs for both energetic and structural properties, and presenting the overall objectives of this work.
- **Chapter 2** summarizes the theoretical background and the QC methods employed throughout this thesis, with particular emphasis on CC theory, F12 methods, and the FNO approximation.
- **Chapter 3** reviews the framework of composite methods, highlighting representative schemes that serve as reference standards for accurate thermochemical and rotational-vibrational spectroscopic predictions.
- **Chapter 4** reports the refinement of the energetic scheme of the junChS-F12 protocol and its extension to the accurate prediction of equilibrium geometries and harmonic frequencies. The performance of the revised model is evaluated through extensive benchmarks and representative applications.
- **Chapter 5** describes the development of the PCS family, a series of cost–effective composite protocols designed to optimize the balance between accuracy and computational

cost across energetic, structural, and vibrational properties. Validation against benchmark and experimental data, along with representative case studies on molecules of increasing size and complexity, demonstrates their reliability and broad applicability.

- **Chapter 6** explores strategies to extend the PCS approach towards spectroscopic accuracy by employing the FNO approximation for post-CCSD(T) contributions. Results show that this strategy achieves substantial computational savings while preserving high accuracy.
- **Chapter 7** outlines the main findings and future perspectives for the further advancement and application of the developed composite schemes.

Chapter 2

Theoretical Background

This Chapter provides an overview of the theoretical foundations of electronic structure theory, providing the theoretical background relevant to the computational approaches employed throughout this dissertation. Rather than presenting an exhaustive treatment, the discussion focuses on the key principles underlying the quantum chemical methods adopted in this PhD thesis. For comprehensive treatments of canonical quantum chemistry, the reader is referred to References 194, 195, 196, 197, 198, while detailed accounts of explicitly correlated methodologies can be found in References 199, 200, 201.

2.1 Foundations of Electronic Structure Methods

2.1.1 The Molecular Hamiltonian

The starting point is the time-dependent non-relativistic Schrödinger equation

$$\hat{\mathcal{H}} \left|\Psi\left(\mathbf{x}, t\right)\right\rangle = i\hbar \frac{\partial}{\partial t} \left|\Psi\left(\mathbf{x}, t\right)\right\rangle \tag{2.1}$$

where $\hbar = h/2\pi$, with h Planck's constant, $\hat{\mathcal{H}}$ the Hamiltonian operator, and $\left|\Psi\left(\mathbf{x}, t\right)\right\rangle$ the wave function which fully describes a state of the system. Equation 2.1 describes the time evolution of a generic N-particles system, with coordinates $\mathbf{x} = \{\mathbf{x}_k\}$, $k = 1, ..., N$. While the wave function has no direct physical interpretation, its squared modulus, $|\Psi(\mathbf{x}, t)|^2$, corresponds to the probability density of locating the particles at specific points in space and time. $\hat{\mathcal{H}}$ is defined as the sum of the kinetic $\hat{\mathcal{T}}$ and the potential $\hat{\mathcal{V}}$ energies,

$$\hat{\mathcal{H}} = \hat{\mathcal{T}} + \hat{\mathcal{V}}. \tag{2.2}$$

In the case of an N-particles system, $\hat{\mathcal{H}}$ can be expressed as

$$\hat{\mathcal{H}} = -\sum_{k=1}^{N} \frac{\hbar^2}{2m_k} \nabla_k^2 + \sum_{k=1}^{N} \sum_{l>k}^{N} \frac{q_k q_l}{4\pi\varepsilon_0 \left|\mathbf{x}_k - \mathbf{x}_l\right|} \tag{2.3}$$

with m_k the mass of the k-th particle, ∇_k^2 the Laplacian operator acting on the coordinates of the k-th particle; the second term on the right-hand side accounts for pairwise Coulomb

interactions, with ε_0 the vacuum electric permittivity, q_k and $\mathbf{x}_k$ representing the charge and the vector collecting the coordinates of the k-th particle, respectively.

As can be observed, Equation 2.3 is not explicitly dependent on time. For this reason, the wave function can be factorized into a part depending exclusively on time, $\Psi_t(t)$, and another one, $\Psi_x(\mathbf{x})$, depending solely on spatial coordinates, as

$$|\Psi(\mathbf{x}_k, t)\rangle = |\Psi_t(t)\rangle\,|\Psi_x(\mathbf{x})\rangle \tag{2.4}$$

where $\Psi_t(t)$ describes the time evolution of the system.

Focusing on the space-dependent part of the wave function, it is possible to obtain the time-independent non-relativistic Schrödinger equation

$$\hat{\mathcal{H}}\,|\Psi_x(\mathbf{x})\rangle = E\,|\Psi_x(\mathbf{x})\rangle \tag{2.5}$$

from which the energy can be obtained as the expectation value of $\hat{\mathcal{H}}$ as

$$E = \left\langle \Psi_x(\mathbf{x}) \middle| \hat{\mathcal{H}} \middle| \Psi_x(\mathbf{x}) \right\rangle. \tag{2.6}$$

A complete description of the static properties of a molecule requires solving Equation 2.5. In the following, attention will be restricted to the time-independent non-relativistic Schrödinger equation, hence from this point onward Ψ will refer to $\Psi_x(\mathbf{x})$.

For a molecular system containing n electrons and M nuclei, the kinetic energy operator can be partitioned into electronic and nuclear contributions, denoted as $\hat{\mathcal{T}}_e$ and $\hat{\mathcal{T}}_n$, respectively. The potential energy consists of three terms: electron–electron ($\hat{\mathcal{V}}_{ee}$), nucleus–nucleus ($\hat{\mathcal{V}}_{nn}$), and electron–nucleus interactions ($\hat{\mathcal{V}}_{en}$). The molecular Hamiltonian can therefore be written in atomic units as

$$\begin{aligned}
\hat{\mathcal{H}} =& \hat{\mathcal{T}} + \hat{\mathcal{V}} \\
=& \hat{\mathcal{T}}_e + \hat{\mathcal{T}}_n + \hat{\mathcal{V}}_{en} + \hat{\mathcal{V}}_{ee} + \hat{\mathcal{V}}_{nn} \\
=& -\sum_{i=1}^{n} \frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M} \frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{n}\sum_{A=1}^{M} \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} + \sum_{i=1}^{n}\sum_{j>i}^{n} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{A=1}^{M}\sum_{B>A}^{M} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|}
\end{aligned} \tag{2.7}$$

where M_A and Z_A denote the mass and atomic number of nucleus A, respectively, while $\mathbf{r}_i$ and $\mathbf{R}_A$ represent the coordinates of the i-th electron and A-th nucleus, respectively.

Among electrostatic interactions, the electron–nucleus term couples the motion of the two types of particles, preventing a straightforward separation of the total wave function into purely electronic and purely nuclear components.

2.1.2 The Born–Oppenheimer Approximation

The Born–Oppenheimer (BO) approximation, a special case of the more general adiabatic approximation, plays a central role in quantum chemistry. It relies on the large mass difference between nuclei and electrons (with nuclei being about 1800 times heavier), which results in much slower nuclear motion. Consequently, to a good approximation, electrons can be regarded as moving in the field of fixed nuclei. Within this framework, the nuclear kinetic energy term is neglected, while the internuclear repulsive interaction is treated as a constant. Under these assumptions, the electronic Hamiltonian, describing the motion of n electrons in the field of M fixed point charges, can be obtained from Equation 2.7 as

$$\hat{\mathcal{H}}_{\text{elec}} = -\sum_{i=1}^{n} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{n}\sum_{A=1}^{M} \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} + \sum_{i=1}^{n}\sum_{j>i}^{n} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \tag{2.8}$$

The solution of the Schrödinger equation involving the electronic Hamiltonian,

$$\hat{\mathcal{H}}_{\text{elec}} |\Psi_{\text{elec}}\rangle = E_{\text{elec}} |\Psi_{\text{elec}}\rangle \tag{2.9}$$

is the electronic wave function,

$$|\Psi_{\text{elec}}\rangle = |\Psi_{\text{elec}}(\{\mathbf{r}_i\}; \{\mathbf{R}_A\})\rangle \tag{2.10}$$

which governs the electronic motion, with an explicit dependence on the electronic coordinates and a parametric one on the nuclear framework, a feature shared by the corresponding electronic energy

$$E_{\text{elec}} = E_{\text{elec}}(\{\mathbf{R}_A\}). \tag{2.11}$$

Parametric dependence implies that, for each arrangement of the nuclei, Ψ_{elec} assumes a different functional form with respect to the electronic coordinates, while the nuclear coordinates do not appear explicitly in Ψ_{elec}. These electronic wave functions are eigenfunctions of $\hat{\mathcal{H}}_{\text{elec}}$ and thus constitute an orthonormal basis set in the electronic Hilbert space.

The total energy for a fixed nuclear configuration must also include the constant contribution of nuclear repulsion

$$E_{\text{tot}} = E_{\text{elec}} + \sum_{A=1}^{M}\sum_{B>A}^{M} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|}. \tag{2.12}$$

Equations 2.8, 2.9, and 2.12 constitute the electronic problem.

Once the electronic problem has been solved, the nuclear motion can be addressed under the same assumptions used in its formulation. Since electrons move much faster than nuclei, it is a reasonable approximation in Equation 2.7 to replace the electronic coordinates with their expectation values, averaged over the electronic wave function. This procedure yields

an effective nuclear Hamiltonian that describes the motion of the nuclei in the mean field generated by the electrons,

$$\begin{aligned}\hat{\mathcal{H}}_{\text{nucl}} &= -\sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 + \left\langle -\sum_{i=1}^{n}\frac{1}{2}\nabla_i^2 - \sum_{i=1}^{n}\sum_{A=1}^{M}\frac{Z_A}{|\mathbf{r}_i-\mathbf{R}_A|} + \sum_{i=1}^{n}\sum_{j>i}^{n}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|}\right\rangle \\ &\quad + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{|\mathbf{R}_A-\mathbf{R}_B|} \\ &= -\sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 + E_{\text{elec}}\left(\{\mathbf{R}_A\}\right) + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{|\mathbf{R}_A-\mathbf{R}_B|} \\ &= -\sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 + E_{\text{tot}}\left(\{\mathbf{R}_A\}\right).\end{aligned} \tag{2.13}$$

The total energy $E_{\text{tot}}(\{\mathbf{R}_A\})$ defines the potential governing nuclear motion, which corresponds to the potential energy surface. Within the BO approximation, the nuclei therefore move on a PES derived from the solution of the electronic problem.
Solutions to a nuclear Schrödinger equation,

$$\hat{\mathcal{H}}_{\text{nucl}}\left|\Psi_{\text{nucl}}\right\rangle = E\left|\Psi_{\text{nucl}}\right\rangle \tag{2.14}$$

describe the translational, vibrational, and rotational motion of a molecule,

$$\left|\Psi_{\text{nucl}}\right\rangle = \left|\Psi_{\text{nucl}}\left(\{\mathbf{R}_A\}\right)\right\rangle \tag{2.15}$$

and E, representing the BO approximation to the total energy, incorporates electronic, vibrational, rotational, and translational degrees of freedom.

The corresponding approximation to the total wave function of Equation 2.5 is

$$\Psi\left(\{\mathbf{r}_i\};\{\mathbf{R}_A\}\right) = \Psi_{\text{elec}}\left(\{\mathbf{r}_i\};\{\mathbf{R}_A\}\right)\Psi_{\text{nucl}}\left(\{\mathbf{R}_A\}\right). \tag{2.16}$$

Assuming Equation 2.9 is solved for each nuclear configuration $\{\mathbf{R}_A\}$, the factorized wave function of Equation 2.16 is substituted and projected onto the Ψ_{elec} basis, with integration carried out over the electronic coordinates. This procedure reveals how the different terms of $\hat{\mathcal{H}}$ act on $\Psi_{\text{nucl}}(\{\mathbf{R}_A\})$ and $\Psi_{\text{elec}}(\{\mathbf{r}_i\};\{\mathbf{R}_A\})$. In particular, the nuclear kinetic energy operator is not diagonal in the Ψ_{elec} basis. The BO approximation consists of neglecting these off-diagonal terms, thereby decoupling electronic and nuclear dynamics. As a consequence, the two types of motion become separable.

In most cases, the BO approximation represents a well-justified assumption, providing an accurate description of molecular systems. However, it may break down when two or more electronic states become nearly degenerate in energy.[202] The residual errors introduced by the approximation are generally small and can be corrected, at least in part, by the diagonal

Born–Oppenheimer correction (DBOC). The DBOC is relatively straightforward to evaluate, as it involves the second derivative of the electronic wave function with respect to the nuclear coordinates, and it is therefore closely related to nuclear gradients and Hessians. Its largest contributions are typically observed in hydrogen-containing molecules, due to the very small nuclear mass of hydrogen. This correction will be addressed in detail in the next Chapter.

When the BO approximation becomes unreliable, non-adiabatic effects can play a dominant role. In such cases, a more appropriate strategy is to explicitly account for the quantum nature of nuclear motion rather than relying solely on adiabatic corrections. Nevertheless, such situations lie beyond the scope of the present PhD thesis. Accordingly, the focus will be restricted to the electronic problem, and the relevant Hamiltonian will from this point on be taken as the electronic Hamiltonian introduced within the BO framework. Unless otherwise specified, $\hat{\mathcal{H}} \equiv \hat{\mathcal{H}}_{\text{elec}}$ and $\Psi \equiv \Psi_{\text{elec}}$.

2.1.3 Pauli Exclusion Principle

Electrons are indistinguishable particles, meaning that any observable derived from the wave function, such as the probability density, must remain invariant under any permutation of the electronic coordinates. For fermionic particles such as electrons, this indistinguishability entails that the total wave function must exhibit antisymmetric behavior under the exchange of any pair of electrons. This requirement is known as the Pauli exclusion principle.

2.1.4 Variational Principle

In quantum chemistry, one of the fundamental principles on which many computational methods are based is the variational principle. It states that, for any normalized trial wave function Φ satisfying the appropriate boundary conditions (typically vanishing at infinity), the expectation value of the Hamiltonian provides an upper bound to the exact ground–state energy. Formally, if $\langle\Phi|\Phi\rangle = 1$, then $\langle\Phi|\hat{\mathcal{H}}|\Phi\rangle \geq E_0$. Equality holds only when the trial function coincides with the exact ground–state wave function, $\Phi = \Phi_0$, for which $\langle\Phi_0|\hat{\mathcal{H}}|\Phi_0\rangle = E_0$.

The central challenge lies in the construction of an appropriate trial wave function. A straightforward possibility for a multi–electron system is to consider a product of single–electron functions, each described by a spin–orbital $|\chi(\mathbf{x})\rangle = |\Psi(\mathbf{r})\rangle\,|\omega(\sigma)\rangle$, where the spin function $\omega(\sigma)$ can take the values α or β. Such a product, however, does not satisfy the Pauli exclusion principle, as it lacks the required antisymmetry under the exchange of two electrons.

A more rigorous choice for a multi-electron wave function is provided by the Slater determinant, which enforces fermionic antisymmetry by construction. By changing sign upon exchange of any two electrons, the SD satisfies the Pauli exclusion principle and thus represents a physically consistent starting point for electronic structure methods. Its formal expression is given by

$$\Psi_{\mathrm{SD}}(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & ... & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & ... & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & ... & \chi_N(\mathbf{x}_N) \end{vmatrix} = |\chi_1 \chi_2 ... \chi_N| \,. \tag{2.17}$$

From this point onward, unless otherwise specified, Ψ_{SD} will be denoted simply as Ψ.

2.2 Hartree-Fock Method

The Hartree-Fock[203] (HF) method provides the foundation of wave function–based quantum chemistry, forming the basis for the development of more advanced electronic structure approaches. In this formulation, the wave function of an n-electron system is approximated by a single Slater determinant, and the electronic energy is obtained as the expectation value of the Hamiltonian,

$$E_{\mathrm{HF}} = \left\langle \Psi \middle| \hat{\mathcal{H}} \middle| \Psi \right\rangle = \sum_{i=1}^{n} h_i + \frac{1}{2} \sum_{i,j=1}^{n} \left(J_{ij} - K_{ij} \right) \tag{2.18}$$

where h_i collects the one-electron contributions, J_{ij} and K_{ij} the two-electrons ones, and the factor $\frac{1}{2}$ allows the double sum to run over all electrons. These terms are defined as

$$\hat{\mathcal{H}} = \sum_{i=1}^{n} \hat{h}_i + \frac{1}{2} \sum_{i,j=1}^{n} \hat{g}_{ij} \tag{2.19}$$

where

$$\hat{h}_i = -\frac{1}{2} \nabla_i^2 - \sum_{A=1}^{M} \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} \tag{2.20}$$

$$\hat{g}_{ij} = \frac{1}{\left|\mathbf{r}_i - \mathbf{r}_j\right|}. \tag{2.21}$$

The integrals in Equation 2.18 can be expressed over the spin-orbitals basis as

$$h_i = \left\langle \chi_i \middle| \hat{h}_i \middle| \chi_i \right\rangle \tag{2.22}$$

$$J_{ij} = \left\langle \chi_i \chi_j \middle| \hat{g}_{ij} \middle| \chi_i \chi_j \right\rangle \tag{2.23}$$

$$K_{ij} = \left\langle \chi_i \chi_j \middle| \hat{g}_{ij} \middle| \chi_j \chi_i \right\rangle \tag{2.24}$$

where the Dirac's notation is employed.

The one-electron operator $\hat{h}_i$ denotes the single-particle part of the Hamiltonian, consisting of the kinetic energy $\hat{\mathcal{T}}_e$ and the nuclear-electron potential $\hat{\mathcal{V}}_{en}$, and it describes the motion of electron i in the field of the nuclei. The two-electron operator $\hat{g}_{ij}$ describes the electron–electron repulsion, from which the terms J_{ij} and K_{ij} derive. The former, referred to as the Coulomb integral, describes the electrostatic interaction between the charge densities $|\chi_i|^2$ and $|\chi_j|^2$. The exchange integral, K_{ij}, arises from the antisymmetric nature of the SD. The three integrals are real quantities with $J_{ij} \geq K_{ij} \geq 0$. In the special case $i = j$, the Coulomb self-interaction term J_{ii} is exactly canceled by the corresponding exchange contribution K_{ii}, thereby eliminating any unphysical self-interaction effects.

Application of HF theory relies on the variational method to determine the set of spin-orbitals that minimize the energy. The minimization is performed under the constraints that the orbitals remain orthogonal and normalized, through the method of Lagrange multipliers. In this context, J_{ij} and K_{ij} integrals can be expressed as single-electron operators, defined by the spin-orbitals themselves

$$\hat{J}_j \left|\chi_i\right\rangle = \left\langle \chi_j \left| \hat{g}_{ij} \right| \chi_j \right\rangle \left|\chi_i\right\rangle \tag{2.25}$$

$$\hat{K}_j \left|\chi_i\right\rangle = \left\langle \chi_j \left| \hat{g}_{ij} \right| \chi_i \right\rangle \left|\chi_j\right\rangle \tag{2.26}$$

which account for the mean-field treatment of two-electron interactions, whereby the influence of all other electrons on a given one is described by an average potential expressed through these one-electron operators.

The electronic energy expressed in Equation 2.18 can be rewritten in terms of one-electron operators as

$$E_{\mathrm{HF}} = \sum_{i=1}^{n} \left\langle \chi_i \left| \hat{h}_i \right| \chi_i \right\rangle + \frac{1}{2} \sum_{i,j=1}^{n} \left(\left\langle \chi_i \left| \hat{J}_j \right| \chi_i \right\rangle - \left\langle \chi_i \left| \hat{K}_j \right| \chi_i \right\rangle \right). \tag{2.27}$$

The Fock operator is defined by combining the two contributions of Equations 2.25 and 2.26 with the core one-electron term, yielding

$$\hat{F} = \hat{h} + \sum_{j=1}^{n} \left(\hat{J}_j - \hat{K}_j \right) \tag{2.28}$$

and the HF orbitals are obtained as solution of the HF equations

$$\hat{F} \left|\chi_i\right\rangle = \left[\hat{h} + \sum_{j=1}^{n} \left(\hat{J}_j - \hat{K}_j \right) \right] \left|\chi_i\right\rangle = \varepsilon_i \left|\chi_i\right\rangle \tag{2.29}$$

where ε_i are the orbital energies. The HF equations (in the basis set representations they are known as the Roothaan–Hall equations[203]) can be solved using iterative schemes, typically of the self-consistent field (SCF) type.

The total electronic energy reported in Equations 2.18 and 2.27 can be reformulated as

$$\begin{aligned} E_{\mathrm{HF}} &= \sum_{i=1}^{n} \varepsilon_i - \frac{1}{2} \sum_{i,j=1}^{n} \left(J_{ij} - K_{ij} \right) \\ \varepsilon_i &= \left\langle \chi_i \left| \hat{F}_i \right| \chi_i \right\rangle = h_i + \sum_{j=1}^{n} \left(J_{ij} - K_{ij} \right) \end{aligned} \tag{2.30}$$

The HF energy is intrinsically approximate, as the electron–electron interaction is treated only in an averaged manner, with each electron moving in the mean field generated by all the other ones, as previously mentioned. This limitation arises from the use of a single SD as the trial wave function, which precludes an explicit treatment of electron correlation.

The formulation of Equations 2.28 and 2.29 holds for closed-shell molecules, that is, systems with an equal number of α and β electrons, where each spatial orbital is doubly occupied by a pair of electrons with opposite spin. This is referred to as restricted Hatree-Fock (RHF). For open-shell molecules, the Fock operator and the corresponding HF equations must be defined separately for each spin component. These are the unrestricted Hatree-Fock (UHF) equations, in which electrons with different spins are allowed to occupy distinct spatial orbitals. Nevertheless, the two sets of spin-dependent orbitals remain coupled through the Coulomb interaction. Open-shell systems may also be described by restricted-type wave functions, where the spatial part of the doubly occupied orbitals is forced to be the same, and this is known as restricted open-shell Hatree-Fock (ROHF). It should be noted that the total RHF and ROHF wave functions are eigenfunctions of the total spin-squared operator $\hat{S}^2$, whereas those obtained within the UHF framework are not. The characteristic orbital occupation patterns of RHF, ROHF, and UHF are illustrated schematically in Figure 2.1.

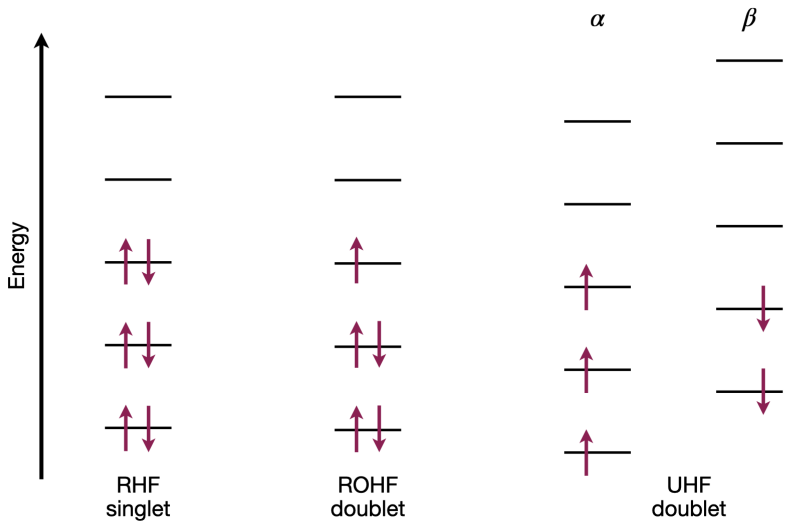


Figure 2.1: Schematic illustration of an RHF singlet, and ROHF and UHF doublet states.

2.2.1 Electron Correlation

The consequence of the mean-field approximation is that each electron experiences only the averaged effect of all the others. As a result, HF theory accounts solely for average electron–electron interactions, while neglecting electronic correlation (EC), leading to an energy estimate that systematically deviates from the exact solution.

EC is defined as the difference between the exact energy and the HF energy

$$E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}}. \tag{2.31}$$

EC can be classified into Coulomb correlation and Fermi correlation, corresponding to interactions between electrons with antiparallel and parallel spins, respectively. Because of Pauli's principle, within the RHF framework, Coulomb correlation can arise both between electrons occupying the same spatial orbital and between those in different orbitals. In contrast, Fermi correlation is restricted to electrons in different orbitals and is therefore typically smaller in magnitude than Coulomb correlation.

EC can also be categorized into *dynamic* and *static* correlation. Dynamic correlation accounts for the instantaneous adjustment of the electronic motion in response to their mutual repulsion, and is particularly relevant for electrons occupying the same orbital. Static correlation, on the other hand, arises in situations involving quasi-degenerate states, where multiple spin-orbitals of similar energy contribute significantly to the electronic structure. Within this framework, UHF can capture a larger portion of static correlation compared to RHF.

Systems dominated by dynamic correlation can be reliably described by a single electronic configuration, usually represented by a single SD derived from the HF wave function. Methods that build upon this description are generally referred to as post-HF approaches, and the corresponding molecular systems are classified as *single-reference*. When static correlation becomes significant, a single determinant no longer provides a sufficient description. In such cases, several electronic configurations of comparable weight contribute to the wave function, requiring a multideterminantal treatment. Systems of this type are defined as *multi-reference*.

In this doctoral work, the focus will be exclusively on single-reference post-HF methods, as they are fully sufficient for the accurate description of the systems under investigation.

2.3 Møller-Plesset Perturbation Theory

The simplest approach to account for EC effects in QC calculations is to apply perturbation theory with the HF wave function taken as the zeroth–order reference. In this framework, EC is introduced as a perturbative correction, represented by $\hat{\mathcal{H}}'$, to the reference HF Hamiltonian $\hat{\mathcal{H}}^{(0)}$. The underlying premise of perturbative methods is that $\hat{\mathcal{H}}'$ is small compared

to $\hat{\mathcal{H}}^{(0)}$, so that corrections can be systematically added to solutions obtained within the independent–particle approximation. This leads to the definition of the total Hamiltonian as

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}^{(0)} + \lambda \hat{\mathcal{H}}' \tag{2.32}$$

where λ is a dimensionless real parameter that controls the magnitude of the perturbation, and the unperturbed Hamiltonian $\hat{\mathcal{H}}^{(0)}$ corresponds to the sum of Fock operators. This formulation constitutes the basis of Møller-Plesset perturbation theory[168] (MP), which can be regarded as a special case of the more general many–body perturbation theory, relying on the Rayleigh–Schrödinger formalism.

Within this formulation, the Schrödinger equation takes the perturbed form

$$\left(\hat{\mathcal{H}}^{(0)} + \lambda \hat{\mathcal{H}}'\right) |\Psi\rangle = E\,|\Psi\rangle\,. \tag{2.33}$$

Since the eigenvalues and eigenfunctions of Equation 2.33 depend continuously on λ, they can be expressed as a Taylor series expansion in powers of the perturbation parameter

$$E(\lambda) = \sum_{k=0}^{\infty} \lambda^k E^{(k)} \tag{2.34}$$

$$\Psi(\lambda) = \sum_{k=0}^{\infty} \lambda^k \left|\Psi^{(k)}\right\rangle \tag{2.35}$$

where each term of the expansion represents a correction of a specific order, and the corresponding wave function components are constructed to be orthogonal to one another.

By substituting Equations 2.34 and 2.35 into Equation 2.33 and collecting terms of equal order in λ yields

$$\begin{aligned}
\lambda^0 :&\quad \hat{\mathcal{H}}^{(0)} \left|\Psi^{(0)}\right\rangle = E^{(0)} \left|\Psi^{(0)}\right\rangle \\
\lambda^1 :&\quad \hat{\mathcal{H}}^{(0)} \left|\Psi^{(1)}\right\rangle + \hat{\mathcal{H}}' \left|\Psi^{(0)}\right\rangle = E^{(0)} \left|\Psi^{(1)}\right\rangle + E^{(1)} \left|\Psi^{(0)}\right\rangle \\
\lambda^2 :&\quad \hat{\mathcal{H}}^{(0)} \left|\Psi^{(2)}\right\rangle + \hat{\mathcal{H}}' \left|\Psi^{(1)}\right\rangle = E^{(0)} \left|\Psi^{(2)}\right\rangle + E^{(1)} \left|\Psi^{(1)}\right\rangle + E^{(2)} \left|\Psi^{(0)}\right\rangle \\
&\ldots \\
\lambda^n :&\quad \hat{\mathcal{H}}^{(0)} \left|\Psi^{(n)}\right\rangle + \hat{\mathcal{H}}' \left|\Psi^{(n-1)}\right\rangle = \sum_{i=0}^{n} E^{(i)} \left|\Psi^{(n-i)}\right\rangle .
\end{aligned} \tag{2.36}$$

These correspond to the zeroth-, first-, second-, and in general the nth-order perturbation equations, respectively. The zeroth-order equation reduces to the Schrödinger equation of the unperturbed system, while the first-order equation contains two unknowns, given by the first-order correction to the energy $E^{(1)}$ and the first-order correction to the wave function $\Psi^{(1)}$. The nth-order energy correction can be calculated by multiplying from left by the

unperturbed wave function and integrating, employing the "turnover rule" $\langle\Psi^{(0)}|\hat{\mathcal{H}}^{(0)}|\Psi^{(i)}\rangle = \langle\Psi^{(i)}|\hat{\mathcal{H}}^{(0)}|\Psi^{(0)}\rangle^*$:

$$\begin{aligned}
\left\langle\Psi^{(0)}\middle|\hat{\mathcal{H}}^{(0)}\middle|\Psi^{(n)}\right\rangle + \left\langle\Psi^{(0)}\middle|\hat{\mathcal{H}}'\middle|\Psi^{(n-1)}\right\rangle &= \sum_{i=0}^{n-1} E^{(i)}\left\langle\Psi^{(0)}\middle|\Psi^{(n-i)}\right\rangle + E^{(n)}\left\langle\Psi^{(0)}\middle|\Psi^{(0)}\right\rangle \\
E^{(0)}\left\langle\Psi^{(n)}\middle|\Psi^{(0)}\right\rangle + \left\langle\Psi^{(0)}\middle|\hat{\mathcal{H}}'\middle|\Psi^{(n-1)}\right\rangle &= E^{(n)}\left\langle\Psi^{(0)}\middle|\Psi^{(0)}\right\rangle \\
E^{(n)} &= \left\langle\Psi^{(0)}\middle|\hat{\mathcal{H}}'\middle|\Psi^{(n-1)}\right\rangle.
\end{aligned} \tag{2.37}$$

Evaluation of successive corrections can be carried out recursively, relying on the knowledge of the preceding terms and the energy of the unperturbed zeroth–order equation.

Upon replacing the unperturbed Hamiltonian with the sum of Fock operators for an n–electron system, the following expressions are obtained

$$\hat{\mathcal{H}}^{(0)} = \sum_{i=1}^{n}\hat{F}_i = \sum_{i=1}^{n}\left[\hat{h}_i + \sum_{j=1}^{n}\left(\hat{J}_{ij} - \hat{K}_{ij}\right)\right] \tag{2.38}$$

$$\hat{\mathcal{H}}' = \hat{\mathcal{H}} - \hat{\mathcal{H}}^{(0)} = \sum_{i=1}^{n}\sum_{j>i}^{n}\hat{g}_{ij} - \sum_{i,j=1}^{n}\left(\hat{J}_{ij} - \hat{K}_{ij}\right). \tag{2.39}$$

The zeroth-order wave function is the HF determinant, and the zeroth-order energy is the sum of orbital energies

$$E^{(0)} = \left\langle\Psi^{(0)}\middle|\hat{\mathcal{H}}^{(0)}\middle|\Psi^{(0)}\right\rangle = \left\langle\Psi^{(0)}\middle|\sum_{i=1}^{n}\hat{F}_i\middle|\Psi^{(0)}\right\rangle = \sum_{i=1}^{n}\varepsilon_i. \tag{2.40}$$

Recall that the orbital energy corresponds to the energy of an electron in the field of all nuclei and already includes its average repulsion from all other electrons (Equations 2.30), which results in a double counting of the electron–electron interaction. The first-order energy correction corresponds to the expectation value of the perturbation operator evaluated over the zeroth-order wave function

$$E^{(1)} = \left\langle\Psi^{(0)}\middle|\hat{\mathcal{H}}'\middle|\Psi^{(0)}\right\rangle = -\frac{1}{2}\sum_{i,j=1}^{n}\left(J_{ij} - K_{ij}\right). \tag{2.41}$$

This correction removes the double counting of electron–electron repulsion present at zeroth order. By comparing Equation 2.41 with the expression for the electronic energy in Equation 2.30, it follows that the sum of $E^{(0)}$ and $E^{(1)}$ reproduces exactly the HF energy. Using the notation $E^{(n)}_{\mathrm{MP}n}$ for the nth–order correction and $E_{\mathrm{MP}n}$ for the electronic energy up to order n, the relations can be written as

$$E_{\mathrm{MP0}} = E^{(0)}_{\mathrm{MP0}} = \sum_{i=1}^{n} \varepsilon_i \tag{2.42}$$

and

$$E_{\mathrm{MP1}} = E^{(0)}_{\mathrm{MP0}} + E^{(1)}_{\mathrm{MP1}} = E_{\mathrm{HF}}. \tag{2.43}$$

Therefore, with this choice of $\hat{\mathcal{H}}^{(0)}$, MP theory does not account for electron correlation until second order.

In the development of perturbation theory, it is assumed that the solutions of the unperturbed problem constitute a complete set. In practice, this would require an infinite number of basis functions, which is clearly not feasible in actual computations. The lowest–energy solution corresponds to the HF wave function, while the higher–energy solutions are given by excited SDs. However, when a finite basis set is employed, only a limited number of excited determinants can be generated, and the expansion of the many–electron wave function must be truncated, typically at a finite order of perturbation. MP methods are therefore denoted as MPn, where n indicates the order at which the expansion is truncated. In practice, the most commonly used levels are MP2, MP3, and MP4.

According to Equations 2.37, the second-order correction to the energy requires the calculation of the expectation value of the perturbation operator $\hat{\mathcal{H}}'$ between the zeroth- and first-order wave functions. The explicit formula for the MP2 energy can be expressed as

$$E_{\mathrm{MP2}} = E^{(0)}_{\mathrm{MP0}} + E^{(1)}_{\mathrm{MP1}} + E^{(2)}_{\mathrm{MP2}} = E_{\mathrm{HF}} - \frac{1}{4} \sum_{i,j=1}^{occ.} \sum_{a,b=1}^{virt.} \frac{\left| \langle \chi_i \chi_j | \chi_a \chi_b \rangle - \langle \chi_i \chi_j | \chi_b \chi_a \rangle \right|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} \tag{2.44}$$

with i, j and a, b denoting occupied and virtual orbitals, respectively. From Equation 2.44 follows that the total electronic energy, once electron correlation is taken into account, is always lower than the corresponding HF energy.

By progressing through higher perturbative orders and incorporating the corresponding excited determinants, it becomes possible to systematically construct all associated energy corrections $E^{(n)}_{\mathrm{MP}n}$.

The main limitation of perturbation methods lies in the assumption that the zeroth–order wave function provides a reasonable approximation to the exact one, meaning that the perturbation operator remains sufficiently small. When the HF reference poorly describes the system, the correction terms become larger, and progressively higher orders must be included to achieve a given accuracy. In extreme cases, where the reference state is inadequate, the convergence of the perturbation series may become so slow or erratic that perturbation methods are no longer applicable.

In practice, it is difficult to determine whether a perturbation expansion is truly convergent. For many systems, the first few terms exhibit an apparently convergent behavior,

but this can be misleading, as the convergence properties depend strongly on the size of the basis set.[204] In the ideal case, the sequence of HF, MP2, MP3, MP4, and higher-order results would display monotonic convergence toward a limiting value, with corrections of decreasing magnitude and consistent sign as the order increases. Unfortunately, such behavior is rarely observed. Even for systems that are well described by a single determinant, oscillations in the energy or other properties with respect to the perturbation order are common. Moreover, since this family of methods is not variational, the computed energy is not guaranteed to be an upper bound to the exact value and can sometimes fall below it. An analysis by Cremer and He[205] showed that smooth convergence of the energy can only be expected for systems with well–separated electron pairs. By contrast, oscillatory behavior arises when this condition is not met. Such cases include systems with lone pairs and/or multiple bonds.

It is important to emphasize that neither monotonic nor oscillatory behavior of the first few terms provides conclusive information about the convergence of the entire perturbation series. In practice, only the lowest orders of perturbation theory can be computed, and it is frequently observed that the HF and MP2 results differ substantially, with MP3 shifting back toward the HF value and MP4 deviating again, as shown in Figure 2.2.

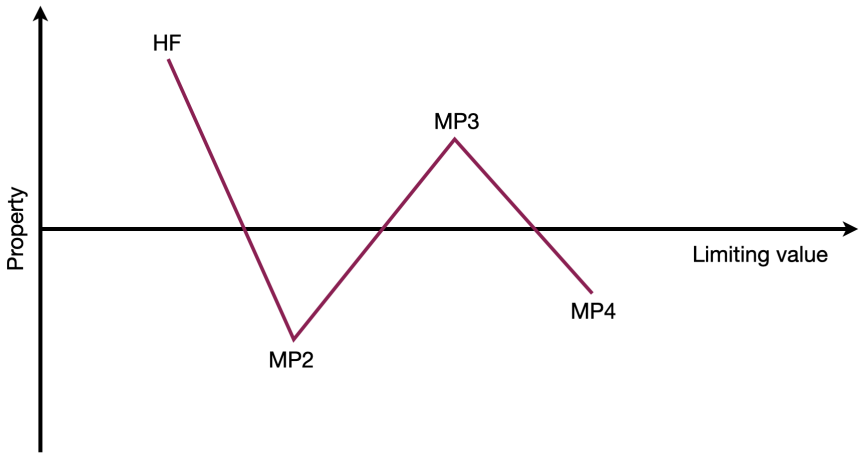


FIGURE 2.2: Typical oscillatory behavior of results obtained with the MP expansion.

For "well-behaved" systems, a good result is often found between the MP3 and MP4 values. MP2 usually overshoots the correlation contribution but nevertheless tends to yield more accurate results than MP3, particularly when medium–sized basis sets are employed. Similarly, because the singles and triples contributions enter the expansion for the first time at fourth order, MP4 often overestimates their effect.

For systems that can be reliably described within a single–reference framework, MP2 remains a computationally inexpensive approach capable of capturing the dominant part

of the electron correlation effects; however, when the reference wave function exhibits significant multi-reference character, a perturbation expansion built on a single determinant generally shows poor or erratic convergence.

2.4 Coupled Cluster Theory

Coupled Cluster theory was originally formulated for the quantum–mechanical treatment of nuclear-physics phenomena. Following its introduction into electronic structure theory, it rapidly evolved into one of the most powerful approaches in quantum chemistry for describing electron correlation and performing high–accuracy computations.

CC theory is founded on the exponential ansatz for the wave function to describe electron correlation. The approach starts from a single SD as the zeroth–order reference state, most commonly obtained from HF theory

$$|\Psi_{\mathrm{CC}}\rangle = e^{\hat{T}} |\Psi_{\mathrm{HF}}\rangle \tag{2.45}$$

where $\hat{T}$ denotes the cluster operator, which generates electronic excitations from the reference determinant. It is expressed as a weighted sum of all possible excitations, and for an n–electron system, it can be written as

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + ... + \hat{T}_n \tag{2.46}$$

in which $\hat{T}_1 + \hat{T}_2 + \dots$ represent the weighted sums of single, double, and higher–order excitations. The $\hat{T}_i$ operator acting on an HF reference wave function Ψ_0 generates all i-th excited SDs, where electrons are promoted from occupied orbitals i, j to virtual ones a, b

$$\begin{aligned} \hat{T}_1\Psi_0 &= \sum_{i=1}^{occ.}\sum_{a=1}^{virt.} t_i^a \Psi_i^a \\ \hat{T}_2\Psi_0 &= \frac{1}{4}\sum_{i,j=1}^{occ.}\sum_{a,b=1}^{virt.} t_{ij}^{ab} \Psi_{ij}^{ab} \\ &\dots \end{aligned} \tag{2.47}$$

with Ψ_i^a and Ψ_{ij}^{ab} denoting singly and doubly excited SDs, respectively.

The corresponding weighting coefficients t are the unknown parameters of the method and are usually referred to as *amplitudes*.

The exponential ansatz of the CC wave function ensures both size-consistency and size-extensivity of the electron–correlation treatment, meaning that energies are additive for non-interacting subsystems and scale linearly with system size.[206,207]

The operator $e^{\hat{T}}$ can be expressed as a power–series expansion

$$e^{\hat{T}} = 1 + \hat{T} + \frac{1}{2!}\hat{T}^2 + \frac{1}{3!}\hat{T}^3 + ... = \sum_{k=0}^{\infty} \frac{1}{k!}\hat{T}^k. \tag{2.48}$$

As follows from Equations 2.46 and 2.48, the exponential operator takes the form

$$\begin{aligned} e^{\hat{T}} = 1 + \hat{T}_1 + \left(\hat{T}_2 + \frac{1}{2}\hat{T}_1^2\right) + \left(\hat{T}_3 + \hat{T}_2\hat{T}_1 + \frac{1}{6}\hat{T}_1^3\right) \\ + \left(\hat{T}_4 + \hat{T}_3\hat{T}_1 + \frac{1}{2}\hat{T}_2^2 + \frac{1}{2}\hat{T}_2\hat{T}_1^2 + \frac{1}{24}\hat{T}_1^4\right) + ... \quad . \end{aligned} \tag{2.49}$$

The first term on the right-hand side corresponds to the reference HF function, while the second accounts for all singly excited states. Doubly excited states arise from the next group of terms, which can be classified as *connected* ($\hat{T}_2$) or *disconnected* ($\hat{T}_1^2$). The following contributions give rise to triply excited states, either as "true" triples ($\hat{T}_3$) or as "product" of lower–order excitations ($\hat{T}_2\hat{T}_1$, $\hat{T}_1^3$). Quadruply excited states are generated analogously, consisting of one connected quadruple and four disconnected contributions. From a physical standpoint, a connected excitation such as $\hat{T}_4$ represents the simultaneous interaction of four electrons, whereas a disconnected term like $\hat{T}_2^2$ describes two independent pairs of interacting electrons.

With the CC wave function in Equation 2.45, the Schrödinger equation becomes

$$\hat{\mathcal{H}} e^{\hat{T}} \Psi_0 = E e^{\hat{T}} \Psi_0. \tag{2.50}$$

The energy is obtained as the expectation value of the CC wave function, with the amplitudes determined through the application of the variational principle

$$\begin{aligned} E_{\mathrm{CC}}^{\mathrm{var}} &= \frac{\left\langle \Psi_{\mathrm{CC}} \middle| \hat{\mathcal{H}} \middle| \Psi_{\mathrm{CC}} \right\rangle}{\langle \Psi_{\mathrm{CC}} | \Psi_{\mathrm{CC}} \rangle} = \frac{\left\langle e^{\hat{T}}\Psi_0 \middle| \hat{\mathcal{H}} \middle| e^{\hat{T}}\Psi_0 \right\rangle}{\left\langle e^{\hat{T}}\Psi_0 \middle| e^{\hat{T}}\Psi_0 \right\rangle} \\ &= \frac{\left\langle \left(1 + \hat{T} + \frac{1}{2}\hat{T}^2 + ... + \frac{1}{n!}\hat{T}^n\right)\Psi_0 \middle| \hat{\mathcal{H}} \middle| \left(1 + \hat{T} + \frac{1}{2}\hat{T}^2 + ... + \frac{1}{n!}\hat{T}^n\right)\Psi_0 \right\rangle}{\left\langle \left(1 + \hat{T} + \frac{1}{2}\hat{T}^2 + ... + \frac{1}{n!}\hat{T}^n\right)\Psi_0 \middle| \left(1 + \hat{T} + \frac{1}{2}\hat{T}^2 + ... + \frac{1}{n!}\hat{T}^n\right)\Psi_0 \right\rangle}. \end{aligned} \tag{2.51}$$

Expanding both the numerator and the denominator in Equation 2.48 generates a series of non-vanishing terms up to the n-th order for an n-electron system. This makes a fully variational formulation of CC theory computationally intractable, but for very small molecules.

The formulation of CC theory proceeds instead through a similarity transformation of the Hamiltonian operator. Starting from the CC Schrödinger equation (Equation 2.50), left-multiplication by $e^{-\hat{T}}$ yields

$$e^{-\hat{T}}\hat{\mathcal{H}}e^{\hat{T}}\Psi_0 = E_{\mathrm{CC}}\Psi_0. \tag{2.52}$$

Analogously to $e^{\hat{T}}$, which acts as an excitation operator on functions to its right, $e^{-\hat{T}}$ serves as a deexcitation operator acting on functions to its left. Projection onto the reference determinant then yields the corresponding energy equation

$$E_{\mathrm{CC}} = \left\langle \Psi_0 \left| e^{-\hat{T}}\hat{\mathcal{H}}e^{\hat{T}} \right| \Psi_0 \right\rangle. \tag{2.53}$$

It is worth noting that Equation 2.53 can be interpreted as the expectation value of a similarity–transformed, non-Hermitian Hamiltonian. In this formulation, the operator $e^{-\hat{T}}$ formally acts as a deexcitation operator on the reference state Ψ_0^*, although such deexcitations cannot be generated in practice. By exploiting the fact that the Hamiltonian contains only one- and two-electron operators, together with Brillouin's theorem, Equation 2.53 can be reformulated as

$$E_{\mathrm{CC}} = E_0 + \frac{1}{4}\sum_{i,j=1}^{occ.}\sum_{a,b=1}^{virt.}\left(t_{ij}^{ab} + t_i^a t_j^b - t_i^b t_j^a\right)\left(\left\langle \chi_i\chi_j|\chi_a\chi_b\right\rangle - \left\langle \chi_i\chi_j|\chi_b\chi_a\right\rangle\right). \tag{2.54}$$

The CC correlation energy is thus fully determined by the amplitudes of single and double excitations together with the two-electron orbital integrals.

The amplitude equations are derived by projecting onto the space of excited determinants, which yields a set of non-linear equations

$$\begin{aligned}
\left\langle \Psi_m^e \left| e^{-\hat{T}}\hat{\mathcal{H}}e^{\hat{T}} \right| \Psi_0 \right\rangle &= 0 \\
\left\langle \Psi_{mn}^{ef} \left| e^{-\hat{T}}\hat{\mathcal{H}}e^{\hat{T}} \right| \Psi_0 \right\rangle &= 0 \\
\left\langle \Psi_{mnl}^{efg} \left| e^{-\hat{T}}\hat{\mathcal{H}}e^{\hat{T}} \right| \Psi_0 \right\rangle &= 0 \\
&\cdots
\end{aligned} \tag{2.55}$$

The action of the deexcitation operator $e^{-\hat{T}}$ on Ψ_m^{e*} generates not only the reference wave function but also singly excited contributions

$$\begin{aligned}
\left\langle \Psi_m^e \left| e^{-\hat{T}}\hat{\mathcal{H}}e^{\hat{T}} \right| \Psi_0 \right\rangle &= 0 \\
\left\langle \Psi_m^e \left(1 - \hat{T}_1\right) \left| \hat{\mathcal{H}} \right| \left[1 + \hat{T}_1 + \left(\hat{T}_2 + \frac{1}{2}\hat{T}_1^2\right) + \left(\hat{T}_3 + \hat{T}_2\hat{T}_1 + \frac{1}{6}\hat{T}_1^3\right)\right] \Psi_0 \right\rangle &= 0.
\end{aligned} \tag{2.56}$$

Taking into account the orthogonality of the SDs and the fact that the Hamiltonian includes only one- and two-electron operators, the expansion reduces to the indicated contributions.

Terms involving the reference and singly excited states vanish as a consequence of Brillouin's theorem, while the remaining ones give rise to a coupled set of equations with single, double, and triple excitation amplitudes as variables. Similarly, the action of $e^{-\hat{T}}$ on Ψ^{ef*}_{mn} produces not only the doubly excited states but also contributions from the reference and singly excited states

$$\begin{aligned} \left\langle \Psi^{ef}_{mn} \left| e^{-\hat{T}} \hat{\mathcal{H}} e^{\hat{T}} \right| \Psi_0 \right\rangle = 0 \\ \left\langle \Psi^{ef}_{mn} \left(1 - \hat{T}_1 - \hat{T}_2 + \frac{1}{2}\hat{T}_1^2\right) \left| \hat{\mathcal{H}} \right| \begin{bmatrix} 1 + \hat{T}_1 + \left(\hat{T}_2 + \frac{1}{2}\hat{T}_1^2\right) + \left(\hat{T}_3 + \hat{T}_2\hat{T}_1 + \frac{1}{6}\hat{T}_1^3\right) \\ + \left(\hat{T}_4 + \hat{T}_3\hat{T}_1 + \frac{1}{2}\hat{T}_2^2 + \frac{1}{2}\hat{T}_2\hat{T}_1^2 + \frac{1}{24}\hat{T}_1^4\right) \end{bmatrix} \Psi_0 \right\rangle = 0. \end{aligned} \tag{2.57}$$

Equation 2.57 also contains quadruple amplitudes, together with additional coupling terms involving lower–order amplitudes. Further relations among amplitudes can be derived by projecting onto triple, quadruple, and higher excited determinants.

If the cluster operator $\hat{T}$ includes all excitation operators up to $\hat{T}_n$, the full set of excited determinants is generated. However, such a treatment is computationally feasible only for the smallest systems. In practice, $\hat{T}$ must be truncated at a given excitation level. As a result, several terms in the amplitude equations vanish, the corresponding amplitudes are obtained only approximately, and the resulting correlation energy is likewise approximate. The severity of the approximation depends on the excitation level retained in $\hat{T}$. Inclusion of only the $\hat{T}_1$ operator yields no improvement over the HF result, since the matrix elements between the HF reference and singly excited states vanish. Consequently, the lowest meaningful level of approximation corresponds to setting $\hat{T} = \hat{T}_2$.

The most common truncated Coupled Cluster models are Coupled Cluster with Double excitations[207,208] (CCD), Coupled Cluster with Single and Double excitations[209] (CCSD), Coupled Cluster with Single, Double, and Triple excitations[210,211] (CCSDT), Coupled Cluster with Single, Double, Triple, and Quadruple excitations[212,213] (CCSDTQ), and Coupled Cluster with Single, Double, Triple, Quadruple, and Pentuple excitations[214] (CCSDTQP), which are summarized in Table 2.1, together with the corresponding computational scaling properties, where N denotes the dimension of the system size.

Table 2.1: Overview of CC models arising from the truncation of the cluster operator $\hat{T}$ to specific excitation levels.[139]

Approximation	Definition of $\hat{T}$	Computational cost
CCD	$\hat{T} = \hat{T}_2$	N^6
CCSD	$\hat{T} = \hat{T}_1 + \hat{T}_2$	N^6
CCSDT	$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3$	N^8
CCSDTQ	$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \hat{T}_4$	N^{10}
CCSDTQP	$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \hat{T}_4 + \hat{T}_5$	N^{12}

The unfavorable scaling of truncated CC schemes can be alleviated by introducing suitable approximations, most commonly through an approximate treatment of the highest excitation contributions. Approximative CC methods can be classified into two basic categories: iterative and non-iterative methods.

In this PhD thesis, attention is devoted to non-iterative corrections of the highest excitations within truncated CC methods, formulated based on MP perturbation theory. Within this category, the most successful method is Coupled Cluster with Single and Double excitations, and a perturbative treatment of Triple excitations, often regarded as the "gold standard" of quantum chemistry. In this approach, the triples contribution is evaluated using the MP4 formula, but with CCSD amplitudes replacing the perturbative coefficients of the wave function corrections, together with an additional term derived from fifth–order perturbation theory that accounts for the coupling between singles and triples. While the formal scaling of CCSD(T) remains steep (N^6 for the iterative part and N^7 for the non-iterative triples), this method represents the best compromise between computational feasibility and accuracy.

This strategy can be generalized to arbitrary excitation levels, leading to methods denoted as CC$(n-1)(n)$, such as Coupled Cluster with Single, Double, and Triple excitations, and a perturbative treatment of Quadruple excitations [CCSDT(Q)] or Coupled Cluster with Single, Double, Triple, and Quadruple excitations, and a perturbative treatment of Pentuple excitations[215] [CCSDTQ(P)].

Further refinements of the CC$(n-1)(n)$ family of methods, which will be employed in the following Chapters, include the CC$(n-1)(n)_\Lambda$,[215–219] CC$(n-1)(n)$/A,[220] and CC$(n-1)(n)$/B[220] variants. Although somewhat more computationally demanding, these approaches generally yield improved accuracy.

An important tool to assess the reliability of CC wave functions is the T_1 diagnostic, which provides a measure of the single-reference character of the system. In general, the magnitude of the Singles amplitudes (t_1) reflects the adequacy of the HF determinant as a reference state. For a system with n electrons, the T_1 diagnostic can be defined as

$$T_1 = \frac{1}{\sqrt{n}} |t_1| . \tag{2.58}$$

As a practical guideline, values of $T_1 < 0.02$ indicate that CCSD(T) is expected to provide good results. Conversely, values above this threshold suggest that the reference wave function exhibits multi-determinant character, in which case multi-reference approaches should be considered. It should be noted, however, that the T_1 diagnostic is not entirely independent of system size, and its interpretation must therefore be made with care.

2.5 Explicitly Correlated Methods

One of the major challenges in correlated electronic structure methods is the notoriously slow convergence of the correlation energy with respect to basis set size. The underlying reason

is that conventional approaches expand the wave function as a linear combination of SDs. Such ansatz fails to properly describe short-range correlation effects, namely, the behavior of the electronic wave function when two electrons approach each other. In particular, the cusp condition, first rigorously formulated by Kato,[221] requires that

$$\left.\frac{\partial \Psi}{\partial r_{12}}\right|_{r_{12}=0} = \frac{1}{2}\Psi\left(r_{12}=0\right) \tag{2.59}$$

where r_{12} denotes the interelectronic distance. This condition is not satisfied by expansions in one–electron orbitals, and as a consequence, very large basis sets, often combined with extrapolation techniques, are needed to approach convergence.

The importance of explicitly accounting for the interelectronic distance was already recognized by Hylleraas[222,223] in his pioneering studies of the He atom, where the inclusion of r_{12} terms in the wave function ansatz dramatically accelerated convergence. However, extending this idea beyond two-electron systems was long considered impractical, since it requires the evaluation of complicated many-electron integrals. A systematic framework was provided by Kutzelnigg,[224] who demonstrated that orbital expansions and explicitly correlated terms should not be regarded as mutually exclusive alternatives. Instead, their combination is particularly powerful: orbital products efficiently recover the medium- and long-range correlation, while the short-range behavior at the electron–electron cusp is best described by functions explicitly depending on r_{12}. Building on this insight, Kutzelnigg, Klopper, and Termath[225–230] developed the R12 methods, whose success rests on approximations that avoid the need for three- and higher-electron integrals. Nevertheless, the original formulations still suffered from technical limitations, such as the requirement for large and unconventional basis sets, which hindered their widespread use.

In the last two decades, explicit correlation has gained renewed attention thanks to major methodological breakthroughs. In particular, the introduction of improved correlation factors into the wave function ansatz enabled accurate and efficient treatments of short-range correlation. Modern explicitly correlated (F12) methods,[231–242] unlike the early linear correlation factor r_{12}, which is unphysical as $r_{12} \to \infty$ and performs poorly for larger systems, employ short-range Slater-type correlation factors of the form $F_{12} = -\frac{1}{\gamma}e^{-\gamma r_{12}}$, with γ geminal coefficient, which ensure improved basis set convergence and numerical stability, while maintaining physical consistency and broad applicability. The parameter γ can take different values, with its optimal choice depending primarily on the basis set size. Benchmark studies have shown that values in the range 0.8–1.4 are generally appropriate, with $\gamma = 1.0$ representing a reliable choice in most applications. In general, larger basis sets benefit from correspondingly higher values of γ.[234,243]

The general form of the wave function applicable to both standard and explicitly correlated second-order Møller-Plesset perturbation theory (MP2-F12) and explicitly correlated Coupled Cluster with Single and Double excitations (CCSD-F12) methods, is given by

$$\Psi_{\text{MP2(-F12)}} = \left(1 + \hat{T}_2\right)\Psi_0 \tag{2.60}$$

$$\Psi_{\text{CCSD(-F12)}} = e^{\hat{T}_1+\hat{T}_2}\Psi_0 \tag{2.61}$$

with Ψ_0 HF reference wave function. In explicitly correlated methods, the double excitation operator $\hat{T}_2$ is augmented by an additional term $\hat{T}_2'$ containing the F12 amplitudes. Since $\hat{T}_2'$ describes the correlation of electron pairs, the single excitation operator $\hat{T}_1$ in Equation 2.61 remains unchanged. The $\hat{T}_1$ and $\hat{T}_2$ operators are defined in Equations 2.47, while the explicitly correlated pair functions are represented as excitations into the complete virtual orbital space (c.v.o.s.), contracted with the integrals over the correlation factor

$$\hat{T}_2'\Psi_0 = \frac{1}{4}\sum_{i,j=1}^{occ.}\sum_{a,b=1}^{virt.}\left(\frac{1}{2}\sum_{\alpha,\beta=1}^{c.v.o.s.}\sum_{k,l=1}^{occ.} c_{ij}^{kl}\left\langle\chi_\alpha\chi_\beta\left|\hat{Q}_{12}F_{12}\right|\chi_k\chi_l\right\rangle\right)\Psi_{ij}^{ab} \tag{2.62}$$

with i, j, k, l and a, b denoting occupied and virtual orbitals, whereas α, β span the complete virtual orbital space. The F12 amplitudes, c_{ij}^{kl}, are constructed through the action of the projector $\hat{Q}_{12}$, which enforces orthogonality between the explicitly correlated functions and conventional SDs, combined with the correlation factor F_{12}. From Equation 2.62 follows that explicitly correlated terms account solely for pair-correlation effects. As a result, improvements in basis set convergence are restricted to these contributions. Given that pair correlation represents the major component of the total correlation energy, this enhancement remains highly relevant.

To reduce the computational cost associated with electron repulsion integrals (ERIs), several strategies have been developed. A widely adopted approach is resolution of identity (RI), which approximates three- and four-electron integrals in terms of one- and two-electron integrals. This approximation becomes exact in the limit of an RI infinite basis set; however, for any finite basis, the representation remains approximate. Early formulations of RI employed the orbital basis itself as the expansion space, but this choice turned out to be inefficient, as orbital bases are optimized primarily for the occupied subspace, while the virtual space overlaps significantly with the orthogonal complement. To overcome this limitation, auxiliary bases tailored for the RI approximation were introduced. The most common is the complementary auxiliary basis set (CABS), constructed to span the orthogonal complement of the orbital basis, and the RI basis is defined as the union of orbital and auxiliary functions.

Closely related to RI, the density fitting (DF) approximation has become a standard tool for reducing the cost of computing and processing ERIs. In DF, electron–repulsion integrals are expressed as Coulomb integrals of generalized electron densities formed by products of basis functions. These densities are then expanded in an auxiliary basis, typically composed of atom-centered Gaussian functions. As in RI, DF requires an additional basis set: the

auxiliary functions share the same general form as the orbital basis but employ different exponents and are usually largely uncontracted.

Further refinements have been proposed to accelerate DF-based and F12 calculations. Among these are the natural auxiliary functions[244] (NAFs), and the more recent natural auxiliary basis[95] (NAB), designed to reduce the size of the auxiliary basis needed for the expansion of F12 geminals.

A wide range of ansatzes and approximations have been developed for both MP2-F12 and CCSD-F12 methods, offering different balances between accuracy and computational efficiency.[234,235,239,241]

For explicitly correlatedMP2 and CCSD calculations, the diagonal FIXed amplitudes approximation[234] (FIX) is generally recommended, as it ensures both size consistency and orbital invariance. In this approximation, the geminal amplitudes are kept fixed rather than optimized, providing a simple yet accurate treatment of the explicit correlation. The errors associated with this approximation depend on the system and the property under investigation, but are typically small. In addition, the method retains the formal N^5 scaling of conventional MP2 and exhibits favorable numerical stability. Importantly, the use of FIX also eliminates the so-called geminal basis set superposition error (BSSE), which arises in the computation of interaction energies. This artifact originates from the incompleteness of the generating basis, restricted to the occupied orbital space, and cannot be corrected by standard counterpoise techniques.

In the case of CCSD-F12, the inclusion of the F12 contributions in the amplitude equations entails a substantial additional computational cost. To address this issue, several simplified variants have been proposed, each based on a rigorous selection of the leading-order terms, including the geminal contributions. Among the most widely employed ones are CCSD(F12),[245,246] CCSD-F12a,[235,239] CCSD-F12b,[235,239] CCSD(2)$_{\overline{F12}}$,[247,248] and CCSD(F12*).[241] For the purposes of this doctoral thesis, it is sufficient to emphasize that CCSD-F12b and CCSD(F12*) differ only in a limited number of terms, while both retain a computational cost essentially comparable to that of a conventional CCSD calculation. CCSD-F12b offers a modest gain in efficiency, whereas CCSD(F12*) constitutes a more rigorous approximation to the full CCSD-F12 formulation. In practice, however, both approaches yield comparable results.

With regard to the perturbative treatment of Triple excitations, an explicitly correlated ansatz has been implemented,[249,250] although it is very expensive. To accelerate the convergence of the (T) energy contribution with respect to the basis set size, a commonly adopted strategy consists of applying a scaling procedure. In particular, Knizia and co-workers[239] proposed to rescale the conventional (T) correction by the ratio between the MP2-F12 and MP2 correlation energies

$$E_{(\mathrm{T}^*)} = \frac{E_{\mathrm{MP2\text{-}F12}}}{E_{\mathrm{MP2}}} E_{(\mathrm{T})}. \tag{2.63}$$

This approach is based on the assumption that explicit correlation affects the triples contribution in the same way as the MP2 correlation energy. Although this correction is essentially cost-free and generally effective, it is not size-consistent, which can become a serious limitation as the system size increases.

Kállay and co-workers[185] proposed a similar heuristic approach, which preserves size-consistency and evaluates the effect of triples as

$$E_{(\mathrm{T}+)} = \sum_{i=1}^{occ.} \frac{\delta E_i^{\mathrm{MP2\text{-}F12}}}{\delta E_i^{\mathrm{MP2}}} \delta E_i^{(\mathrm{T})} \tag{2.64}$$

with i running over the occupied orbitals. Rather than scaling the entire MP2 correlation energy and (T) correction, this approach assumes that the contributions of individual orbitals, δE_i, to these quantities exhibit a similar dependence on the basis set size. Accordingly, each orbital contribution to the (T) correction is rescaled by the ratio of the corresponding second-order terms. Size-consistency is guaranteed by decomposing the correlation energy into orbital contributions and scaling them separately.

2.6 Density Functional Theory

Wave function–based methods rely on the many-electron wave function Ψ, which depends on $4N$ variables, where N is the number of electrons ($3N$ spatial coordinates and N spin coordinates). An alternative approach is provided by Density Functional Theory, which describes the system in terms of the electron density $\rho(\mathbf{r})$, a function of only three spatial variables. This dramatic reduction in complexity makes DFT particularly attractive for the study of larger molecules and extended systems.

The electron density can be formally obtained from the many-electron wave function as

$$\rho(\mathbf{r}) = N \int \ldots \int |\Psi(\mathbf{x}_1, .., \mathbf{x}_N)|^2 \, d\mathbf{x}_1 d\mathbf{x}_2 \ldots d\mathbf{x}_N \tag{2.65}$$

which is a non-negative function, vanishes at infinity, and integrates over the entire space to the total number of electrons N,

$$\int \rho(\mathbf{r}) \, d\mathbf{r} = N. \tag{2.66}$$

Moreover, the cusps of $\rho(\mathbf{r})$ coincide with the nuclear positions and their heights reflect the nuclear charges.

The total energy of a molecular system can be expressed as a functional of the electron density

$$E[\rho(\mathbf{r})] = V_{nn} + E_e[\rho(\mathbf{r})] = V_{nn} + V_{en}[\rho(\mathbf{r})] + V_{ee}[\rho(\mathbf{r}, \mathbf{r}')] + T[\rho(\mathbf{r})] \tag{2.67}$$

where V_{nn} is the nucleus–nucleus interaction and $E_e[\rho(\mathbf{r})]$ is the electronic energy. The latter can be decomposed into the electron–nucleus interaction $V_{en}[\rho(\mathbf{r})]$, the electron–electron repulsion $V_{ee}[\rho(\mathbf{r},\mathbf{r}')]$, and the electronic kinetic energy $T[\rho(\mathbf{r})]$. Furthermore, the $V_{ee}[\rho(\mathbf{r},\mathbf{r}')]$ term can be divided into Coulomb and exchange parts, $J_e[\rho(\mathbf{r},\mathbf{r}')]$ and $K_e[\rho(\mathbf{r},\mathbf{r}')]$, implicitly including correlation energy in all the terms. The $V_{en}[\rho(\mathbf{r})]$ and $J_e[\rho(\mathbf{r},\mathbf{r}')]$ functionals are given by their classical expression

$$V_{en}[\rho(\mathbf{r})] = -\sum_{A=1}^{M} Z_A \int \frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{R}_A|} d\mathbf{r} \tag{2.68}$$

and

$$J_e[\rho(\mathbf{r},\mathbf{r}')] = \frac{1}{2}\int\int \frac{\rho(\mathbf{r})\,\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \tag{2.69}$$

with M number of nuclei, Z_A nuclear charges, and $\mathbf{R}_A$ their positions. The prefactor $\frac{1}{2}$ in $J_e[\rho(\mathbf{r},\mathbf{r}')]$ corrects for double counting, allowing integration over the entire space for both variables.

It can be demonstrated that the ground-state electron density uniquely defines the system: from it, knowledge of the number of electrons and $V_{en}[\rho(\mathbf{r})]$, the so-called *external* potential, can be determined, which fully specify the Hamiltonian. Once established, the Hamiltonian unambiguously determines the total energy, the many-electron wave function, and all related observables. Consequently, there exists a one-to-one correspondence between the electron density and the energy of the system.

The connection between the electron density ρ and the total energy of the system was first established by the Hohenberg–Kohn theorems.[251] The first theorem guarantees the existence of a universal functional of the electron density that yields the ground-state energy, while the second states that this energy can be obtained variationally, i.e., for any trial density $\tilde{\rho}(\mathbf{r})$ the relation $E[\tilde{\rho}(\mathbf{r})] \geq E[\rho(\mathbf{r})]$ is satisfied. In addition, an unrestricted formulation of DFT can be introduced by decomposing the total density into spin components, $\rho(\mathbf{r}) = \rho_\alpha(\mathbf{r}) + \rho_\beta(\mathbf{r})$, corresponding to α and β electrons, respectively.

While the existence and universality of the functional $E[\rho(\mathbf{r})]$ are guaranteed, its explicit form is unknown. The difficulty arises because the functionals $T[\rho(\mathbf{r})]$ and $K_e[\rho(\mathbf{r},\mathbf{r}')]$ cannot be expressed directly in terms of the density $\rho(\mathbf{r})$. Knowledge of these contributions in closed form would render DFT an exact theory for the ground-state energy.

Early attempts to construct approximate functionals for the kinetic and exchange energies were based on the uniform electron gas, leading to the orbital-free Thomas–Fermi and Thomas–Fermi–Dirac models. While the uniform electron gas approximation can provide a reasonable description of valence electrons in certain metallic or periodic systems, it performs poorly for atoms and molecules. From a chemical perspective, a fundamental shortcoming is that these models do not predict molecular bonding: in such a framework, molecules simply do not exist. The inclusion of gradient corrections partially improves the

Thomas–Fermi description, allowing bonding to be reproduced; however, the absence of sufficient error cancellation renders this strategy unsuitable for developing DFT models with accuracy comparable to that of wave function–based methods.

To circumvent this limitation, Kohn and Sham[252] proposed an ingenious construction based on a fictitious system of non-interacting electrons moving in an effective external potential $V_{en}[\rho(\mathbf{r})]$, where the density of this auxiliary system is constrained to coincide with the true electron density of the interacting one.

The main limitation of orbital-free models lies in the poor representation of the kinetic energy. The Kohn-Sham (KS) formalism overcomes this issue by partitioning the kinetic energy functional into two parts: a dominant contribution that can be evaluated exactly and a small correction term. This approach requires the reintroduction of orbitals, thereby increasing the dimensionality of the problem from three to $3N$ variables and treating electron correlation as a separate contribution. The KS framework is closely related to the HF method, with identical expressions for the kinetic, electron–nuclear, and Coulomb electron–electron energies.

As a consequence of the Hohenberg–Kohn theorems, the external potential $V_{en}[\rho(\mathbf{r})]$ is well defined, even if its exact form is unknown, and in the non-interacting electron model, the Hamiltonian reduces to a sum of one-electron operators

$$\hat{\mathcal{H}}_s = \sum_{i=1}^{n} \left[-\frac{1}{2}\nabla_i^2 + V_{en}[\rho(\mathbf{r})] \right] = \sum_{i=1}^{n} \hat{h}_i^{\mathrm{KS}}. \tag{2.70}$$

The ground-state of this non-interacting system can therefore be represented as a single Slater determinant of the lowest-energy KS spin-orbitals χ_i^{KS}, whose spatial parts are eigenfunctions of the one-electron KS operator $\hat{h}_i^{\mathrm{KS}}$. In the KS formalism, the kinetic energy functional $T_s[\rho(\mathbf{r})]$ is also calculated under the assumption of non-interacting electrons (in the same sense that HF orbitals in wave mechanics describe non-interacting electrons), as

$$T_s[\rho(\mathbf{r})] = -\sum_{i=1}^{n} \left\langle \chi_i^{\mathrm{KS}} \left| \frac{1}{2}\nabla_i^2 \right| \chi_i^{\mathrm{KS}} \right\rangle. \tag{2.71}$$

In reality, the electrons are interacting, and Equation 2.71 does not provide the total kinetic energy. However, just as HF theory provides 99% of the correct answer, the difference between the exact kinetic energy and that calculated by assuming non-interacting orbitals is small. The remaining kinetic energy is incorporated into an exchange–correlation term, $E_{xc}[\rho(\mathbf{r})]$, and a general $E_e[\rho(\mathbf{r})]$ DFT energy expression can be written as

$$E_e^{\mathrm{DFT}}[\rho(\mathbf{r})] = V_{en}[\rho(\mathbf{r})] + T_s[\rho(\mathbf{r})] + J_e[\rho(\mathbf{r},\mathbf{r}')] + E_{xc}[\rho(\mathbf{r})]. \tag{2.72}$$

At this stage, the only unknown contribution is $E_{xc}[\rho]$. By equating $E_e^{\mathrm{DFT}}[\rho(\mathbf{r})]$ to the exact $E_e[\rho(\mathbf{r})]$ energy, the exchange-correlation contribution can be defined as

$$E_{xc}[\rho(\mathbf{r})] = (T[\rho(\mathbf{r})] - T_s[\rho(\mathbf{r})]) + (V_{ee}[\rho(\mathbf{r},\mathbf{r}')] - J_e[\rho(\mathbf{r},\mathbf{r}')]). \tag{2.73}$$

The first term on the right-hand side of Equation 2.73 can be interpreted as the kinetic correlation energy, whereas the second encompasses both the potential correlation and the exchange contributions. In orbital-free models, the central challenge is to develop approximations for the kinetic, exchange, and correlation energy functionals separately. By contrast, in KS theory, the problem reduces to devising approximations solely for the exchange–correlation functional.

The variational principle applied to the electron density determines the ground-state density by optimizing the KS orbitals of the non-interacting reference system to minimize the energy functional. The exchange–correlation potential is then defined as the functional derivative of the exchange–correlation energy

$$V_{xc}\left(\left[\rho\left(\mathbf{r}\right)\right]\right) = \frac{\delta E_{xc}\left[\rho\left(\mathbf{r}\right)\right]}{\delta\rho\left(\mathbf{r}\right)}. \tag{2.74}$$

To make DFT a practical tool, numerous approximate functionals have been developed to represent $E_{xc}\left[\rho\left(\mathbf{r}\right)\right]$. In practice, this term is separated in two components: $E_{xc}\left[\rho\left(\mathbf{r}\right)\right] = E_x\left[\rho\left(\mathbf{r}\right)\right] + E_c\left[\rho\left(\mathbf{r}\right)\right]$, where $E_x\left[\rho\left(\mathbf{r}\right)\right]$ and $E_c\left[\rho\left(\mathbf{r}\right)\right]$ are the exhange and correlation functionals, respectively, and suitable formulations are developed for both of them. Although hundreds of such approximations have been proposed, there exists no systematic route to improve their accuracy. To rationalize this diversity, Perdew and Schmidt[253] introduced a classification into five families, commonly referred to as the "Jacob's ladder" of DFT. This metaphorical ladder connects the "hell" of the non-interacting electron gas to the "heaven" of chemical accuracy: ascending the ladder generally improves accuracy, but at the expense of increased computational cost. The rungs of the ladder, ordered by increasing complexity, can be summarized as follows.

- The local density approximation (LDA) assumes that $\rho\left(\mathbf{r}\right)$ varies slowly with position, such that the uniform electron gas provides a suitable local approximation for the non-uniform density. Within this framework, the exchange–correlation functional can be written as

$$E_{xc}^{\mathrm{LDA}}\left[\rho\left(\mathbf{r}\right)\right] = \int \rho\left(\mathbf{r}\right)\epsilon_{xc}\left(\rho\left(\mathbf{r}\right)\right)d\mathbf{r} \tag{2.75}$$

  where $\epsilon_{xc}(\rho)$ denotes the exchange–correlation energy per electron. For open-shell systems, improved accuracy is achieved by the local spin density approximation (LSDA), which allows electrons with opposite spin to occupy distinct spatial components of the KS orbitals.

- The generalized gradient approximation (GGA) improves upon the LDA by incorporating corrections for density inhomogeneity through explicit dependence on the density gradients. In this framework, the exchange–correlation functional can be expressed as

$$E_{xc}^{\mathrm{GGA}}\left[\rho\left(\mathbf{r}\right)\right]=\int f\left(\rho\left(\mathbf{r}\right),\nabla\rho\left(\mathbf{r}\right)\right)d\mathbf{r} \tag{2.76}$$

where approximate forms of f are constructed on theoretical constraints based on the exact functional, often supplemented with empirical parameterization to optimize accuracy.

- The meta-generalized gradient approximation (meta-GGA) family further refines the description of the exchange–correlation functional by including higher-order derivatives of the electron density or, more frequently, the kinetic-energy density obtained from the auxiliary orbitals. In general terms, meta-GGA functionals can be formulated as

$$E_{xc}^{\text{meta-GGA}}\left[\rho\left(\mathbf{r}\right)\right]=\int f(\rho\left(\mathbf{r}\right),\nabla\rho\left(\mathbf{r}\right),\nabla^{2}\rho\left(\mathbf{r}\right),\tau)d\mathbf{r} \tag{2.77}$$

where τ denotes the KS kinetic-energy density, defined as

$$\tau=\frac{1}{2}\left|\nabla\chi_{i}^{\mathrm{KS}}\right|^{2}. \tag{2.78}$$

Although meta-GGA functionals are computationally more demanding than GGA, they generally yield improved accuracy.

- Hybrid functionals address the locality limitations of the previous families by incorporating a fraction of non-local "exact" exchange, evaluated in the same manner as in the HF method but using the KS orbitals. The exact exchange energy in this framework can be formulated, for an n-electron system, as

$$E_{x}^{\mathrm{HF}}=-\frac{1}{4}\sum_{i,j=1}^{n}\left\langle\chi_{i}^{\mathrm{KS}}(1)\chi_{j}^{\mathrm{KS}}(2)\left|\frac{1}{r_{12}}\right|\chi_{j}^{\mathrm{KS}}(1)\chi_{i}^{\mathrm{KS}}(2)\right\rangle. \tag{2.79}$$

The central idea of hybrid functionals is to combine this orbital-dependent exchange with the local or semi-local approximations of the LSDA, GGA, or meta-GGA families, thereby improving accuracy with only a moderate increase in computational cost.

- Double-hybrid functionals extend the hybrid framework by incorporating an additional estimate of dynamical correlation from perturbative post-HF methods, most commonly MP2, at the expense of increased computational cost. In practice, a hybrid-GGA functional is first defined as

$$E_{xc}^{\text{hybrid-GGA}}=a_{1}E_{x}^{\mathrm{GGA}}+\left(1-a_{1}\right)E_{x}^{\mathrm{HF}}+a_{2}E_{c}^{\mathrm{GGA}} \tag{2.80}$$

which is used self-consistently to determine the KS orbitals. The exchange–correlation energy is then refined by adding a perturbative correction

$$E_{xc} = E_{xc}^{\text{hybrid-GGA}} + (1 - a_2)\, E_c^{\text{KS-MP2}} \tag{2.81}$$

where $E_c^{\text{KS-MP2}}$ is evaluated from second-order Møller-Plesset perturbation theory using the KS orbitals. The resulting double-hybrid functional provides improved ground-state energies and properties compared to conventional hybrid approaches.

2.6.1 Dispersion Energy

Density functional approximations generally require explicit corrections to properly account for London dispersion forces. Except for double-hybrid functionals, in which perturbative correlation partially captures dispersion, most standard DFT approaches fail to describe long-range interactions adequately. In particular, LDA and GGA predict exponential decays instead of the correct r^{-6} asymptotic behavior, while hybrid functionals do not remedy this deficiency, as the inclusion of HF-like exchange may even lead to artificial repulsion. Among the available strategies, the most widely used are those proposed by Grimme, which are attractive for their simplicity and negligible additional computational cost. The DFT-D3[254,255] scheme improves the description of long-range compared to earlier generations. In addition, the D3(BJ)[256] version employs the Becke–Johnson damping function to regulate the balance between short- and long-range contributions, the former already being described by DFT. More recently, the DFT-D4 model has been proposed,[257,258] which improves the description of non-covalent interactions mostly for systems involving metal centers, while providing only small differences for neutral molecules with respect to DFT-D3.

For a system composed of M nuclei, the correction with zero-damping (D3) is expressed as

$$\Delta E_{\text{disp}}^{\text{D3}} = -\frac{1}{2} \sum_{n=6,8} \sum_{B>A}^{M} s_n \frac{C_n^{AB}}{R_{AB}^n} f_{\text{damp}}\left(R_{AB}\right) \tag{2.82}$$

where s_n denotes a scaling factor that depends on the chosen functional, C_n^{AB} are the n-th order dispersion coefficients associated with each atomic pair AB, R_{AB} represents the inter-nuclear distance, and the damping function is defined as

$$f_{\text{damp}}\left(R_{AB}\right) = \frac{1}{1 + \exp\left[-\gamma\left(\frac{R_{AB}}{s_{r,n} R_0^{AB}} - 1\right)\right]}$$

are damping functions, where R_0^{AB} denote the cut-off radii, $s_{r,6}$ are functional-dependent scaling factors, $s_{r,8}$ is fixed to 1 for all functionals, and γ are constants, with values of 14 for $n = 6$ and 16 for $n = 8$.

The DFT-D3 correction with Becke–Johnson D3(BJ) damping takes the form

$$\Delta E_{\text{disp}}^{\text{D3(BJ)}} = -\frac{1}{2} \sum_{n=6,8} \sum_{B>A}^{M} s_n \frac{C_n^{AB}}{R_{AB}^n + f_{\text{damp}}^n \left(R_{AB}^0\right)} \tag{2.83}$$

with $f\left(R_{AB}^0\right) = a_1 R_{AB}^0 + a_2$, where a_i are empirical coefficients referred to as BJ parameters, and $R_{AB}^0 = \sqrt{\frac{C_8^{AB}}{C_6^{AB}}}$.

In both D3 and D3(BJ), the fitting parameters ($s_{r,6}$ and $s_{r,8}$ for the former, a_1, s_8, and a_2 for the latter) are optimized by a least-squares procedure against a reference set of dispersion interaction energies.

2.7 Basis Sets

All *ab initio* methods rely on the use of a basis set to represent molecular orbitals (MOs). In principle, an exact description would require a complete basis, corresponding to an infinite set of functions, a condition that is computationally infeasible. Using a finite basis set restricts the orbital representation to the chosen functions, so that both the size and the quality of the basis strongly influence the accuracy. More efficient basis functions reduce the size required for a given accuracy, which is crucial since the computational cost of *ab initio* methods scales at least as the fourth power with respect to the basis set size. Therefore, it is essential to balance the compactness of the basis with the accuracy of the results.

A correlation basis set consists of functions, commonly referred to as atomic orbitals, which are linearly combined to construct MOs.

From a general perspective, two main types of atom-centered basis functions are employed in quantum chemical calculations: Slater-type orbitals (STOs) and Gaussian-type orbitals (GTOs).

STOs, in atom-centered polar coordinates, have the functional form

$$\chi_{\zeta,n,l,m}^{\text{STO}}(r,\theta,\varphi) = N \underbrace{Y_{l,m}(\theta,\varphi)}_{\text{angular part}} \underbrace{r^{n-1} e^{-\zeta r}}_{\text{radial part}} \tag{2.84}$$

where n is the principal quantum number, N is a normalization constant, r is the distance of the electron from the atomic nucleus, ζ is the Slater orbital exponent, and $Y_{l,m}$ are spherical harmonics. The exponential dependence of STOs on the electron–nucleus distance mirrors the behavior of hydrogen-like orbitals and ensures relatively rapid convergence with increasing basis set size. A limitation of this type of functions is the lack of radial nodes, which can be incorporated only by forming linear combinations of multiple functions. Moreover, the evaluation of three- and four-center two-electron integrals cannot be performed analytically. For these reasons, STOs are primarily employed in atomic and diatomic systems where very high accuracy is required, in semi-empirical approaches that neglect multi-center integrals,

or in certain density functional methods where the Coulomb interaction is treated by density fitting. To overcome these limitations, GTOs were introduced into MO calculations.

GTOs can be represented in polar form as

$$\chi^{\mathrm{GTO}}_{\alpha,n,l,m}(r,\theta,\varphi) = N \underbrace{Y_{l,m}(\theta,\varphi)}_{\text{angular part}} \underbrace{r^{2n-2-l}e^{-\alpha r^2}}_{\text{radial part}} \tag{2.85}$$

with α being the exponential factor. The quadratic dependence on the electron-nucleus distance in the exponential makes GTOs less accurate than STOs in two respects. First, at the nucleus, a GTO exhibits zero slope, whereas a STO displays the correct cusp behavior, resulting in a more faithful description near the nucleus. Second, GTOs decay too rapidly at large distances, providing a poor representation of the long-range tail of the wave function. Consequently, more GTOs are required to achieve the same level of accuracy as STOs. This drawback, however, is offset by the fact that integrals over GTOs can be evaluated analytically, making them computationally far more efficient. As a result, GTOs have become the basis functions of choice in electronic structure calculations, typically centered on nuclei, although off-nuclear functions are occasionally introduced, for example, at bond midpoints, to improve the description of weak interactions.

When linear combinations of Gaussian primitives with different exponents are employed to approximate the radial behavior of STOs, the resulting functions are referred to as contracted Gaussian-type orbitals (CGTOs), while the individual primitives are termed primitive Gaussian-type orbitals (PGTOs). This procedure, called *basis set contraction*, reduces the number of independent functions required without significantly compromising accuracy. The general form of a CGTO can be represented as

$$\mathrm{CGTO} = \sum_{i=1}^{M} a_i \mathrm{PGTO}_i \tag{2.86}$$

where M denotes the number of primitive Gaussians combined and the coefficients a_i are selected to optimize the overall shape of the function while preserving normalization. The degree of contraction depends on the number of primitive functions combined to form each contracted orbital.

In 1969, Hehre, Stewart, and Pople[259] systematically determined optimal contraction coefficients and exponents for approximating STOs with CGTOs across a large portion of the periodic table. The resulting basis sets were denoted STO-MG, where each STO is represented by a contraction of M PGTOs (G). These constitute minimal basis sets in which the exponents of the primitives are fitted to reproduce the corresponding STO, rather than being optimized variationally. Although STO-MG basis sets with $M = 2-6$ had been derived, it was found that using more than three primitives yields little additional improvement. Consequently, the STO-3G basis set has become the most widely used minimal basis set.

The requirement for a proper representation of chemical bonding, which necessitates greater flexibility in the valence space than in the core, led to the development of split-valence, or valence-multiple-ζ, basis sets. Within this framework, core orbitals are represented by a single contracted function, while valence orbitals are expanded into two or more functions, improving the description of bond formation and electronic distribution. Pople and co-workers systematically developed such basis sets, denoted M-nlmG, where M indicates the number of PGTOs used for the core orbitals and n, l, m specify the splitting of the valence space (two numbers for double-ζ, three for triple-ζ), while the letter G denotes the use of GTOs. Polarization and diffuse functions are commonly added: the former increase the angular flexibility of the basis set, allowing orbitals to distort and better reproduce the anisotropy of chemical bonding and electron density, whereas the latter introduce radially extended orbitals with small exponents, enabling an accurate description of loosely bound electrons, long-range interactions, and Rydberg or charge-transfer states. Additional polarization functions are indicated in parentheses after the G, whereas diffuse functions are denoted by a "+" (one "+" for heavy atoms, "++" to also include hydrogen). In cases where only one set of polarization functions is included, an alternative notation with an asterisk is used. A characteristic of these basis sets is that the same exponent is applied to s- and p-valence functions, a choice that increases computational efficiency but reduces flexibility. Representative examples include 3-21G, 6-21G, 4-31G, 6-31G, and 6-311G.[260–262] The exponents and contraction coefficients were optimized variationally at the Hartree-Fock level over selected atomic and molecular test sets, and segmented contractions were employed, meaning that primitive Gaussians used in one function are not shared with others of the same angular momentum.

In contrast to segmented contractions, general contractions are constructed from a common set of primitive functions that contribute to all contracted orbitals with different coefficients. This scheme offers technical advantages, as integrals involving the same primitives need to be computed only once and can subsequently be reused, improving computational efficiency. Split-valence basis sets employing general contractions were introduced by Dunning and co-workers[263] in the correlation-consistent polarized valence n-tuple zeta (cc-pVnZ) family. The designation "correlation-consistent" reflects two key aspects of these basis sets. First, the exponents and contraction coefficients were variationally optimized with the inclusion of electron correlation. Second, additional functions are systematically introduced according to their relative contribution to the correlation energy, independent of angular momentum. Since s- and p-type functions alone lack the flexibility to represent electron correlation adequately, polarization functions of higher angular momentum are required. For first-row atoms, the most effective additions are d-type Gaussians, while for hydrogen, they are p-type Gaussians. Thus, balanced double-ζ sets include d functions on heavy atoms and p functions on hydrogen, while triple-ζ sets extend this to one f and two d functions on heavy atoms, and one d and two p functions on hydrogen, with higher-ζ expansions following the same systematic scheme.[264]

Standard correlation-consistent basis sets may lack the ability to properly describe weakly bound or spatially diffuse electrons, leading to significant errors in computed energies and molecular properties. To overcome this limitation, such basis sets can be augmented with diffuse functions, indicated by the prefix aug-.[265] In this scheme, one additional function with a small exponent is added for each angular momentum already present. For example, aug-cc-pVDZ includes one extra s-, p-, and d-function on heavy atoms, while aug-cc-pVTZ contains additional s, p, d, and f functions (with appropriate subsets for hydrogen and helium).

Correlation-consistent basis sets can also be enriched with tight functions (large exponents) to better recover core–core and core–valence correlation effects. These sets are denoted cc-pCVnZ (n = D, T, Q, 5), where the number and type of additional tight functions increases systematically with the cardinal number. For instance, cc-pCVDZ includes one additional tight s- and one p-function, cc-pCVTZ adds $2s$, $2p$, and $1d$, cc-pCVQZ includes $3s$, $3p$, $2d$, and $1f$, and cc-pCV5Z extends this further with $4s$, $4p$, $3d$, $2f$, and $1g$ functions on non-hydrogen atoms.[266] While the cc-pCVnZ basis sets exhibit systematic convergence toward the all-electron correlation energy at the CBS limit, the core–valence contribution converges more slowly than the core–core term. To address this limitation, weighted core–valence basis sets[267] (cc-pwCVnZ) were developed, in which the additional functions are optimized primarily with respect to core–valence correlation. These provide improved convergence of energetic and spectroscopic properties and are generally preferred over cc-pCVnZ when accurate treatment of core correlation is required.

For some molecular properties, like ionization potentials and equilibrium geometries, the effect of diffuse functions has been demonstrated to be limited. Since their inclusion not only increases the computational cost but may also lead to self-consistent field convergence difficulties and an enhancement of basis set superposition error, their use should be restricted to cases where they are essential. For this reason, Truhlar and co-workers[268] developed the so-called calendar basis sets, a family of minimally augmented basis sets in which diffuse functions are systematically optimized but kept to the minimum necessary. Named after the months and culminating in the fully augmented "aug-ust" level, these basis sets provide a convergent hierarchy of partially augmented expansions, offering a balanced compromise between accuracy and computational efficiency. In many cases, such intermediate augmentation proves more suitable than either minimal or full augmentation.

For third-row elements, all correlation-consistent basis sets can be augmented with an additional tight d function,[264] denoted by the "+d" suffix, which improves the description of polarization effects.

Unlike conventional methods, explicitly correlated approaches do not require steep, high-angular-momentum functions, as these primarily account for short-range correlation effects that are already explicitly described by the correlation factor. To this end, Peterson and co-workers[269] developed the cc-pVnZ-F12 series, a family of basis sets specifically optimized for F12 calculations, for which dedicated auxiliary basis sets have also been provided.[270]

While the cc-pVnZ-F12 series is tailored to valence correlation, a corresponding all-electron extension, denoted cc-pCVnZ-F12, was subsequently introduced to account for core–valence effects as well.[271] Furthermore, basis set extrapolation schemes have been proposed for these families, enabling improved accuracy at reduced computational cost.[272]

2.8 Frozen Natural Orbital Approximation

Achieving accurate and reliable results in computational chemistry requires the use of high-level electron correlation methods, such as the Coupled Cluster hierarchy. However, their applicability is limited to relatively small systems due to the steep computational scaling with system size, as summarized in Table 2.1, and by the intrinsically slow convergence with respect to the one-electron basis set. For quantitative accuracy, relatively large basis sets are required, with their size becoming a limiting factor. In the case of CCSD(T), for instance, the rate-determining steps scale with the fourth power of the basis set size for both the CCSD and the perturbative (T) correction.

Over the years, numerous strategies have been introduced to address these challenges, as outlined in the Introduction. On the one hand, efficient implementations such as density fitting-based CC[273] and massively parallel algorithms[92,274–276] reduce the prefactor without altering the formal scaling. On the other hand, more advanced approaches achieve an actual reduction in scaling, either by employing data compression techniques[277] or by exploiting local correlation approximations, which may approach linear complexity.[112,278–281]

The slow basis set convergence of CC methods has also been intensively investigated. The complete basis set limit of conventional CC energies can be reliably approximated using basis set extrapolation techniques.[282] Alternatively, finite-basis CC results may be improved by applying corrections obtained at a lower level of theory[283] or at different points of the potential energy surface.[284] A further and particularly powerful solution is offered by explicitly correlated approaches at the CCSD level, which incorporate the interelectronic distance directly into the wave function, as discussed in Section 2.5.

An alternative strategy to address the basis set problem is provided by methods based on the transformation and subsequent truncation of the one-electron basis set. In these approaches, the virtual molecular orbital space is partitioned into an active and an inactive subspace, with the latter being neglected in the most computationally demanding steps of the correlation treatment. To minimize the loss of correlation energy, a transformation of the virtual orbitals is performed before truncation, typically by optimizing a functional that depends on the orbital rotation parameters within the virtual space. The orbitals obtained in this manner are commonly referred to as optimized virtual orbitals (OVOs).[285–288] Several functional forms have been proposed for this purpose, including modified second-order Hylleraas functionals,[285–287] as well as formulations based on the overlap of CC wave functions expanded in the full and in the truncated virtual spaces.[287,289] In addition, the

use of CCSD rather than MP1 amplitudes to define the retained virtual subspace has been investigated, revealing potential advantages in accuracy.[290]

A simpler yet efficient strategy to compress the virtual space is provided by the frozen natural orbital approximation. In this scheme, the orbital transformation matrix is obtained by diagonalizing the virtual–virtual block of the one-particle density matrix constructed from a lower-level wave function. The resulting orbitals, known as natural orbitals, are characterized by eigenvalues that represent their occupation numbers.[291–296] Truncation is performed according to a predefined threshold on the occupation numbers, whereby weakly populated NOs are discarded and the more significantly occupied ones retained to define the active virtual space. The reduced set of FNOs thus replaces the full virtual space in the high-level calculation, significantly lowering the computational cost. A schematic overview of the procedure is provided in Figure 2.3. To reduce the computational cost of CCSD and CCSD(T), a common strategy is to employ NOs derived from an underlying Møller-Plesset perturbation theory wave function.[288,297–299]

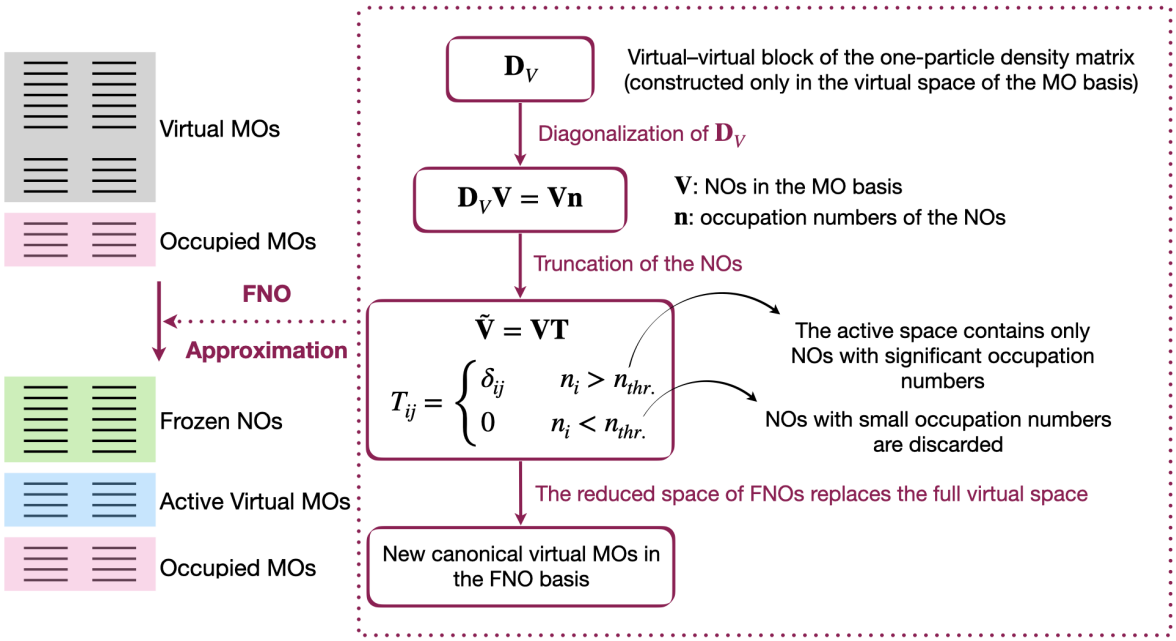


FIGURE 2.3: Outline of the frozen natural orbital approximation procedure.

Although the OVO and FNO approaches show comparable performance in benchmark studies,[288,300] the simplicity of the latter one has led to its broader adoption. Since its introduction, the frozen natural orbital methodology has been extended to larger systems through reduced-scaling density matrix construction techniques[301–304] and integrated in DF-CCSD(T) implementations.[305] Further developments include the implementation of analytic gradients,[298] excited-state extensions,[299,306–309] and extrapolation strategies toward the virtual CBS limit.[93,299,310]

The accuracy of the frozen natural orbital approximation can be further improved by

error-correction schemes. The most widely used is the ΔMP2 correction, based on the difference between MP2 energies in the full and truncated virtual spaces, which provides an efficient estimate of the truncation error and significantly enhances reliability through error compensation.[285,297,298,311]

Kállay and co-workers have substantially advanced the FNO framework, integrating it with several complementary strategies to enhance both accuracy and applicability. They combined the FNO approach with the NAF approximation[244] in an efficient, parallel, integral-direct DF-CCSD(T) algorithm,[92] and assessed MP2-based corrections and extrapolation schemes, showing that such hybrid approaches considerably extend the applicability of FNO-CCSD(T) to systems with 50–75 atoms and triple- to quadruple-ζ basis sets.[93] They also proposed a size-consistent variant of the multiplicative basis-set correction originally developed by Irmler and Grüneis,[312,313] which exploits the parallel convergence behavior of the second-order and Particle-Particle Ladder (PPL) contributions to the CCSD correlation energy.[1,314–317] In their scheme, a ΔMP2 correction, defined as the difference between large- and small-basis MP2 correlation energies, is supplemented with a scaled PPL contribution, thereby reducing the basis set incompleteness error while preserving size-consistency when applied within the NO framework.[94] The PPL approximation isolates the dominant particle–particle correlation effects between virtual orbitals and provides an efficient way to recover the correlation energy lost due to the truncation of the virtual space in the FNO approach.

Finally, Kállay and co-workers extended the (T+) correction (Section 2.5),[185] adapting it to the FNO approximation to accelerate the convergence of the perturbative triples contributions with respect to the number of NOs. In combination with extrapolation techniques, this strategy has been validated through extensive benchmarks, further demonstrating the robustness of the enhanced FNO methodology.[94]

Chapter 3

Composite Schemes for Quantitative Quantum Chemistry

As already discussed in previous Chapters, achieving accurate predictions of molecular properties through *ab initio* electronic structure theory requires careful consideration of both methodological and numerical approximations. Although CCSD(T) is widely regarded as the gold standard for systems with predominantly single-reference character, an individual calculation generally does not achieve spectroscopic precision and, depending on the system under investigation, may not even reach chemical accuracy. In practice, results of this quality demand a balanced and systematic improvement of both the one-electron and N-electron spaces. To obtain the exact solution of the time-independent non-relativistic Schrödinger equation (Equation 2.5), two fundamental requirements must be fulfilled:

1. the orbitals used to construct the determinants must form a complete set in the one-electron space;

2. the expansion of the N-electron wave function must include all determinants that can be generated from these orbitals.

Failure to satisfy either condition introduces errors into the wave function. Importantly, the errors arising from truncation of the one-electron space are qualitatively different from those associated with approximations in the N-electron description. Poor treatment in one of these spaces cannot be compensated for by a better description in the other, although in some instances, error cancellation may occur. The one-particle space corresponds to the size and quality of the basis set used to describe the orbitals in the Hartree-Fock wave function, while the N-particle space defines the level of correlation treatment in the many-electron wave function.

Systematically improving both spaces simultaneously leads, in principle, to convergence toward the exact non-relativistic electronic solution.[29] This relationship is depicted in Figure 3.1, which illustrates how systematic improvements in the one- and N-electron descriptions progressively approach the exact limit. In practice, computational constraints restrict calculations to a finite region of this one-electron-N-electron diagram.

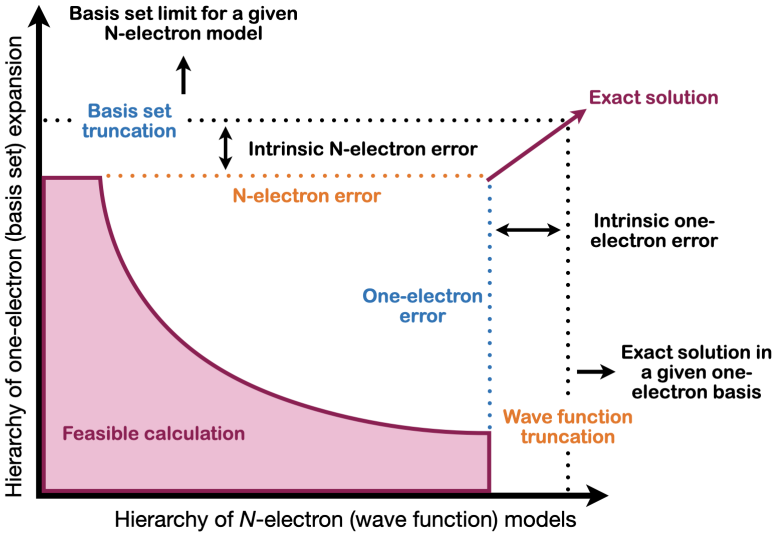


FIGURE 3.1: Approach toward the exact solution of the time-independent non-relativistic Schrödinger equation and schematic representation of the main sources of error in *ab initio* calculations. The overall deviation arises from the truncation of both the one-electron basis set and the N-electron wave function.

Consequently, the total error in an *ab initio* calculation can be regarded as the sum of two principal contributions: (i) the one-electron or basis set truncation error, and (ii) the N-electron or wave function model truncation error.

These two primary errors are unavoidable in any *ab initio* calculation, as neither an infinite basis set nor an exhaustive expansion of the wave function is computationally feasible. Their conceptual interplay is also illustrated in Figure 3.1, where the total deviation from the exact limit originates from both basis set truncation and wave function incompleteness.

Additionally, minor sources of error may arise from neglecting effects such as core-valence correlation, relativistic contributions, or other factors beyond those included in the chosen correlation model.

The most effective strategy to minimize both the principal and minor sources of error inherent in *ab initio* calculations is the use of composite schemes. These methods rely on the *additivity approximation*, whereby the target quantity is decomposed into multiple contributions that can be computed independently and later combined. This framework allows each term to be evaluated at the most suitable balance between theoretical level and basis set size, achieving high overall accuracy but at a significantly reduced computational cost with respect to the single, fully converged calculation. Typically, the main contribution is evaluated using a wave function method such as CCSD(T) in combination with a large basis set or through CBS extrapolation, whereas successive corrections are computed with more sophisticated methods and smaller basis sets. The conceptual workflow of CSs is illustrated in Figure 3.2.

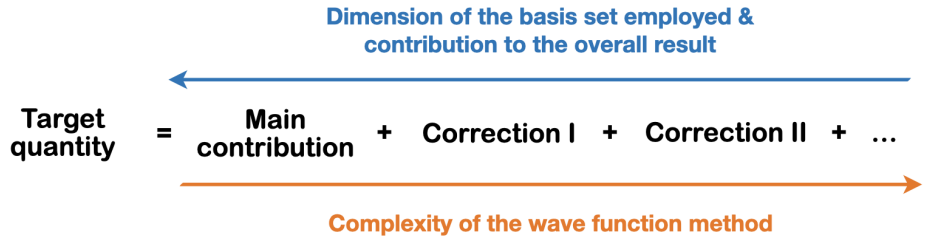


FIGURE 3.2: Conceptual representation of the composite scheme approach. The target quantity is expressed as the sum of a main contribution and successive corrections, each evaluated at an appropriate level of theory and basis set size.

Specific strategies have been developed to systematically reduce both the N-electron (wave function) and one-electron (basis set) truncation errors. The wave function truncation error is mitigated through the hierarchical framework of Coupled Cluster theory.[318] As illustrated in Table 2.1, this family of methods defines a systematic series of correlation models in which the progressive inclusion of excitation levels allows convergence toward the exact solution within a given basis set. The one-electron truncation error can be effectively reduced through basis set extrapolation techniques, which exploit hierarchical sequences of basis sets exhibiting systematic and monotonic convergence toward the complete basis set limit. By evaluating a given property with successively larger basis sets and extrapolating the results, it becomes possible to closely approximate the CBS value without employing prohibitively large expansions. This concept is illustrated in Figure 3.3, where the calculated property approaches the CBS limit as the basis set size increases, leaving only the intrinsic correlation error associated with the electronic structure method.

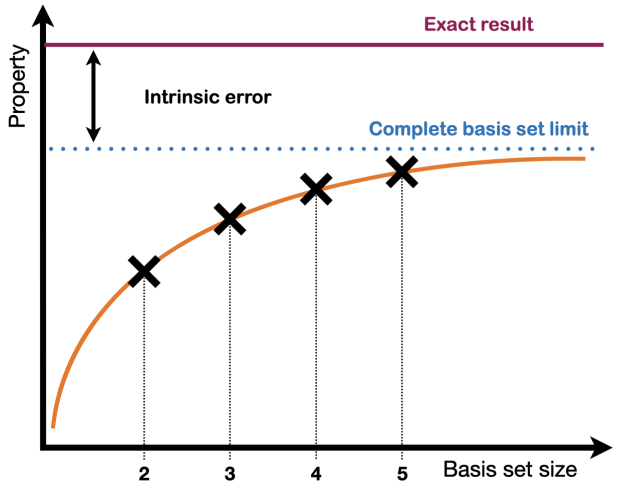


FIGURE 3.3: Convergence of a molecular property with increasing basis set size. Hierarchical basis sets enable extrapolation toward the complete basis set limit, thereby reducing the one-electron truncation error.

A particularly effective class of basis sets for this purpose is represented by the correlation-consistent family.[263] These systematically convergent polarized valence basis sets (cc-pVnZ, with n = D, T, Q, 5, etc.) were designed to recover the correlation energy in a hierarchical and balanced manner, making them ideally suited for extrapolation schemes. The cardinal number n defines the number of contracted functions per atomic valence orbital, and the highest angular momentum increases linearly with n, resulting in an increase of computational cost by approximately one order of magnitude at each successive level.[319,320] The use of these basis sets not only enables reliable estimation of the CBS limit and mitigation of the basis set incompleteness error, but also minimizes basis set superposition errors and provides a clear separation between one-particle and N-particle expansions.[194,195]

Besides the principal sources of error, additional contributions, such as CV correlation, scalar relativistic, spin-orbit, diagonal Born–Oppenheimer corrections, higher-order excitations, and zero-point vibrational energy contributions, can be added to yield the best estimate of the target property. The conceptual structure of this additive strategy beyond CS is illustrated in Figure 3.4, where the target property is expressed as the sum of a main contribution and several distinct correction terms, each accounting for a specific physical effect.

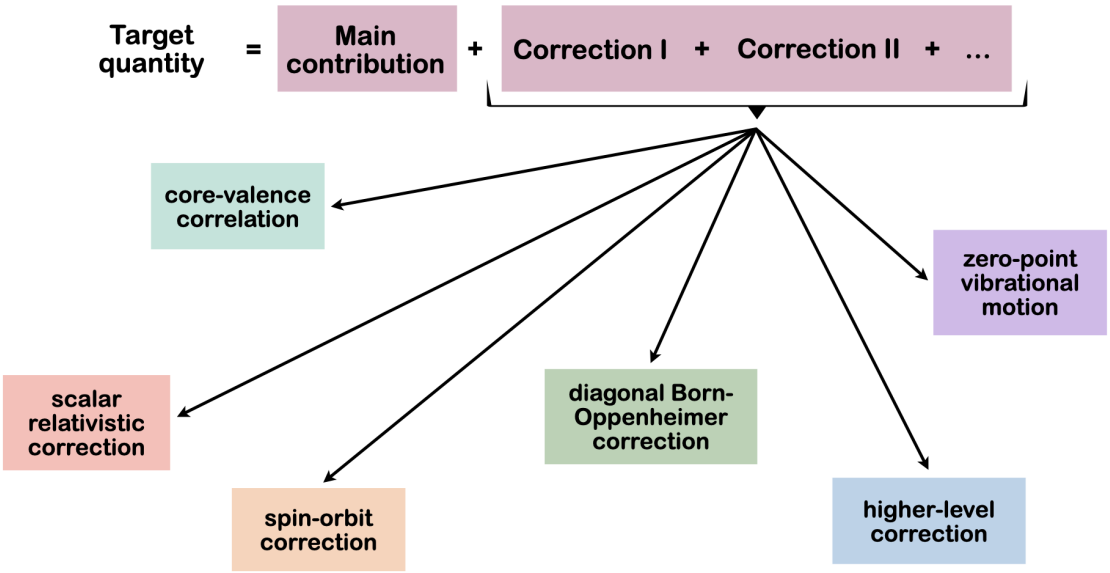


Figure 3.4: Additive decomposition of a target molecular property into the main contribution and several correction terms.

As mentioned in the Introduction, the extensive diversification and refinement of composite *ab initio* methods reflect the continuous effort to achieve high accuracy in computational thermochemistry and molecular spectroscopy. Many series of CSs have been proposed, and within each, several variants have been designed to balance accuracy and computational cost across different chemical applications. As a result, well over one hundred composite *ab initio*

schemes are now available, making the selection of the most appropriate one increasingly challenging.

The choice of a composite method depends on several key factors, including the size and composition of the system, the presence of multi-reference character, the electronic state under investigation, the target property, and the desired level of accuracy. Additional aspects such as bond polarity, formal oxidation state, and overall molecular charge may further constrain the selection of suitable basis sets and correlation treatments.

The various CSs span a broad range of accuracies and differ widely in their domains of applicability. Some are restricted to light-atom systems, whereas others extend across the periodic table. Their computational feasibility also varies, ranging from small molecules containing fewer than five heavy atoms to medium-sized species with several dozens of non-hydrogen atoms. Beyond thermochemical quantities, such as atomization or reaction energies, composite methods have also been developed for molecular structures and spectroscopic observables, like vibrational frequencies and intensities. In this way, they enable quantitatively reliable quantum chemical predictions across a wide variety of chemical contexts.[29,320] Table 3.1 summarizes the commonly adopted definitions of chemical, and spectroscopic or benchmark accuracies for both thermochemical and spectroscopic properties.

Table 3.1: Typical accuracy thresholds defining chemical, spectroscopic, and benchmark accuracy.[29]

Property	Chemical accuracy	Spectroscopic/Benchmark accuracy
Heats of formation	4.18 kJ mol^{-1} = 1 kcal mol^{-1}	1 kJ mol^{-1} = 0.24 kcal mol^{-1}
Weak interactions	0.42 kJ mol^{-1} = 0.10 kcal mol^{-1}	0.10 kJ mol^{-1} = 0.024 kcal mol^{-1}
Bond distances	0.005 Å	0.001 Å
Vibrational frequencies	5.0 cm^{-1}	1.0 cm^{-1}

In the following, state-of-the-art CSs relevant to this PhD work will be presented, beginning with those developed for energetics and subsequently extending to molecular geometries and vibrational frequencies. Particular attention will be devoted to energetic and structural composite methods since, as presented in the next Chapters, they are targeted by the research described in this thesis. For completeness, CSs employed for vibrational frequencies will also be briefly presented.

3.1 Energetics

The general expression for the total energy, including all relevant corrections, is given by

$$E_{\text{tot}} = E^{\text{CBS}}_{\text{valence}} + \Delta E_{\text{CV}} + \Delta E_{\text{HLC}} + \Delta E_{\text{Rel}} + \Delta E_{\text{SO}} + \Delta E_{\text{DBOC}} + \Delta E_{\text{ZPVE}} \tag{3.1}$$

where $E^{\text{CBS}}_{\text{valence}}$ denotes the valence electronic energy extrapolated to the CBS limit, including the HF and valence correlation contributions. The subsequent additive terms represent the individual corrections for core-valence (CV) correlation, higher-level correlation beyond CCSD(T) (HLC), scalar relativistic (Rel) and spin-orbit (SO) effects, and nuclear-motion contributions arising from the diagonal Born–Oppenheimer correction (DBOC) and zero-point vibrational energy (ZPVE).

The first and most important contribution, $E^{\text{CBS}}_{\text{valence}}$, represents the valence electronic energy extrapolated to the CBS limit. In some CSs, the HF and valence correlation components are extrapolated separately, each using an appropriate functional form for their convergence behavior.[11,65] In other approaches, a single extrapolation is performed directly on the correlated energy without explicitly separating the HF contribution.[67] In all cases, because this term dominates the total energy and determines the overall accuracy of the scheme, large and systematically convergent basis sets are employed for its evaluation.

The ΔE_{CV} term accounts for the correlation between valence and core electrons, which becomes increasingly significant for systems containing second-row or heavier elements. It is defined as the difference between an all-electron (*ae*) calculation, in which all electrons are correlated, and a frozen core (*fc*) calculation, where only the valence electrons are included in the correlation treatment, both performed with the same basis set. Thus, this correction quantifies the energetic contribution of the core-valence coupling neglected in the *fc* approximation commonly adopted in correlated calculations. In high-accuracy CSs, the CV term is often extrapolated to the CBS limit and evaluated using basis sets specifically optimized for core-valence correlation, such as the cc-pCVnZ or cc-pwCVnZ families.

The ΔE_{HLC} term accounts for higher-order correlation effects beyond the CCSD(T) level. These corrections compensate for the residual truncation of the electronic wave function and are essential for reaching spectroscopic or benchmark accuracy in high-level composite methods. As the excitation rank included in the correlation treatment increases (see Table 2.1), the computational cost rises steeply, necessitating the use of progressively smaller basis sets for the corresponding additive terms. Consequently, lower-order contributions, such as the inclusion of full triple excitations (ΔE_{CCSDT}), are often extrapolated to the CBS limit, whereas higher-order increments involving connected quadruple or quintuple excitations are typically evaluated with smaller basis sets. In this hierarchy of excitations, approximate CC methods including perturbative corrections are frequently employed to reduce the computational cost while retaining a balanced and systematically improvable description of correlation effects.

The ΔE_{Rel} and ΔE_{SO} terms account for relativistic effects, which become increasingly important for systems containing heavy atoms. The scalar relativistic correction (ΔE_{Rel}) arises from the relativistic motion of the electrons in the strong electrostatic field of the nuclei and incorporates the mass–velocity and Darwin terms of the Dirac Hamiltonian.[321,322] These

effects lead to a contraction of the inner s and p orbitals, which enhances the screening of the nuclear charge and indirectly causes an expansion of the outer d and f shells. The resulting redistribution of the electronic density stabilizes the valence orbitals and alters the computed total energy. The spin-orbit correction (ΔE_{SO}) originates from the interaction between the electron spin and the magnetic field induced by its motion in the electrostatic field of the nuclei and other electrons. This interaction couples the spin and spatial degrees of freedom, leading to the splitting of atomic and molecular energy levels. In most composite protocols, the Rel correction is evaluated as the energy difference between relativistic and non-relativistic calculations performed with the Douglas–Kroll–Hess or exact two-component Hamiltonians, typically at the MP2 or CCSD(T) level of theory.[27] The SO term is generally obtained from atomic fine-structure data or computed explicitly for open-shell systems when necessary.[11,65]

The ΔE_{DBOC} and ΔE_{ZPVE} terms account for corrections associated with nuclear motion. The diagonal Born–Oppenheimer correction arises from the partial breakdown of the Born–Oppenheimer approximation and represents the coupling between electronic and nuclear degrees of freedom within the ground electronic state. It provides a small yet systematic adjustment to the electronic energy, which becomes relevant mainly for light-element systems, particularly those containing hydrogen.[29,323–326] The zero-point vibrational energy, in contrast, reflects the quantum mechanical motion of the nuclei in their vibrational ground state and is required to convert the purely electronic energy into the total energy at 0 K. In most composite protocols, DBOCs are evaluated at the HF or CCSD level, while ZPVEs are derived either from scaled harmonic frequencies or from anharmonic force field calculations, depending on the desired level of accuracy and computational feasibility.[72]

Depending on the level of theory adopted for the evaluation of each individual contribution, as well as on their inclusion or omission within a given composite protocol, the overall performance of the method can be systematically tuned. Consequently, the resulting accuracy and the range of applicability in terms of molecular size and complexity are determined by the chosen balance between computational cost and methodological sophistication.

In the following, CSs particularly relevant to the present doctoral work will be presented.

3.1.1 W4 Theory

The Weizmann-4[11] composite method was designed to achieve sub-kJ mol^{-1} accuracy in molecular energetics through the explicit inclusion of post-CCSD(T) correlation increments. As a high-level composite scheme, W4 provides spectroscopic precision for small molecules, typically up to four or five non-hydrogen atoms, representing one of the most rigorous benchmarks available in modern computational thermochemistry. The expression of the total energy is given by

$$\begin{aligned} E_{\mathrm{W4}} = E_{\mathrm{HF}}^{\mathrm{CBS}} + \Delta E_{\mathrm{CCSD}}^{\mathrm{CBS}} + \Delta E_{\mathrm{CCSD(T)}}^{\mathrm{CBS}} + \Delta E_{\mathrm{CCSDT}} + \Delta E_{\hat{T}_4} + \Delta E_{\hat{T}_5} \\ + \Delta E_{\mathrm{CV}} + \Delta E_{\mathrm{Rel}} + \Delta E_{\mathrm{SO}} + \Delta E_{\mathrm{DBOC}} + \Delta E_{\mathrm{ZPVE}} \end{aligned} \tag{3.2}$$

where each individual contribution is evaluated according to a well-defined hierarchy of correlation levels and basis sets.

- The HF energy extrapolated to the CBS limit, $E_{\mathrm{HF}}^{\mathrm{CBS}}$, is obtained using the two-point Karton–Martin modification[327] of Jensen's extrapolation formula,[328] together with aug-cc-pVnZ basis sets with $n = 5, 6$:

$$E_{\mathrm{HF}}^{\mathrm{CBS}} = E_{\mathrm{HF}}^{n} - a(n+1)e^{-9\sqrt{n}} \tag{3.3}$$

  where E_{HF}^{n} represents the HF-SCF energy obtained with the n-cardinal number basis and a is the scaling coefficient determined from the two E_{HF}^{n} and $E_{\mathrm{HF}}^{(n+1)}$ energies.

- The CBS valence CCSD correlation energy, $\Delta E_{\mathrm{CCSD}}^{\mathrm{CBS}}$, is determined from aug-cc-pVnZ with $n = 5, 6$ using the two-point Klopper extrapolation formula:[329]

$$\Delta E_{\mathrm{CCSD}}^{\mathrm{CBS}} = \Delta E_{\mathrm{CCSD}}^{n} + \frac{a}{n^{\alpha}} \tag{3.4}$$

  with $\Delta E_{\mathrm{CCSD}}^{n} = E_{\mathrm{CCSD}}^{n} - E_{\mathrm{HF}}^{n}$, a obtained from the $\Delta E_{\mathrm{CCSD}}^{n}$ and $\Delta E_{\mathrm{CCSD}}^{(n-1)}$ energy differences, and the exponent $\alpha = 3/5$ for singlet/triplet pairs.

- The CBS valence (T) correction, $\Delta E_{\mathrm{CCSD(T)}}^{\mathrm{CBS}}$, is extrapolated from aug-cc-pVnZ with n = Q, 5, using the same two-point Klopper scheme with $\alpha = 3$:

$$\Delta E_{\mathrm{CCSD(T)}}^{\mathrm{CBS}} = \Delta E_{\mathrm{CCSD(T)}}^{n} + \frac{a}{n^{3}} \tag{3.5}$$

  where $\Delta E_{\mathrm{CCSD(T)}}^{n} = E_{\mathrm{CCSD(T)}}^{n} - E_{\mathrm{CCSD}}^{n}$ and a evaluated from the $\Delta E_{\mathrm{CCSD(T)}}^{n}$ and $\Delta E_{\mathrm{CCSD(T)}}^{(n+1)}$ energy differences.

- The $\Delta E_{\mathrm{CCSDT}}$ term accounts for the energy difference between full triple and perturbative triple excitations. It is extrapolated from cc-pVnZ with n = D, T, using the two-point Klopper formula with α = 3(singlet)/5(triplet):

$$\Delta E_{\mathrm{CCSDT}}^{\mathrm{CBS}} = \Delta E_{\mathrm{CCSDT}}^{n} + \frac{a}{n^{\alpha}} \tag{3.6}$$

  with $\Delta E_{\mathrm{CCSDT}}^{n} = E_{\mathrm{CCSDT}}^{n} - E_{\mathrm{CCSD(T)}}^{n}$ and a obtained from the $\Delta E_{\mathrm{CCSDT}}^{n}$ and $\Delta E_{\mathrm{CCSDT}}^{(n-1)}$ energy differences

- The $\Delta E_{\hat{T}_4}$ correction represents the contribution from connected quadruple excitations and is evaluated as

$$\Delta E_{\hat{T}_4} = 1.1 \cdot \left(\Delta E^{\text{cc-pVDZ}}_{\text{CCSDTQ}} + \Delta E^{\text{cc-pVTZ}}_{\text{CCSDT(Q)}} \right) \tag{3.7}$$

  where $\Delta E^{\text{cc-pVDZ}}_{\text{CCSDTQ}} = E^{\text{cc-pVDZ}}_{\text{CCSDTQ}} - E^{\text{cc-pVDZ}}_{\text{CCSDT(Q)}}$ and $\Delta E^{\text{cc-pVTZ}}_{\text{CCSDT(Q)}} = E^{\text{cc-pVTZ}}_{\text{CCSDT(Q)}} - E^{\text{cc-pVTZ}}_{\text{CCSDT}}$.

- The $\Delta E_{\hat{T}_5}$ term includes connected quintuple excitations:

$$\Delta E_{\hat{T}_5} = E^{\text{cc-pVDZ}}_{\text{CCSDTQP}} - E^{\text{cc-pVDZ}}_{\text{CCSDTQ}}. \tag{3.8}$$

- The core-valence correlation correction, ΔE_{CV}, is computed as the energy difference between *ae*- and *fc*-CCSD(T) calculations, extrapolated to the CBS limit using aug-cc-pwCVnZ (n = T, Q) and the Klopper formula with $\alpha = 3$.

- The scalar relativistic correction, ΔE_{Rel}, is obtained from the difference between non-relativistic and second-order Douglas–Kroll–Hess CCSD(T) calculations in conjunction with the aug-cc-pV(Q+d)Z and aug-cc-pV(Q+d)Z-DK basis sets, respectively.

- Atomic SO coupling corrections, ΔE_{SO}, are taken from experimental fine-structure splittings.

- The diagonal Born–Oppenheimer correction, ΔE_{DBOC}, is evaluated at the HF/aug-cc-pVTZ level.

- The zero-point vibrational energy, ΔE_{ZPVE}, is derived from experimental fundamental frequencies or from CCSD(T)/cc-pV(Q+d)Z anharmonic force field calculations.

The reference geometries are optimized at the CCSD(T)/cc-pV(Q+d)Z level. Unless otherwise specified, all calculations are performed within the frozen core approximation.

3.1.2 Explicitly correlated W1-F12 and W2-F12

The explicitly correlated protocols W1-F12 and W2-F12[47] belong to the Wn family of CSs. They are among the most widely used explicitly correlated approaches in the literature, combining the accuracy of CC methodologies with the accelerated basis set convergence provided by the F12 formalism. This allows near-CBS quality results to be achieved at a substantially reduced computational cost. Both schemes are truncated at the CCSD(T) level and do not include post-CCSD(T) excitations, ensuring an optimal compromise between accuracy and efficiency. In both approaches, all the extrapolations are performed with the two-point Klopper extrapolation formula[329] where the exponents were reoptimized to ensure consistency with the cc-pVnZ-F12 basis sets and to preserve the systematic behavior of the composite framework.

The computational structure of the W1-F12 and W2-F12 protocols can be outlined as follows:

- Reference geometry and zero-point vibrational energy: in both variants, the reference geometries and the ZPVE are obtained using the B3LYP[330–332] hybrid functional in conjunction with the cc-pV(T+*d*)Z basis set. The ZPVE contribution is scaled by a factor of 0.985.

- HF-SCF energy: in both W1-F12 and W2-F12, the SCF energy is evaluated using the so-called explicitly correlated Hartree-Fock approach, denoted as HF*. In this formulation, the CABS singles correction is included to allow the HF orbitals to relax within the complementary auxiliary basis set space. This correction substantially reduces the BSIE in the SCF energy, leading to faster convergence toward the CBS limit. In the W1-F12 scheme, the HF* energies are extrapolated to the CBS limit using the two-point extrapolation formula based on the cc-pV*n*Z-F12 (n = D, T) basis sets with an exponent of 5.00. In contrast, the W2-F12 protocol employs a single HF* calculation at the cc-VQZ-F12 level, which provides results already close to the CBS limit due to the accelerated basis set convergence of the F12 methodology.

- Valence CCSD: in the W1-F12 protocol, the valence CCSD-F12b correlation energy is extrapolated from the cc-V*n*Z-F12 (n = D, T) basis set pair using an exponent of 3.67, while in W2-F12 it is extrapolated from the n = T, Q pair with an exponent of 5.94.

- Valence (T): in the W1-F12 protocol, the perturbative triple excitation (T) contribution is extrapolated from the aug'-cc-pV*n*Z (n = D, T) basis set pair using the two-point extrapolation formula with an exponent of $\alpha = 3.22$. In the W2-F12 variant, the corresponding contribution is obtained directly as the (T*)/cc-VTZ-F12 correlation energy scaled by a factor of 0.987. The notation aug'-cc-pV*n*Z denotes a properly defined basis set designed to ensure a balanced treatment of different elements across the periodic table. Specifically, it combines the standard cc-pV*n*Z basis set for hydrogen, the aug-cc-pV*n*Z basis for second-row elements, and the aug-cc-pV(*n*+*d*)Z basis for third-row elements.

- Core-valence: in both variants, the CCSD and (T) components are obtained from standard calculations. In the W1-F12 protocol, the CCSD component is evaluated using the cc-pwCVTZ basis set, while in W2-F12 it is computed with the aug'-cc-pwCVTZ basis set. In both cases, the resulting CCSD contribution is scaled by a factor of 1.1. The (T) component is calculated with the cc-pwCVTZ basis set removing the f functions in both W1-F12 and W2-F12 protocols.

- The scalar relativistic correction is obtained as the difference between non-relativistic and relativistic CCSD(T) calculations, performed with the aug'-cc-pV*n*Z and aug'-cc-pV*n*Z-DK basis sets, respectively, using n = D in the W1-F12 protocol and n = T in the W2-F12 variant.

- Spin-orbit and diagonal Born–Oppenheimer corrections: in both variants, the SO splitting are taken from experimental fine-structure data, while the DBOC is evaluated at the HF/aug'-cc-pVTZ level of theory.

3.1.3 HEAT Protocol

The High-accuracy Extrapolated *Ab initio* Thermochemistry[65] protocol represents one of the most accurate and systematically formulated composite *ab initio* methods currently available, developed to deliver energetics with sub-kJ mol^{-1} accuracy. In terms of performance and predictive reliability, HEAT is comparable to the W4 protocol, while offering a distinct formulation. The total energy is obtained as the sum of the following contributions

$$\begin{aligned} E_{\mathrm{HEAT}} = E_{\mathrm{HF}}^{\mathrm{CBS}} + \Delta E_{ae\text{-}\mathrm{CCSD(T)}}^{\mathrm{CBS}} + \Delta E_{\mathrm{CCSDT}} + \Delta E_{\mathrm{HLC}} \\ + \Delta E_{\mathrm{Rel}} + \Delta E_{\mathrm{SO}} + \Delta E_{\mathrm{DBOC}} + \Delta E_{\mathrm{ZPVE}} \end{aligned} \tag{3.9}$$

where the terms are evaluated as follows.

- The HF energy extrapolated to the CBS limit, $E_{\mathrm{HF}}^{\mathrm{CBS}}$, is obtained from aug-cc-pCVnZ basis sets with n = T, Q, 5 using the three-point Feller extrapolation formula:[333]

$$E_{\mathrm{HF}}^{\mathrm{CBS}} = E_{\mathrm{HF}}^{n} - ae^{-bn}. \tag{3.10}$$

  where the parameters a and b are fitting coefficients. The parameter a represents the magnitude of the residual basis set error, while b controls the exponential rate at which the HF energy converges toward the CBS limit. Both are determined from a non-linear fit to E_{HF}^{n}, $E_{\mathrm{HF}}^{(n+1)}$, and $E_{\mathrm{HF}}^{(n+2)}$ energies.

- The all-electron CCSD(T) correlation energy, $\Delta E_{\mathrm{CCSD(T)}}^{\mathrm{CBS}}$, is extrapolated to the CBS limit from aug-cc-pCVnZ basis sets with n = Q, 5 using the two-point Helgaker formula.

- The $\Delta E_{\mathrm{CCSDT}}$ correction accounts for the difference between full and perturbative triple excitations and is extrapolated using the two-point Helgaker formula with the cc-pVTZ and cc-pVQZ basis sets

$$\Delta E_{\mathrm{CCSDT}} = E_{\mathrm{CCSDT}}^{\text{cc-pV(T,Q)Z}} - E_{\mathrm{CCSD(T)}}^{\text{cc-pV(T,Q)Z}}. \tag{3.11}$$

  where the notation cc-pV(T,Q)Z denotes an extrapolation performed using the cc-pVnZ basis sets with n = T and Q.

- The higher-level correction term, ΔE_{HLC}, represents the contribution from connected quadruple or quintuple excitations, depending on the desired level of accuracy, evaluated using the cc-pVDZ basis set

$$\Delta E_{\mathrm{HLC}} = E^{\text{cc-pVDZ}}_{\mathrm{CCSDTQ/P}} - E^{\text{cc-pVDZ}}_{\mathrm{CCSDT/Q}} \tag{3.12}$$

- The scalar relativistic correction, ΔE_{Rel}, is computed by contracting the one-particle density matrix obtained at the CCSD(T)/aug-cc-pCVTZ level with the one-electron Darwin and mass–velocity terms.
- The spin-orbit contribution, ΔE_{SO}, is evaluated using the SO configuration interaction procedure in conjunction with cc-pVDZ basis sets and relativistic effective core potentials.
- The diagonal Born–Oppenheimer correction, ΔE_{DBOC}, is calculated at the HF/aug-cc-pVTZ level of theory.
- The zero-point vibrational energy, ΔE_{ZPVE}, is derived from an *ae*-CCSD(T)/cc-pVQZ anharmonic force field.

The reference geometries are optimized at the *ae*-CCSD(T)/cc-pVQZ level of theory. Unless otherwise specified, all calculations are performed within the frozen core approximation.

The main differences with respect to the W4 protocol lie in the extrapolation schemes employed, the treatment of electron correlation, and the evaluation of SO effects. In the HEAT methodology, the CCSD and (T) components are not extrapolated separately; instead, the complete CCSD(T) correlation energy is obtained in a single step. Moreover, the calculation is performed by correlating all electrons, and thus no distinct CV correction is required. Another distinguishing feature is the explicit evaluation of the SO contribution through dedicated calculations.

3.1.4 HEAT-like and CBS + CV Approaches

The HEAT-like and CBS + CV approaches[156,157] were developed as simplified yet accurate alternatives to the full HEAT protocol, aiming to extend the applicability of composite methods to small to medium-size molecular systems. These schemes retain the hierarchical and additive structure of the HEAT methodology, while reducing the overall computational cost by employing lower levels of theory for the individual contributions and by omitting some corrections included in the parent protocol.

In the HEAT-like framework, the total energy can be expressed as

$$E_{\text{HEAT-like}} = E^{\mathrm{CBS}}_{\mathrm{HF}} + \Delta E^{\mathrm{CBS}}_{\mathrm{CCSD(T)}} + \Delta E_{\mathrm{CV}} + \Delta E_{\mathrm{CCSDT}} + \Delta E_{\mathrm{CCSDTQ}} + \Delta E_{\mathrm{DBOC}} + \Delta E_{\mathrm{Rel}} + \Delta E_{\mathrm{ZPVE}} \tag{3.13}$$

where $E^{\mathrm{CBS}}_{\mathrm{HF}}$ and $\Delta E^{\mathrm{CBS}}_{\mathrm{CCSD(T)}}$ denote the extrapolations of the HF and valence CCSD(T) correlation energies to the CBS limit, respectively. The former is obtained using the exponential formula proposed by Feller,[333] while the latter is evaluated using Helgaker's expression.[282]

Correlation-consistent cc-pVnZ basis sets are employed for these calculations, typically with n = T, Q, and 5 for the HF extrapolation, and n = T and Q for the CCSD(T) correlation energy. The ΔE_{CV} term accounts for the core-valence correlation correction, obtained as the energy difference between all-electron and frozen core CCSD(T) calculations performed with the cc-pCVTZ basis set. Similarly, higher-order correlation effects beyond the CCSD(T) level are included through the $\Delta E_{\mathrm{CCSDT}}$ and $\Delta E_{\mathrm{CCSDTQ}}$ terms, computed as the energy differences between *fc*-CCSDT/cc-pVTZ and *fc*-CCSD(T)/cc-pVTZ, and between *fc*-CCSDTQ/cc-pVDZ and *fc*-CCSDT/cc-pVDZ, respectively. Because full CCSDTQ computations are computationally demanding even with double-ζ bases, the effect of connected quadruple excitations can alternatively be estimated using the CCSDT(Q) model. The DBOC term is usually evaluated at the HF/aug-cc-pVnZ level with n = D or T, while ΔE_{Rel} is computed at the *ae*-CCSD(T)/aug-cc-pCVnZ level, n = D or T, including the one-electron Darwin and mass–velocity terms. The SO correction can be included by using experimental fine-structure data. The molecular equilibrium structures and the anharmonic zero-point vibrational energy are obtained at the DFT level using DFT-D3 dispersion-corrected[254,256] B2PLYP[334,335] or revDSD-PBEP86[336] double-hybrid functionals in conjunction with triple-ζ quality basis sets.

Compared to the original HEAT protocol, the HEAT-like scheme differs in several respects: it includes a separate evaluation of the CV correction, thereby reducing the computational cost of the CCSD(T) extrapolation with respect to the all-electron treatment; it employs lower levels of theory for post-CCSD(T) contributions, which are therefore not extrapolated to the CBS limit; and it omits the explicit calculation of the SO correction, whose inclusion depends on the system under investigation. Furthermore, differences also arise in the level of theory adopted for the equilibrium geometries and the corresponding ZPVE corrections.

A more compact variant, referred to as the CBS + CV scheme, includes only the extrapolation to the complete basis set limit and the core-valence correlation contribution (i.e., first three terms on the right-hand side of Equation 3.13), typically yielding total energies within 1–2 kJ mol^{-1} of benchmark or experimental values. Equilibrium geometries and ZPVEs are evaluated at the same level of theory as in the HEAT-like protocol.

Both HEAT-like and CBS + CV approaches represent efficient and systematically improvable approximations to the full HEAT scheme, offering a flexible hierarchy of accuracy levels applicable to a broad range of molecular sizes and chemical compositions. HEAT-like methods are typically applicable to systems containing up to about 5-7 heavy atoms, whereas the CBS + CV protocol can be reliably extended to molecules with up to 7–10 non-hydrogen atoms.

3.1.5 The Cheap Scheme

The so-called Cheap Scheme[67] was introduced as a simplified composite protocol aimed at extending high-accuracy *ab initio* thermochemical methodologies to medium-sized molecular systems while keeping the computational cost affordable. Following the same hierarchical and additive philosophy as HEAT-like and CBS + CV schemes, ChS replaces the

most demanding CC high-level correlation and extrapolation steps with MP2-based corrections, while retaining CCSD(T) as the reference level of theory. The total energy in the ChS framework is expressed as

$$E_{\mathrm{ChS}} = E_{\mathrm{CCSD(T)}}^{\text{cc-pVTZ}} + \Delta E_{\mathrm{MP2}}^{\mathrm{CBS}} + \Delta E_{\mathrm{MP2}}^{\mathrm{CV}} \tag{3.14}$$

where:

- $E_{\mathrm{CCSD(T)}}^{\text{cc-pVTZ}}$ is the frozen core CCSD(T) energy computed with the cc-pVTZ basis set;
- $\Delta E_{\mathrm{MP2}}^{\mathrm{CBS}}$ corresponds to the MP2-based correction for the BSIE, evaluated by means of the two-point Helgaker extrapolation formula[282] using the cc-pVnZ basis sets with n = T and Q;
- $\Delta E_{\mathrm{MP2}}^{\mathrm{CV}}$ represents the core-valence correlation contribution, computed as the energy difference between all-electron and frozen core MP2 calculations with the cc-pCVTZ basis set.

In contrast to the HEAT-like and CBS + CV approaches, the ChS protocol adopts a slightly different philosophy, as it starts from a single, non-extrapolated CCSD(T) calculation and subsequently adds a correction for the BSIE, rather than extrapolating the HF and correlation components separately. This correction is evaluated at the MP2 level, which also provides the basis for estimating the CV contribution. Consequently, the scheme omits post-CCSD(T) and other corrections, reduces the number of extrapolations, and relies on more compact basis sets throughout the protocol.

The ChS scheme achieves results within 1–2 kJ mol^{-1} of benchmark HEAT-like or CBS + CV values, offering a scalable compromise between accuracy and computational feasibility. Its robustness and low cost make it applicable to molecular systems containing up to 10–20 non-hydrogen atoms.

3.1.6 junChS and junChS-F12

Building upon the framework of the ChS approach, the jun Cheap Scheme[89] and its F12 version[90] were developed to further improve the balance between accuracy and computational cost, particularly for non-covalent molecular complexes.

The junChS model introduces partially augmented jun-cc-pVnZ basis sets of the calendar family,[268] designed to include a reduced number of diffuse functions while preserving the correct description of long-range correlation effects, which are essential for weak intermolecular interactions. This modification allows a substantial reduction in the computational effort compared to the use of fully augmented aug-cc-pVnZ sets, with negligible loss in accuracy.

In the junChS protocol, the basic ChS structure is retained, with the reference energy obtained from a frozen core CCSD(T)/jun-cc-pV(T+d)Z calculation and the CBS and CV

corrections both evaluated at the MP2 level. The CBS correction is derived from the two-point Helgaker extrapolation formula[282] using the jun-cc-pV(T+*d*)Z and jun-cc-pVQZ basis sets, whereas the CV contribution is computed as the energy difference between all-electron and frozen core MP2 calculations with the cc-pwCVTZ basis set. Geometries and anharmonic ZPVEs are determined at the revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+*d*)Z level of theory.

To further accelerate the convergence toward the CBS limit, the junChS-F12 variant incorporates explicitly correlated methods, replacing the conventional CCSD(T) and MP2 calculations with their F12 counterparts. In particular, the F12b[235,239] and the FIX[234] approximations are employed for the CCSD and MP2 computations, respectively.

Extensive benchmarks[90] demonstrated that junChS-F12 achieves sub-chemical accuracy for interaction energies significantly improving over junChS, while reducing the computational cost by more than an order of magnitude compared to high-level CCSD(T)/CBS calculations.

3.1.7 SVECV-f12

The Simple Version of an Extrapolated, Core-Valence correlation corrected, CCSD(T)-f12[68] represents a cost-effective methodology, designed for accurate thermochemical characterizations. Molecular geometries and harmonic frequencies are computed using the hybrid M06-2X-D3[254,337] functional in combination with the aug-cc-pVTZ basis set, which provides a reliable description of equilibrium structures.

The energetics within the SVECV-f12 protocol is evaluated through the explicitly correlated CCSD(T) method, employing the F12b[235,239] approximation. The extrapolation to the CBS limit is performed using Martin's two-point formula[338] in conjunction with the cc-pVDZ-F12 and cc-pVTZ-F12 basis sets. CV correlation effects are included as an additive correction, evaluated as in the ChS.

Overall, the SVECV-f12 protocol can be applied to molecular systems of comparable size to those accessible to the junChS and junChS-F12 schemes.

By construction, these last methods include the evaluation of CBS and CV corrections without additional contributions, and therefore occupy an intermediate position between the domains of chemical and spectroscopic, or benchmark, accuracy, effectively bridging the gap between computational efficiency and high-level precision.

3.2 Molecular Structures

The concept of composite *ab initio* methods, originally devised for the highly accurate prediction of thermochemical properties, can be extended to the determination of equilibrium molecular geometries. In this context, the same partitioning as Equation 3.1 is retained with exclusion of the ZPVE term, which represents a vibrational contribution. When this formalism is applied to geometry evaluations, the same hierarchical decomposition of the energy is retained. In this case, analytic energy gradients require the first derivatives of each energy

contribution with respect to the nuclear coordinates. Although the general framework remains unchanged, the computational cost increases considerably due to the need for gradient evaluations instead of single-point energies. The equilibrium geometry can be obtained by minimizing the overall composite gradient, which guarantees internal consistency among all corrections and yields structures corresponding to a well-defined potential energy surface.

A rigorous implementation of this approach, referred to as the *gradient scheme*, was developed by Heckert, Kállay, and Gauss,[339] who proposed a CS for molecular geometries based on the following additive energy expression

$$E_{\text{tot}} = E_{\text{CCSD(T)}} + \Delta E_{\text{CCSDT}} + \Delta E_{\text{CCSDTQ}} + \Delta E_{\text{CV}}. \tag{3.15}$$

In the previous Equation, according to Reference 339, the reference term is evaluated at the *fc*-CCSD(T) level with a large cc-pV6Z basis set, while higher-order correlation corrections are included as additive terms, both within the frozen core approximation, as

$$\Delta E_{\text{CCSDT}} = E^{\text{cc-pVTZ}}_{\text{CCSDT}} - E^{\text{cc-pVTZ}}_{\text{CCSD(T)}} \tag{3.16}$$

and

$$\Delta E_{\text{CCSDTQ}} = E^{\text{cc-pVDZ}}_{\text{CCSDTQ}} - E^{\text{cc-pVDZ}}_{\text{CCSDT}} \tag{3.17}$$

where the first Equation accounts for the effect of full triples and the second one that of the connected quadrupoles. Finally, the core-valence correlation correction is computed as

$$\Delta E_{\text{CV}} = E^{\text{cc-pCVQZ}}_{\textit{ae}\text{-CCSD(T)}} - E^{\text{cc-pCVQZ}}_{\textit{fc}\text{-CCSD(T)}}. \tag{3.18}$$

The analytic gradient employed during the geometry optimization can therefore be written as the derivatives of Equation 3.15 with respect to the nuclear coordinates x

$$\frac{dE_{\text{tot}}}{dx} = \frac{dE_{\text{CCSD(T)}}}{dx} + \frac{d\Delta E_{\text{CCSDT}}}{dx} + \frac{d\Delta E_{\text{CCSDTQ}}}{dx} + \frac{d\Delta E_{\text{CV}}}{dx}. \tag{3.19}$$

Each correction term is defined analogously to the corresponding energetic contribution, that is, as the difference between analytic gradients evaluated at two successive levels of theory. This formulation ensures that the additive structure of the total energy is consistently preserved at the gradient level, allowing a systematic and coherent refinement of the equilibrium geometry at each step. This differential formulation enables the determination of equilibrium structures that reproduce experimental geometries with spectroscopic accuracy, which is achieved only when the additive correlation corrections are explicitly included, as demonstrated by Heckert and co-workers.[339]

In a subsequent refinement, Heckert and co-workers[340] introduced a basis set extrapolation scheme for analytic gradients to reduce BSIE in optimized geometries. In this approach, the HF and electron-correlation components are treated separately through differentiation of

the corresponding energy extrapolation formulas. Accordingly, the analytical formulation of the composite gradient takes the form

$$\frac{dE_{\text{tot}}}{dx} = \frac{dE_{\text{HF}}^{\text{CBS}}}{dx} + \frac{d\Delta E_{\text{CCSD(T)}}^{\text{CBS}}}{dx} + \frac{d\Delta E_{\text{CCSDT}}}{dx} + \frac{d\Delta E_{\text{CCSDTQ}}}{dx} + \frac{d\Delta E_{\text{CV}}}{dx}. \tag{3.20}$$

Each term in this expression follows the same hierarchical structure adopted for the corresponding energy contributions. Accordingly to the formulation of Reference 340, the HF component is extrapolated to the CBS limit using the three-point Feller exponential formula[333] with cc-pVnZ basis sets and n = Q, 5, 6, while the valence CCSD(T) correlation contribution is obtained from a two-point Helgaker extrapolation[282] employing the same family of basis sets with n = Q, 5 or 5, 6. The higher-order correlation terms accounting for connected triple and quadruple excitations are computed using cc-pVTZ and cc-pVDZ basis sets, respectively, whereas the core-valence correction is evaluated as the difference between the analytic gradients of the all-electron and frozen core CCSD(T)/cc-pCVQZ calculations.

This refined composite gradient scheme demonstrates that CBS extrapolation can be successfully applied to gradient-based optimizations. By partitioning the evaluation of the valence contribution into separate extrapolated components, rather than performing a single CCSD(T) optimization with a very large basis set (e.g., cc-pV6Z), the overall computational cost is significantly reduced. In this way, the method achieves faster basis set convergence while retaining the internal consistency of the additive framework, yielding equilibrium geometries of spectroscopic accuracy.

Accurate prediction of spectroscopic observables represents one of the most stringent and meaningful validations of modern QC methods. Among these, rotational constants play a central role, providing a direct bridge between the molecular structure and experimentally measurable quantities. Indeed, rotational constants B_i (with $i = a, b, c$ being the principal axes of inertia) are inversely proportional to the corresponding principal moments of inertia I_i, according to

$$B_i = \frac{h}{8\pi^2 I_i} \tag{3.21}$$

where h is Planck's constant. As a consequence, even sub-milliångström variations in bond lengths or tenths of a degree in bond angles lead to measurable shifts in the rotational constants. For isolated systems, experimentally measured rotational constants provide reference values to test theoretical predictions, while accurate QC results are indispensable for the interpretation and assignment of increasingly complex spectra. In this mutual interplay, highly accurate CSs enable theoretical methods to reach the precision required to guide and assist experimentalists, and they also provide predictions for systems for which experimental data are scarce or not available.

From a theoretical standpoint, the ground vibrational state rotational constants (B_i^0) derived experimentally can be expressed as

$$B_i^0 = B_i^e + \Delta B_i^{\mathrm{vib}} + \Delta B_i^{\mathrm{el}} \tag{3.22}$$

where B_i^e represents the equilibrium rotational constant derived from the corresponding geometry, $\Delta B_i^{\mathrm{vib}}$ accounts for vibrational corrections arising from zero-point vibrational motion, and ΔB_i^{el} includes electronic contributions associated with the rotational g-tensor. The equilibrium term B_i^e, depending only on isotopic masses and geometrical parameters, is obtained from geometry optimizations, while $\Delta B_i^{\mathrm{vib}}$ can be evaluated through second-order vibrational perturbation theory[341] (VPT2) based on anharmonic force fields including semi-diagonal force constants. The electronic correction, though typically small, can be determined from the expectation value of the rotational g-tensor and becomes relevant for light-atom or open-shell systems.

Benchmark studies have shown that this combined approach, in which composite equilibrium geometries are complemented by vibrational and electronic corrections, yields rotational constants with spectroscopic accuracy. MADs as low as 0.04% (with a standard deviation of 0.07%) have been reported across a wide set of small molecules, including isotopologues, confirming that the residual differences with respect to experimental data fall well within the limits of high-resolution rotational spectroscopy.[152,342]

The accuracy of the predicted rotational constants depends critically on the level of theory adopted for each contribution within the composite framework. In this respect, a *fc*-CCSD(T)/cc-pVTZ calculation exhibits relatively large deviations from experiment that progressively decrease when larger basis sets are employed, from cc-pVQZ to cc-pV6Z, and are further reduced by extrapolation to the CBS limit. The subsequent inclusion of core-valence correlation brings the MAD within the 0.1% accuracy threshold. Additional improvements are obtained by incorporating higher-order correlation effects, such as connected triple and quadruple excitations, while the DBOC, although generally small, becomes relevant for light-atom systems. Relativistic contributions are minor for first- and second-row elements, but become increasingly important for heavier atoms. For instance, in bromine-containing molecules such as bromomethane (CH_3Br) or bromofluoromethane (CH_2BrF), relativistic effects lead to a shortening of the C – Br bond by about 0.001–0.002 Å, exceeding the magnitude of post-CCSD(T) corrections and thus significantly affecting the corresponding equilibrium rotational constants.[343]

Overall, these results clearly demonstrate that the hierarchical inclusion of physically motivated corrections, ranging from basis set enlargement and CBS extrapolation to CV, post-CCSD(T), relativistic, and other small effects, is essential to achieve quantitative spectroscopic accuracy and to establish a direct connection between computed equilibrium structures and experimentally observed rotational spectra.

Building upon the composite gradient scheme, which determines equilibrium geometries through the analytic differentiation of the total energy expansion, an effective *geometry scheme*

was introduced assuming that the additivity approximation can be applied directly to geometrical parameters.[344] In this formulation, all additive components of the composite expansion, including the HF, correlation, CV, and other terms, are evaluated directly at the optimized geometries. This approach preserves the hierarchical structure of the method while substantially lowering the computational cost.

The CBS + CV geometry scheme retains the first, second, and fifth terms of the right-hand side of Equation 3.20, with additivity exploited at the level of structural parameters r, which therefore is expressed as

$$r_{\text{tot}} = r_{\text{HF}}^{\text{CBS}} + \Delta r_{\text{CCSD(T)}}^{\text{CBS}} + \Delta r_{\text{CV}} \tag{3.23}$$

where $r_{\text{HF}}^{\text{CBS}}$ is the CBS value of the structural parameter obtained at the HF level of theory, $\Delta r_{\text{CCSD(T)}}^{\text{CBS}}$ represents the difference in CCSD(T) geometrical parameters extrapolated to the CBS limit, and Δr_{CV} denotes the core-valence correction evaluated as the difference between all-electron and frozen core optimized geometries. The geometry and gradient schemes yield very similar equilibrium structures, with discrepancies below 0.001 Å for bond lengths and 0.1° for bond angles.[344]

The same conceptual framework can be extended to all composite methodologies, allowing the formulation of geometry-based counterparts of energetic schemes. CSs for molecular geometries rooted in CC ansatz provide spectroscopic accuracy but are computationally affordable only for small systems. To extend the applicability of composite approaches toward medium-sized molecules, geometry-based formulations of the ChS[67] and junChS[89] schemes have been developed. These protocols reproduce the theoretical hierarchy and basis set structure of their energetic counterparts, applying the same additive rationale directly to optimized geometrical parameters. In both cases, the CCSD(T) method serves as the reference level, while basis set incompleteness and core-valence effects are accounted for through MP2 corrections, ensuring accurate and computationally efficient equilibrium structures. In the explicitly correlated junChS-F12 variant,[90] a slightly different strategy is employed. In this case, extrapolation to the CBS limit is neither required nor beneficial, as F12 methods intrinsically yield structural parameters very close to basis set convergence. Therefore, the core-valence correlation correction, computed at the MP2-F12 level, is directly added to the *fc*-CCSD(T)-F12b/jun-cc-pV(T+*d*)Z geometries, providing a computationally efficient yet accurate description of molecular structures.

3.3 Vibrational Frequencies

The extension of CSs to harmonic vibrational frequencies relies on the same hierarchical and additive philosophy adopted for molecular energies and geometries. Tew and co-workers[345] carried out a systematic investigation of the basis set convergence of harmonic frequencies computed at the frozen core CCSD(T) level for a set of closed-shell diatomic molecules. Their results showed that harmonic frequencies may exhibit non-monotonic behavior; therefore,

the extrapolation to the CBS limit should be performed on second derivatives of the PES at a fixed geometry, since they exhibit a smooth convergence. However, a number of applications have shown that for organic molecules, accurate harmonic frequencies can be obtained by performing the extrapolation directly using the eigenvalues of the Hessian matrix, employing composite methods similar to those developed for molecular structures. In particular, the ChS[346] approach has also been successfully extended to the calculation of harmonic vibrational frequencies. In this case, additional corrections are introduced to account for the influence of diffuse basis functions, particularly relevant for stretching modes involving electronegative atoms. An empirical additive term is defined as the difference between MP2 harmonic frequencies computed using the aug-cc-pVTZ and cc-pVTZ basis sets. Although lacking a rigorous theoretical foundation, once the CBS extrapolation has been performed, its inclusion has been shown to improve the overall agreement with experimental data.

While harmonic frequencies provide a first approximation to the vibrational transition energies, quantitative predictions require inclusion of anharmonic effects.

The evaluation of anharmonic corrections to vibrational frequencies can be performed by resorting to VPT2 for which sampling of the PES to obtain at cubic and semi-diagonal quartic force constants is needed. Because of the steep computational scaling of these quantities, this task becomes very expensive at high levels of theory such as CCSD(T), which limits the feasible molecular size and basis set quality. To overcome this limitation, hybrid QM/QM' approaches have been developed.[347] These methods partition the PES into harmonic and anharmonic components, based on the assumption that the dominant source of error arises from the harmonic term. Accordingly, the harmonic force field, which provides the dominant contribution to vibrational frequencies, is evaluated at the higher QM level of theory (typically CCSD(T) or CSs), whereas anharmonic corrections are computed at a lower level of theory, denoted as QM' (usually DFT or MP2). Within this framework, the hybrid anharmonic frequency of each normal mode ν_i^{hyb} is obtained as the sum of the high-level harmonic term $\omega_i^{\text{harm, QM}}$ and the low-level anharmonic shift $\Delta\nu_i^{\text{anh, QM'}}$

$$\nu_i^{\text{hyb}} = \omega_i^{\text{harm, QM}} + \Delta\nu_i^{\text{anh, QM'}}. \tag{3.24}$$

This hybrid strategy significantly reduces the computational cost while preserving high accuracy, providing a practical and reliable route to compute anharmonic frequencies for small- and medium-sized molecular systems.

Chapter 4

Advancing junChS-F12: From Energetics to Geometries and Vibrational Frequencies

The first part of the research carried out during this PhD focuses on refining the junChS-F12 model,[90] already presented in Chapter 3. The junChS-F12 protocol, developed in our research group as the explicitly correlated extension of the junChS[89] within the ChS family,[67] was originally designed to provide accurate thermochemical data at a moderate computational cost. Although the initial formulation achieved an excellent balance between accuracy and efficiency, some aspects could still be improved. In particular, the use of the F12b approximation[235,239] for the CCSD-F12 computations and the conventional perturbative treatment of triple excitations limited the accuracy of the model, since the latter do not benefit from the F12 ansatz.

The refined model[184] adopts the CCSD(F12*)(T+) approach,[185] which combines the (F12*)[241] approximation with a size-consistent treatment of perturbative triples. Together, these improvements significantly reduce the BSIE associated with the (T) term and improve the internal consistency of the composite scheme. As a result, the new formulation preserves the computational efficiency of the original junChS-F12 approach while enhancing its accuracy.

The performance of the improved protocol was validated on accurate thermochemical data for molecules containing elements from the first three rows of the periodic table. Representative applications to non-covalent interactions, conformational, and tautomeric equilibria have confirmed the overall reliability and transferability of the new junChS-F12 scheme.[184]

Beyond energetics, the model is extended to the accurate prediction of equilibrium geometries and harmonic vibrational frequencies.

The energetic protocol can be coupled with two levels of structural accuracy. At the first, computationally efficient level, geometries are optimized using the revDSD-PBEP86-D3(BJ)[254,256,336] (revDSD) double-hybrid functional, which was shown to yield reliable structures and harmonic force fields for most applications.[148] At the second, higher-accuracy level,

geometries are refined by adding CV correlation corrections, evaluated by MP2-F12, to *fc*-CCSD(T)-F12b structures. This strategy avoids explicit CBS extrapolations while providing a robust and cost-effective route toward near-spectroscopic structural accuracy (see Chapter 3).

Harmonic frequencies computed at the *fc*-CCSD(T)-F12b/jun-cc-pV(T+*d*)Z level are likewise highly reliable without additional corrections. For both geometries and harmonic force fields, only the F12b and (T) options are currently available for closed- and open-shell systems. However, the influence of the (F12*) and (T+) refinements is expected to be much less significant than for energies.[348]

4.1 Methodological Framework

The MRCC program[349,350] was employed for the MP2-F12 and CCSD(F12*)(T+) frozen core calculations. DFT computations were carried out with the Gaussian package,[351] while all other calculations were performed using MOLPRO.[352,353] Default options of MRCC and MOLPRO were adopted for all explicitly correlated computations, assimilating the partially augmented "jun" basis sets to their fully augmented counterparts for the selection of the corresponding auxiliary basis sets.[90] All open-shell species were treated using the restricted open-shell formalism.

4.1.1 Energies

In its refined formulation, the total electronic energy of the junChS-F12 model is expressed as

$$E_{\text{junChS-F12}} = E^{\text{jun-cc-pV(T+}d\text{)Z}}_{\text{CCSD(F12*)(T+)}} + \Delta E^{\text{CBS}}_{\text{MP2-F12}} + \Delta E^{\text{CV}}_{\text{MP2-F12}} \tag{4.1}$$

where the leading term corresponds to the frozen core explicitly correlated CC energy computed with the jun-cc-pV(T+*d*)Z basis set.[268] The CBS correction is obtained from MP2-F12 energies computed with the jun-cc-pV(*n*+*d*)Z basis sets, with n = T and Q, using the two-point Helgaker formula[282]

$$\Delta E^{\text{CBS}}_{\text{MP2-F12}} = \frac{4^3 E^{\text{jun-cc-pV(Q+}d\text{)Z}}_{\text{MP2-F12}} - 3^3 E^{\text{jun-cc-pV(T+}d\text{)Z}}_{\text{MP2-F12}}}{4^3 - 3^3} - E^{\text{jun-cc-pV(T+}d\text{)Z}}_{\text{MP2-F12}}. \tag{4.2}$$

In this context, the CBS extrapolation can be performed in a single step for both the HF and MP2-F12 components, since in the former case the contribution arising from the CABS is included.[235] The core-valence correction is estimated as the difference between all-electron and frozen core MP2-F12 energies computed with the cc-pwCVTZ basis set[267]

$$\Delta E^{\text{CV}}_{\text{MP2-F12}} = E^{\text{cc-pwCVTZ}}_{\textit{ae}\text{-MP2-F12}} - E^{\text{cc-pwCVTZ}}_{\textit{fc}\text{-MP2-F12}}. \tag{4.3}$$

This hierarchical formulation provides a balanced and cost-effective description of electron correlation, ensuring a fast convergence toward the CBS limit and an accurate treatment of core-valence effects.

4.1.2 Geometries

Geometrical parameters are obtained according to the following expression

$$r_{\text{junChS-F12}} = r_{\text{CCSD(T)-F12b}}^{\text{jun-cc-pV(T+}d\text{)Z}} + \Delta r_{\text{MP2-F12}}^{\text{CBS}} + \Delta r_{\text{MP2-F12}}^{\text{CV}} \tag{4.4}$$

where, in analogy with Equations 4.2 and 4.3, the CBS and CV corrections are defined as

$$\Delta r_{\text{MP2-F12}}^{\text{CBS}} = \frac{4^3 r_{\text{MP2-F12}}^{\text{jun-cc-pV(Q+}d\text{)Z}} - 3^3 r_{\text{MP2-F12}}^{\text{jun-cc-pV(T+}d\text{)Z}}}{4^3 - 3^3} - r_{\text{MP2-F12}}^{\text{jun-cc-pV(T+}d\text{)Z}} \tag{4.5}$$

and

$$\Delta r_{\text{MP2-F12}}^{\text{CV}} = r_{ae\text{-MP2-F12}}^{\text{cc-pwCVTZ}} - r_{fc\text{-MP2-F12}}^{\text{cc-pwCVTZ}} . \tag{4.6}$$

Reference geometries are calculated either at the junChS-F12 level or by adding only the CV contribution, evaluated at the MP2-F12 level, to the CCSD(T)-F12b results. The latter model will be referred to in the following as $r_{\text{CCSD(T)-F12b}}^{\text{jun-cc-pV(T+}d\text{)Z}} + \Delta r_{\text{MP2-F12}}^{\text{CV}}$.

4.1.3 Harmonic Frequencies

The junChS-F12 approach can also be employed for the calculation of harmonic vibrational frequencies. In this case, the core-valence contribution can be safely neglected, since in explicitly correlated CCSD(T) calculations the omission of CV effects improves the results owing to a favorable error compensation with the missing higher-order excitations.[348] To ensure stable computations, the cc-pVTZ/JKFIT basis set[354] was employed for density fitting instead of its augmented counterpart, which is used for energy and geometry calculations.[355] When required, the model including only the CBS extrapolation and neglecting CV correlation will be denoted as $\omega_{\text{CCSD(T)-F12b}}^{\text{jun-cc-pV(T+}d\text{)Z}} + \Delta\omega_{\text{MP2-F12}}^{\text{CBS}}$.

4.2 Results and Discussion

4.2.1 Energetics

The impact of replacing the F12b and (T) models with their (F12*) and (T+) counterparts was examined using the Knizia–Adler–Werner (KAW) test set.[239] Subsequently, the performance of the new junChS-F12 formulation was assessed for the calculation of interaction energies in the non-covalent complexes of the A14 database,[90,356] as well as for evaluating the relative stabilities of the cytosine and guanine tautomers and rotamers and exploring the conformational landscape of glycine.

4.2.1.1 KAW test set

The KAW test set provides high-quality reference data obtained at the frozen core CCSD(T) level and extrapolated to the CBS limit with the aug-cc-pV(5,6)Z basis sets. It comprises 49 atomization energies and 28 and 48 reaction energies for closed- and open-shell systems, respectively. The performance of the new junChS-F12 model was compared with that of its F12b predecessor and with the conventional junChS formulation. Since the reference values of the test set do not include core-valence correlation, all ChS-type schemes reduce to their CC + CBS counterparts.

The errors associated with the different model chemistries are summarized in Table 4.1, where max unsigned errors (MAXs), mean unsigned errors (MUEs), and root-mean-square deviations (RMSDs) are reported.

Atomization energies represent some of the most challenging benchmarks in computational thermochemistry. In fact, the CCSD(T)/jun-cc-pV(T+*d*)Z approach yields rather unsatisfactory results, with a MUE of about 30 kJ mol^{-1}. Upon inclusion of the CBS extrapolation at the MP2 level, the error decreases by nearly an order of magnitude, approaching chemical accuracy (MUE $\approx$ 4 kJ mol^{-1}). This remarkable improvement, which lies at the core of the success of the junChS composite model, comes at negligible additional cost compared to the underlying CCSD(T) calculations.

The explicitly correlated CCSD(T)-F12b method enhances accuracy with respect to the conventional CCSD(T), yielding a MUE of approximately 6 kJ mol^{-1}. However, the corresponding CBS extrapolation proves much less effective than in conventional approaches. The final accuracy is slightly worse than that of the non-explicitly correlated scheme (MUE $\approx$ 5 kJ mol^{-1} compared to $\approx$ 4 kJ mol^{-1}), likely due to an unbalanced treatment of the various energetic contributions.

The CCSD(F12*)(T+) model provides a further and more consistent improvement, especially when combined with the CBS extrapolation performed at the MP2-F12 level, reaching a MUE of about 2.7 kJ mol^{-1}. The same trends are observed for both closed- and open-shell reaction energies, with the latter being particularly relevant as a measure of the robustness and generality of the proposed computational framework.

The overall error statistics for the KAW test set lie well within chemical accuracy and are comparable to those obtained using much larger basis sets specifically developed for explicitly correlated methods.[185] Finally, Tables A.1, A.2, and A.3 in Appendix A provide a systematic analysis about the effect of the adopted CBS extrapolation equation, showing that the n^{-3} performs better than alternative expressions.

TABLE 4.1: MAX, MUE, and RMSD values for the KAW test set with respect to the CCSD(T)/aug-cc-pV(5,6)Z reference data.[239] Results are obtained using the CCSD(T), CCSD(T)-F12b, and CCSD(F12*)(T+) methods combined with the jun-cc-pV(T+*d*)Z basis set, as well as their counterparts including the CBS correction. In all cases, the jun-cc-pV(*n*+*d*)Z basis sets with n = T and Q are employed for the extrapolation. All values are in kJ mol^{-1}.

	$E^{\text{jun-cc-pV(T+}d\text{)Z}}_{\text{CCSD(T)}}$	$E^{\text{jun-cc-pV(T+}d\text{)Z}}_{\text{CCSD(T)}}+\Delta E^{\text{CBS}}_{\text{MP2}}$	$E^{\text{jun-cc-pV(T+}d\text{)Z}}_{\text{CCSD(T)-F12b}}$	$E^{\text{jun-cc-pV(T+}d\text{)Z}}_{\text{CCSD(T)-F12b}}+\Delta E^{\text{CBS}}_{\text{MP2-F12}}$	$E^{\text{jun-cc-pV(T+}d\text{)Z}}_{\text{CCSD(F12*)(T+)}}$	$E^{\text{jun-cc-pV(T+}d\text{)Z}}_{\text{CCSD(F12*)(T+)}}+\Delta E^{\text{CBS}}_{\text{MP2-F12}}$
Atomization energy						
MAX	69.97	10.58	16.10	13.93	12.89	10.33
MUE	30.69	4.01	6.62	5.25	4.18	2.69
RMSD	33.50	4.77	7.37	6.03	5.08	3.78
Closed-shell reaction energy						
MAX	30.13	5.15	6.19	3.36	4.98	5.31
MUE	5.39	1.54	1.15	0.84	1.38	1.06
RMSD	8.29	2.11	1.74	1.20	1.95	1.72
Open-shell reaction energy						
MAX	70.10	11.57	7.75	13.67	6.40	6.57
MUE	18.47	3.02	3.05	2.83	2.02	2.14
RMSD	22.46	4.40	3.78	3.72	2.58	2.82
Overall						
MAX	70.10	11.57	16.10	13.93	12.89	10.33
MUE	20.33	3.08	4.02	3.33	2.72	2.11
RMSD	25.48	4.03	5.54	4.46	3.67	3.05

4.2.1.2 Non-covalent Interactions, Conformational Landscapes and Tautomeric Equilibria

After the new junChS-F12 model was validated for the thermochemistry of small molecules, it is essential to assess its performance for both inter- and intramolecular non-covalent interactions, as well as for tautomeric equilibria, which play a key role in many biological processes. The three benchmark tests employed for this purpose are the A14 set of non-covalent interaction energies (see Figure 4.1), the relative stabilities of glycine conformers (see Figure 4.2), and the relative energies of cytosine (see Figure 4.3) and guanine (see Figure 4.4) tautomers and rotamers. All conformers, tautomers, and rotamers are presented in order of decreasing stability, and their labeling follows References 169, 357, 358.

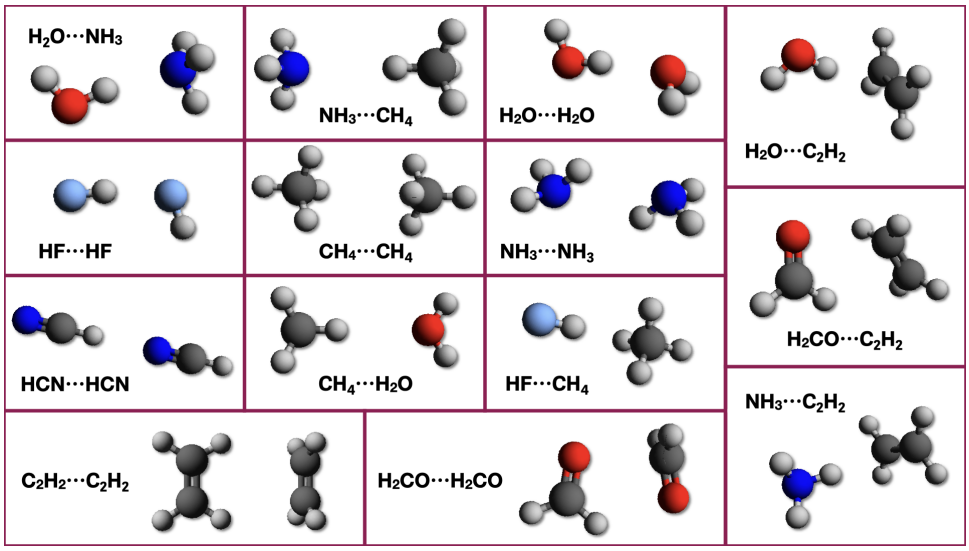


FIGURE 4.1: Non-covalent complexes of the A14 database.

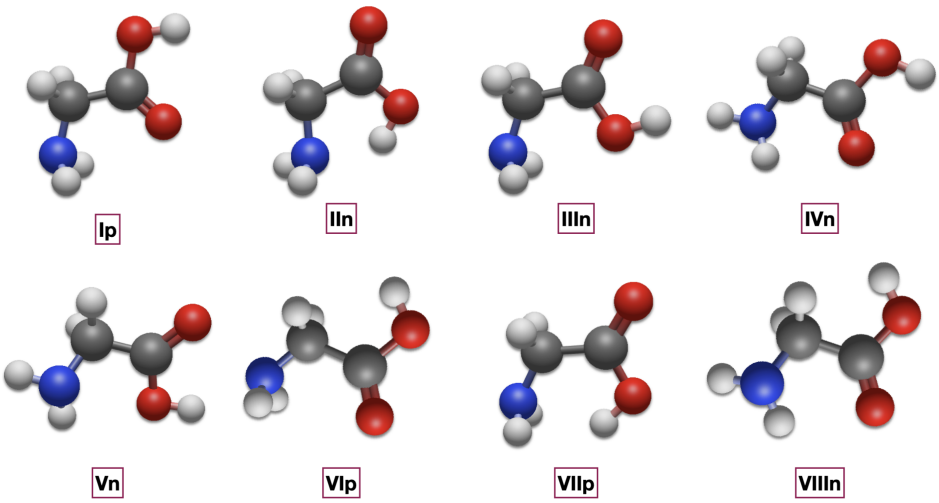


FIGURE 4.2: Structure of glycine conformers in order of relative stability evaluated at the refined junChS-F12 level of theory.

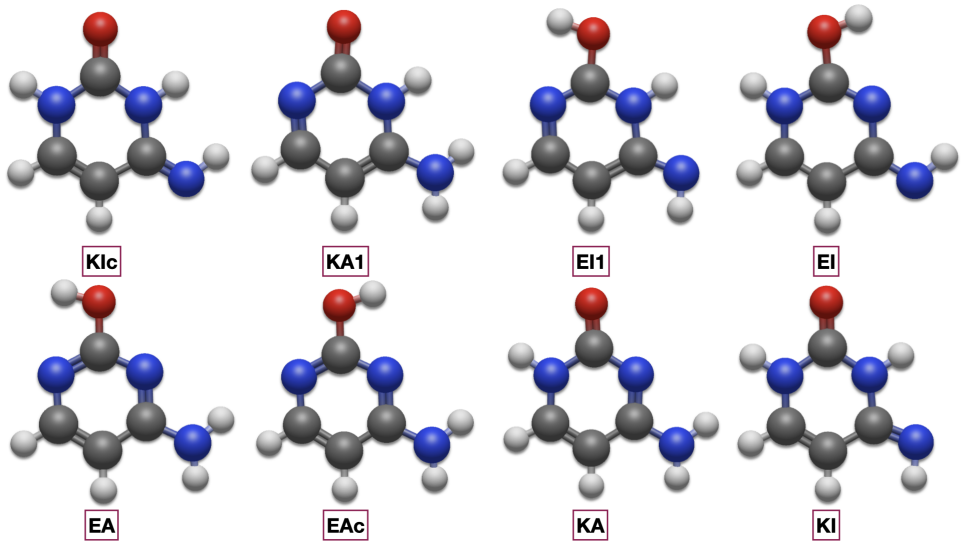


FIGURE 4.3: Structure of cytosine tautomers and rotamers in order of relative stability evaluated at the refined junChS-F12 level of theory.

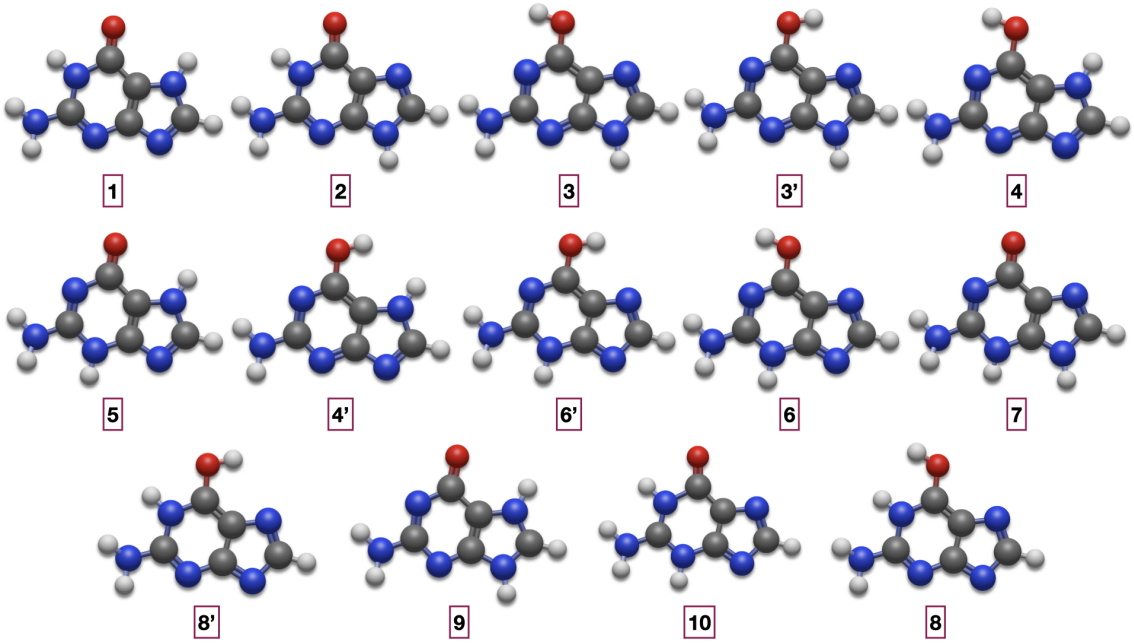
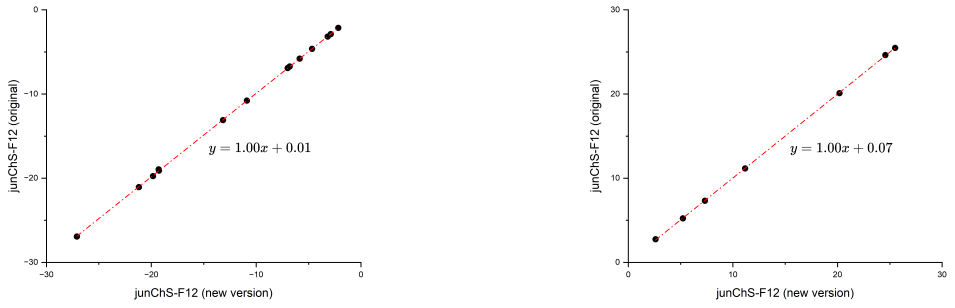


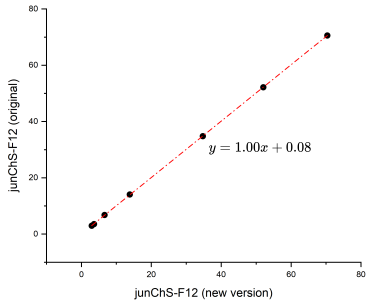
FIGURE 4.4: Structure of guanine tautomers and rotamers in order of relative stability evaluated at the refined junChS-F12 level of theory.

Figure 4.5 shows the correlation between the original and the improved junChS-F12 models, whereas a statistical analysis is given in Table 4.2.

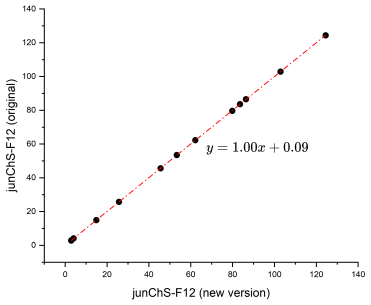


(A)

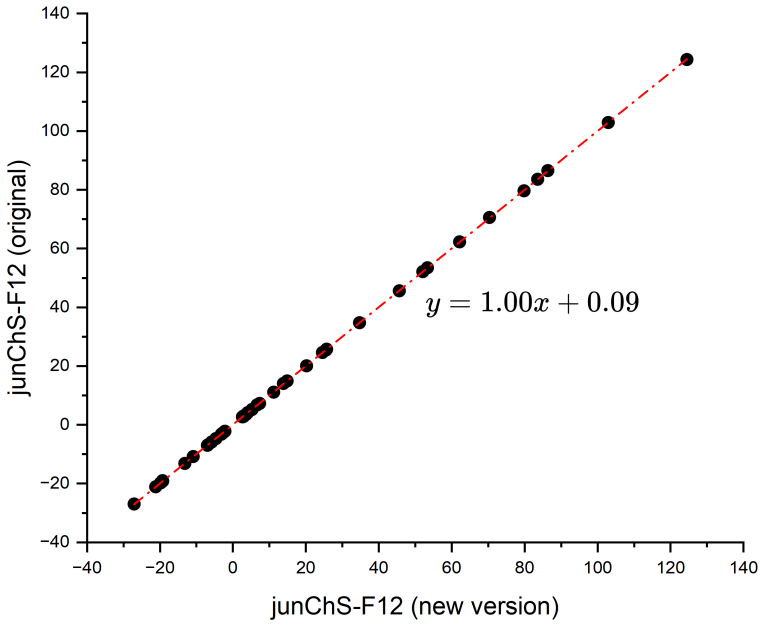
(B)



(C)



(D)



(E)

FIGURE 4.5: Linear correlation between interaction, conformational, and tautomerization energies (kJ mol^{-1}) obtained with the original and the revised version of the junChS-F12 model: (A) A14 database of non-covalent complexes, (B) glycine conformers, (C) cytosine tautomers and rotamers, (D) guanine tautomers and rotamers, (E) overall dataset.

TABLE 4.2: MAX, MUE, and mean deviation (MD) in kJ mol^{-1}, together with slope (a), intercept (b), and coefficient of determination (R^2) values of the linear regressions of the new junChS-F12 model with respect to its previous version.

	$y = ax + b$			MAX	MUE	MD
	a	b	R^2			
A14 interaction energies	1.00	0.01	1.00	0.35	0.10	0.10
Relative energies glycine conformers	1.00	0.07	1.00	0.15	0.05	0.02
Relative energies of cytosine tautomers and rotamers	1.00	0.08	1.00	0.26	0.13	0.12
Relative energies of guanine tautomers and rotamers	1.00	0.09	1.00	0.19	0.10	0.05
Overall	1.00	0.09	1.00	0.35	0.09	0.07

Table 4.2 and Figure 4.5 show a perfect correlation between the two schemes, with mean deviations of about 0.1 kJ mol^{-1} and maximum differences below 0.4 kJ mol^{-1}. These results demonstrate that the improved accuracy achieved for covalent interactions does not come at the expense of larger errors in the description of non-covalent interactions, a result that is far from trivial, particularly for the (T+) contribution. Since comparisons with previous state-of-the-art data have shown that the expected accuracy of the original junChS-F12 model lies well within 0.1 kJ mol^{-1},[90] the same level of reliability can be ascribed to the new version.

4.2.2 Geometrical Parameters

As already mentioned, the energetic protocol can be coupled with revDSD-optimized geometries using the jun-cc-pV(T+d)Z basis set, or with the higher-accuracy structural variant described above.

These two structural variants were validated by comparing their results with accurate semi-experimental equilibrium geometries for 32 molecules containing elements from the first three rows of the periodic table. Unless otherwise specified, all reference data were taken from the SE100 database.[148] The test set includes CH_4,[142] CO_2, HCN, HNC, H_2O, NH_3, C_2H_2, C_2H_4, H_2CO, CH_2O_2 (trans formic acid), CH_2NH, BH_3NH_3, C_2H_4O (oxirane), C_2H_4NH (aziridine), CH_2F_2, HOF, CH_2CHF, cyc$-C_3H_6$, NH_2OH, BH_2OH, HNO,[359] H_2O_2,[360] SO_2, H_2S,[361] PH_3, H_2CS, CH_2PH, C_2H_4S (thiirane), H_2S_2, HClCO, CH_2Cl_2, and CH_2ClF.

Error statistics for the various model chemistries are summarized in Table 4.3. Overall, the accuracy of revDSD geometries, typically within a few mÅ for bond lengths, is confirmed to be sufficient for most thermochemical applications. All post-HF approaches tested further reduce the deviations by a factor of three to four, achieving the precision required for the most demanding applications.

TABLE 4.3: MAX, MUE, and RMSD of geometrical parameters computed at the revDSD, CCSD(T), and CCSD(T)-F12b level using the jun-cc-pV(T+d)Z basis set. Results including the MP2 and MP2-F12 CV correction with the cc-pwCVTZ basis set, as well as those from the junChS and junChS-F12 models, are also reported. All values are given with respect to SE reference data. Bond lengths (d) are given in Å, and valence (θ) and dihedral (φ) angles in degrees.

	r_{revDSD}	$r_{\text{CCSD(T)}}$	$r_{\text{CCSD(T)}}+\Delta r^{\text{CV}}_{\text{MP2}}$	r_{junChS}	$r_{\text{CCSD(T)-F12b}}$	$r_{\text{CCSD(T)-F12b}}+\Delta r^{\text{CV}}_{\text{MP2-F12}}$	$r_{\text{junChS-F12}}$
	jun-cc-pV(T+d)Z	jun-cc-pV(T+d)Z	jun-cc-pV(T+d)Z cc-pwCVTZ		jun-cc-pV(T+d)Z	jun-cc-pV(T+d)Z cc-pwCVTZ	
2nd-row (22 molecules)							
MAX(d)	0.0093	0.0112	0.0067	0.0042	0.0045	0.0044	0.0045
MAX(θ,φ)	0.69	0.64	0.65	1.52	0.61	0.62	0.80
MUE(d)	0.0032	0.0048	0.0030	0.0008	0.0020	0.0006	0.0006
MUE(θ,φ)	0.16	0.17	0.13	0.10	0.12	0.07	0.08
RMSD(d)	0.0035	0.0053	0.0034	0.0014	0.0022	0.0012	0.0013
RMSD(θ,φ)	0.23	0.24	0.19	0.29	0.18	0.18	0.21
3rd-row (10 molecules)							
MAX(d)	0.0113	0.0209	0.0156	0.0020	0.0055	0.0021	0.0029
MAX(θ,φ)	0.38	0.29	0.34	0.10	0.23	0.22	0.63
MUE(d)	0.0044	0.0065	0.0038	0.0006	0.0024	0.0005	0.0006
MUE(θ,φ)	0.13	0.10	0.11	0.04	0.08	0.08	0.09
RMSD(d)	0.0051	0.0082	0.0054	0.0007	0.0028	0.0007	0.0009
RMSD(θ,φ)	0.16	0.13	0.13	0.05	0.10	0.10	0.16

While CBS extrapolation plays a crucial role in conventional approaches, this is not the case for F12-based models, where it can introduce small oscillations that slightly deteriorate the resulting geometries,[90] as illustrated in Table 4.3 in the transition from the CCSD(T) or CCSD(T)-F12b + CV schemes to the corresponding junChS or junChS-F12 formulations. In contrast, inclusion of the CV correlation remains essential for achieving accurate equilibrium structures: although its impact on valence and dihedral angles is minor, it significantly improves bond lengths, reducing the corresponding MUE by roughly a factor of two for both conventional and F12 methods.[149] The overall accuracy achieved is remarkable, with bond lengths benefiting from a consistent, though partly fortuitous, error compensation. The lengthening induced by the perturbative inclusion of quadruple excitations is, in most cases, counterbalanced by a comparable shortening arising from the differences between full and perturbative treatments of triple and quadruple excitations [T–(T) and Q–(Q)]. Among the molecules considered, the only exception to this general trend is the O-O bond in hydrogen peroxide, which is significantly underestimated. In this case, the unusually large lengthening caused by the perturbative treatment of quadruple excitations (0.0031 Å) cannot be compensated by the much smaller shortenings associated with the T–(T) (0.0002 Å) and Q–(Q) (0.0004 Å) contributions.[362]

4.2.3 Harmonic Vibrational Frequencies

The final step of the refinement of the junChS-F12 model chemistry is the accurate prediction of harmonic vibrational frequencies.

The benchmark set comprises 29 molecules containing elements from the first three rows of the periodic table, taken from Reference 348. Experimental harmonic frequencies were employed as references for the diatomics CO,[363] F_2,[363] HF,[363] N_2,[363] H_2,[75] $OH^\bullet$,[75]

Cl_2,[363] ClF,[364,365] CS,[363] HCl,[363] SiO,[365] and polyatomic molecules C_2H_2,[366–368] ClCN,[369] CO_2,[370] H_2O,[371] HCN,[372] and N_2O.[373] For NH_3,[374] C_2H_4,[375] CCl_2,[376] CF_2,[377] CH_4,[378] H_2CO,[379] H_2CS,[380] PH_3,[381] C_2H_2O,[382] and HNO,[383] reference frequencies were derived from computed force fields subsequently refined through variational or perturbational calculations for the best agreement with experimental spectra. Finally, for HOCl[384] and BH_3,[385] high-level *ab initio* calculations of near-spectroscopic quality were employed as reference data.

The CCSD(T)-F12b/jun-cc-pV(T+*d*)Z harmonic frequencies and their counterparts including the CBS extrapolation are compared in Table 4.4 with those obtained using the revDSD, CCSD(T)-F12a/jun-cc-pV(T+*d*)Z, and CCSD(T)-F12x/aug-cc-pV(T+*d*)Z (with x = a, b) models.

TABLE 4.4: MAX, MUE, and RMSD (cm^{-1}) of harmonic frequencies computed at the CCSDT(T)-F12x (x = a, b) level with aug-cc-pV(T+*d*)Z and jun-cc-pV(T+*d*)Z basis sets, the latter also including the CBS MP2-F12/jun-cc-pV((T+*d*),(Q+*d*))Z correction, and at the revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+*d*)Z level.

	$\omega_{\text{CCSDT(T)-F12x}}$[a]		$\omega_{\text{CCSDT(T)-F12x}} + \Delta\omega^{\text{CBS}}_{\text{MP2-F12}}$[a]	ω_{revDSD}
	aug-cc-pV(T+*d*)Z	jun-cc-pV(T+*d*)Z	jun-cc-pV(T+*d*)Z jun-cc-pV((T+*d*),(Q+*d*))Z	jun-cc-pV(T+*d*)Z
2nd-row (19 molecules)				
MAX	17.8 (18.4)	11.5 (12.3)	26.0 (27.1)	74.8
MUE	4.2 (4.0)	3.6 (3.4)	3.5 (3.4)	7.2
RMSD	5.3 (5.2)	4.5 (4.4)	5.1 (5.0)	12.5
3rd-row (10 molecules)				
MAX	13.6 (14.3)	12.3 (12.8)	13.8 (14.6)	22.2
MUE	3.7 (4.0)	3.5 (3.6)	3.6 (3.6)	8.4
RMSD	5.2 (5.4)	5.0 (5.0)	5.0 (5.3)	10.5

[a] F12b and (in parenthesis) F12a results.

Since the F12b and F12a models yield comparable errors (slightly smaller for F12b), the following discussion focuses on the former, which provides the lowest RMSD value when used with the jun-cc-pV(T+*d*)Z basis set. The inclusion of diffuse functions on first-row atoms and f-type diffuse functions on second- and third-row atoms yields no improvement in the calculated harmonic frequencies and even slightly degrades their accuracy. Hence, the consistent use of the jun-cc-pV(T+*d*)Z basis set is recommended for reliable vibrational frequency calculations.

Regarding the CBS extrapolation, it is not recommended for computing harmonic frequencies with F12-based methods. As already noted for geometries, this procedure does not improve the results and can introduce non-negligible numerical instabilities. The RMSD obtained with the CCSD(T)-F12b/jun-cc-pV(T+*d*)Z model is about 5 cm^{-1} for the entire dataset, a level of accuracy fully adequate for the assignment and interpretation of vibrational spectra. The revDSD functional also performs remarkably well, with an RMSD of about 10 cm^{-1},

a precision sufficient for most spectroscopic applications, particularly for the evaluation of zero-point energies and partition functions.

4.3 Conclusions

In this Chapter, the refinement of the junChS-F12 composite scheme has been presented, together with its validation against accurate reference data, with the overall goal of achieving accurate electronic energies, geometrical parameters, and harmonic frequencies for reliable characterizations of medium-sized molecules. The new energetic scheme replaces the CCSD(T)-F12b formulation with the more complete CCSD(F12*)(T+) approach, which substantially reduces the basis set incompleteness error without increasing the computational cost. Benchmark results demonstrated that this refinement achieves MUEs within chemical accuracy for atomization and reaction energies of both closed- and open-shell systems, while maintaining excellent performance for non-covalent interactions, conformational, and tautomeric equilibria, with deviations typically below 0.1 kJ mol^{-1} relative to the original junChS model. For equilibrium geometries, sub-milliångström accuracy can be reached by adding MP2-F12 core-valence corrections to explicitly correlated CCSD(T)/jun-cc-pV(T+*d*)Z optimizations, whereas CBS extrapolation can be safely neglected in F12-based schemes. Similarly, harmonic vibrational frequencies computed at the CCSD(T)/jun-cc-pV(T+*d*)Z level exhibit an RMSD of about 5 cm^{-1} across a diverse test set, ensuring reliable vibrational assignments without the need for additional corrections.

Chapter 5

Development of the PCS Family

Following the refinement of the junChS-F12 model, the second part of this PhD research focused on the development of the Pisa Composite Scheme family,[176,182] a new generation of composite schemes capable of improving the accuracy of previous ChS-type models while maintaining comparable computational demands.

The PCS protocols provide a unified and flexible framework for the evaluation of accurate energetic, structural, and vibrational properties of medium- and large-size molecular systems, ensuring an optimal balance between accuracy and computational cost. Their main features include the systematic use of cc–pVnZ–F12 basis sets,[269] separate CBS extrapolations for MP2 and post–MP2 contributions, the account for MP2 core-valence correlation in molecular geometries, and the explicit treatment of anharmonic effects in zero–point vibrational corrections and thermal contributions.

Although the protocol defines consistent levels of theory for structures and vibrational frequencies, its core lies in the energetic component, which constitutes the foundation of its methodological development. Building upon the original design of the ChS-type schemes, the valence correlation energy was redefined through a more sophisticated strategy to obtain a more reliable model.

The initial development focused on the conventional (non–F12) variant of the energetic scheme, with only a preliminary extension to its explicitly correlated counterpart.[176] In the latter version, the conventional valence terms were replaced by their F12 counterparts, specifically employing the CCSD(F12*)(T+) method, while keeping the overall structure of the scheme unchanged.

Subsequent efforts led to the refined explicitly correlated PCS model,[182] which achieved higher accuracy in electronic energies through improved CBS extrapolation strategies, optimized basis sets for third–row elements, and a more advanced treatment of CV correlation. These enhancements significantly extended the robustness and transferability of the protocol across a wide range of chemical systems.

Finally, the explicitly correlated PCS model was further extended to large molecules through the introduction of the FNO approximation. This approach markedly reduced the computational cost of CC correlation treatments while preserving accuracy, enabling the study of molecules containing up to about 50 atoms without resorting to local correlation techniques.

For equilibrium geometries, explicit CBS extrapolations were avoided by evaluating the CV correction at the MP2 level and adding it to the valence geometry obtained from explicitly correlated CC optimizations. Since analytical gradients are not available for the FNO implementation, low-cost yet accurate alternatives were obtained by combining double–hybrid functionals with MP2–based CV corrections, or by employing one–parameter approximations within the bond–corrected PCS model.[167,187,188]

Zero–point vibrational and thermal corrections were evaluated using double–hybrid or hybrid functionals, depending on the molecular size, while anharmonic effects were included via VPT2 for semi–rigid systems.[67]

The organization of the PCS family allows selecting the most appropriate compromise between accuracy and efficiency depending on the dimension and complexity of the molecular system. This Chapter focuses on molecules in the 5-15 and 15–60 atom ranges, for which the optimized protocols are summarized in Table 5.1, where the level of theory decreases from left to right with the derivative order and computational cost, and from top to bottom with the increasing number of atoms

Table 5.1: Overview of the PCS protocols for molecular systems of different dimensions.

Number of Atoms	Energetics	Geometry	Harmonic Frequencies	Anharmonic Contributions
5-15	PCS3	PCS2	revDSD–PBEP86–D3(BJ)/cc–pVTZ–F12	B3LYP–D3(BJ)/6–31+G*
15–60	FPCS3	BDPCS3	B3LYP–D3(BJ)/6–31+G*	Scaled Harmonic B3LYP–D3(BJ)/6–31+G*

The *n* label in the PCS*n* acronym refers to the cardinal number of the largest basis set used in the computation of the valence electronic energy, whereas the use of the FNO approximation is indicated by the prefix *F*, and *BD* stands for bond-corrected double-hybrid. The hierarchy of PCS models is designed to progressively reduce computational cost along the workflow, from electronic energies to geometries, harmonic frequencies, and anharmonic corrections.

The PCS3 variant, which employs CC energies with a triple–ζ basis set, serves as the workhorse for systems containing up to about 15 atoms. In this range, geometries are computed using the composite PCS2 scheme, which combines explicitly correlated CC optimizations with MP2–based CV corrections.

For larger molecules, the FNO–based schemes ensure a dramatic reduction in computational cost while preserving accuracy. In this size range, molecular structures are typically obtained using the BDPCS3,[167,188] while harmonic and anharmonic vibrational properties are evaluated with the B3LYP–D3(BJ)[254,256,330–332] functional instead of the double–hybrid revDSD–PBEP86–D3(BJ)[254,256,336] one, in conjunction with the smaller 6–31+G* basis set rather than the modified cc–pVTZ–F12$^-$ triple–ζ set (whose details are discussed below).

The PCS protocols were extensively benchmarked, demonstrating accuracy comparable to that of significantly more demanding methods and consistent performance for molecules containing first–, second–, and third–row elements. Following this validation, the models

were successfully applied to representative "molecular bricks of life" and to systems of current fundamental and technological relevance.

5.1 The PCS Model

5.1.1 Energetics

Considering the generic formulation of a ChS scheme, the total electronic energy can be expressed as the sum of valence and core-valence contributions

$$\begin{aligned} E_{\mathrm{ChS}} &= E^{\mathrm{T}}_{\mathrm{CCSD(T)}} + \Delta E^{\mathrm{CBS,(T,Q)}}_{\mathrm{MP2}} + \Delta E^{\mathrm{CV,T}}_{\mathrm{MP2}} \\ &= \underbrace{E^{\mathrm{T}}_{\mathrm{CCSD(T)}} + E^{\mathrm{CBS,(T,Q)}}_{\mathrm{MP2}} - E^{\mathrm{T}}_{\mathrm{MP2}}}_{\text{valence}} + \underbrace{\Delta E^{\mathrm{CV,T}}_{\mathrm{MP2}}}_{\text{core–valence}} . \end{aligned} \tag{5.1}$$

where the methods are reported in their conventional form, and only the cardinal numbers of the basis sets are specified to keep the Equation general.

In this formulation, the correction to the CBS extrapolation is made explicit, and the cardinal numbers of the basis sets are indicated as superscripts, making Equation 5.1 applicable to all ChS schemes, with conventional methods taken as reference.

By reorganizing the terms, taking the MP2 extrapolation as the leading contribution instead of the CCSD(T) evaluation and grouping the remaining valence corrections together, thus distinguishing between the low–level (MP2) and high–level (post–MP2) contributions to the electronic energy, the same expression can be rewritten as

$$\begin{aligned} E_{\mathrm{ChS}} &= E^{\mathrm{CBS,(T,Q)}}_{\mathrm{MP2}} + \left(E^{\mathrm{T}}_{\mathrm{CCSD(T)}} - E^{\mathrm{T}}_{\mathrm{MP2}} \right) + \Delta E^{\mathrm{CV,T}}_{\mathrm{MP2}} \\ &= \underbrace{E^{\mathrm{CBS,(T,Q)}}_{\mathrm{MP2}}}_{\text{low–level}} + \underbrace{\Delta E^{\mathrm{T}}_{\mathrm{CCSD(T)}}}_{\text{high–level}} + \underbrace{\Delta E^{\mathrm{CV,T}}_{\mathrm{MP2}}}_{\text{low–level}} . \end{aligned} \tag{5.2}$$

The PCS model originates from the general formulation given in Equation 5.2, introducing two main innovations: (i) the use of cc–pVnZ–F12 basis sets,[269] and (ii) a different estimation of the valence correlation energy.

In the PCS schemes, the MP2 extrapolation with triple- and quadruple-ζ basis sets is retained, while the post–MP2 valence correction is, unlike in the ChS model, extrapolated to the CBS limit using double- and triple-ζ basis sets. This modification provides a more reliable estimate of the valence contribution with only a negligible increase in computational cost, since the additional double–ζ computation is much cheaper than the triple–ζ one.

The first model developed within this framework was the non–F12 PCS3-A scheme (referred to simply as PCS in References 176, 186, and 177). Compared to the ChS schemes described in Chapter 3, PCS3-A retains the same treatment of the CV contribution and

employs the two–point Helgaker formula[282] for all CBS extrapolations, preserving the design philosophy of the ChS family. The number "3" in the model name refers to the CCSD(T)/cc–pVTZ–F12 calculation used in the high–level valence term.

Following the same procedure adopted for extending junChS to its junChS-F12 counterpart, the PCS3-B variant (denoted PCS-F12 in References 176 and 186) represents the first explicitly correlated extension of PCS3-A, where CCSD(F12*)(T+) and MP2-F12 replace their conventional analogues in the valence contribution.

The refinement of PCS3-B resulted in the definition of the PCS3 model, which at the time of this writing represents the reference energetic scheme within the PCS family for molecular systems containing 5-15 atoms. With respect to PCS3-B, three main methodological improvements were introduced in the PCS3 formulation.

- The modified two–point Helgaker extrapolation with an n^{-5} dependence was adopted for the high–level valence term, replacing the previous n^{-3} expression. This choice ensures a smoother and faster convergence toward the CBS limit.

- The standard cc–pVDZ–F12 basis set was replaced by its modified cc–pVDZ–F12^{+} counterpart (see below), providing an improved description of third–row elements.

- The CV contribution, previously evaluated at the MP2 level as the difference between all-electron and frozen core calculations with the cc–pwCVTZ basis, was refined by partitioning it into MP2 and post–MP2 components, following the same philosophy adopted for the valence energy. The MP2 part is extrapolated to the CBS limit using the n^{-3} two–point formula with double– and triple–ζ basis sets, whereas the post–MP2 correction is computed at the CCSD(T) level with a double–ζ basis. Although this formulation results in a marginal increase in computational cost, since the additional MP2 and CCSD(T) calculations are performed with a double–ζ basis, it yields a more accurate treatment of CV effects.

The rationale for improving the valence and CV treatments will be discussed in the following.

Building upon the PCS3 formulation, the FPCS3 model integrated the FNO approximation into all CC calculations, encompassing both the valence and CV contributions. This refinement maintained the methodological consistency of the original scheme while extending its applicability to large molecular systems. The standard cc–pVDZ–F12 basis was used for the high–level valence correction. In all variants, the CV contribution is consistently evaluated using conventional methods together with cc–pwCVnZ basis sets.[267]

The main characteristics of the PCS3 variants are summarized in Table 5.2, which details the computational methods, basis sets, and extrapolation exponents employed in the evaluation of the valence and core-valence contributions.

TABLE 5.2: Definition of the PCS3 composite schemes for electronic energy evaluations: computational methods, basis sets, and extrapolation exponents used for valence and core-valence terms.

	Valence		Core–Valence	
	low–level	high–level	low–level	high–level
PCS3–A	$E^{\mathrm{CBS}}_{\mathrm{MP2}}$ cc–pV(T,Q)Z–F12 n^{-3}	$\Delta E^{\mathrm{CBS}}_{\mathrm{CCSD(T)}}$ cc–pV(D,T)Z–F12 n^{-3}	$\Delta E^{\mathrm{CV}}_{\mathrm{MP2}}$ cc–pwCVTZ	–
PCS3-B	$E^{\mathrm{CBS}}_{\mathrm{MP2–F12}}$ cc–pV(T,Q)Z–F12 n^{-3}	$\Delta E^{\mathrm{CBS}}_{\mathrm{CCSD(F12*)(T+)}}$ cc–pV(D,T)Z–F12 n^{-3}	$\Delta E^{\mathrm{CV}}_{\mathrm{MP2}}$ cc–pwCVTZ	–
PCS3	$E^{\mathrm{CBS}}_{\mathrm{MP2–F12}}$ cc–pV(T,Q)Z–F12 n^{-3}	$\Delta E^{\mathrm{CBS}}_{\mathrm{CCSD(F12*)(T+)}}$ cc–pV(D$^+$,T)Z–F12 n^{-5}	$\Delta E^{\mathrm{CV,CBS}}_{\mathrm{MP2}}$ cc–pwCV(D,T)Z n^{-3}	$\Delta E^{\mathrm{CV}}_{\mathrm{CCSD(T)}}$ cc–pwCVDZ
FPCS3	$E^{\mathrm{CBS}}_{\mathrm{MP2–F12}}$ cc–pV(T,Q)Z–F12 n^{-3}	$\Delta E^{\mathrm{CBS}}_{\mathrm{FNO–CCSD(F12*)(T+)}}$ cc–pV(D,T)Z–F12 n^{-5}	$\Delta E^{\mathrm{CV,CBS}}_{\mathrm{MP2}}$ cc–pwCV(D,T)Z n^{-3}	$\Delta E^{\mathrm{CV}}_{\mathrm{FNO–CCSD(T)}}$ cc–pwCVDZ

A detailed analysis of the PCS3 scheme is particularly relevant, as a significant part of this PhD work was devoted to its development and to the implementation of the corresponding FNO-based extension. The energy expression for the model can be written as

$$E_{\mathrm{PCS3}} = E^{\mathrm{CBS}}_{\mathrm{MP2\text{-}F12}} + \Delta E^{\mathrm{CBS}}_{\mathrm{CCSD(F12*)(T+)}} + \Delta E^{\mathrm{CV,CBS}}_{\mathrm{MP2}} + \Delta E^{\mathrm{CV}}_{\mathrm{CCSD(T)}}. \tag{5.3}$$

The valence MP2-F12 energy, $E^{\mathrm{CBS}}_{\mathrm{MP2\text{-}F12}}$, is extrapolated to the CBS limit using the two-point Helgaker formula:[282]

$$E^{\mathrm{CBS}}_{\mathrm{MP2\text{-}F12}} = \frac{4^3 E^{\text{cc-pVQZ-F12}}_{\mathrm{MP2\text{-}F12}} - 3^3 E^{\text{cc-pVTZ-F12}}_{\mathrm{MP2\text{-}F12}}}{4^3 - 3^3}. \tag{5.4}$$

The valence post-MP2-F12 correlation energy correction, $\Delta E^{\mathrm{CBS}}_{\mathrm{CCSD(F12*)(T+)}}$, accounting for the CC correlation, is evaluated through a modified two-point Helgaker extrapolation employing an exponent of 5 and the CCSD(F12*)(T+) ansatz[185]

$$\Delta E^{\mathrm{CBS}}_{\mathrm{CCSD(F12*)(T+)}} = \frac{3^5 \left(E^{\text{cc-pVTZ-F12}}_{\mathrm{CCSD(F12*)(T+)}} - E^{\text{cc-pVTZ-F12}}_{\mathrm{MP2\text{-}F12}} \right) - 2^5 \left(E^{\text{cc-pVDZ-F12}^+}_{\mathrm{CCSD(F12*)(T+)}} - E^{\text{cc-pVDZ-F12}^+}_{\mathrm{MP2\text{-}F12}} \right)}{3^5 - 2^5} \tag{5.5}$$

where the superscript + on cc-pVDZ-F12$^+$ denotes the addition of a single set of f functions (taken from the cc-pVTZ basis set) to the standard cc-pVDZ-F12 for third-row atoms.

The core-valence MP2 contribution, $\Delta E^{\mathrm{CV,CBS}}_{\mathrm{MP2}}$, is extrapolated to the CBS limit using the same two-point Helgaker expression:

$$\Delta E_{\mathrm{MP2}}^{\mathrm{CV,CBS}} = \frac{3^3\left(E_{ae\text{-MP2}}^{\text{cc-pwCVTZ}} - E_{fc\text{-MP2}}^{\text{cc-pwCVTZ}}\right) - 2^3\left(E_{ae\text{-MP2}}^{\text{cc-pwCVDZ}} - E_{fc\text{-MP2}}^{\text{cc-pwCVDZ}}\right)}{3^3 - 2^3}. \tag{5.6}$$

The CV post-MP2 correlation energy correction, $\Delta E_{\mathrm{CCSD(T)}}^{\mathrm{CV}}$, represents the CCSD(T)–MP2 difference

$$\Delta E_{\mathrm{CCSD(T)}}^{\mathrm{CV}} = \left(E_{ae\text{-CCSD(T)}}^{\text{cc-pwCVDZ}} - E_{fc\text{-CCSD(T)}}^{\text{cc-pwCVDZ}}\right) - \left(E_{ae\text{-MP2}}^{\text{cc-pwCVDZ}} - E_{fc\text{-MP2}}^{\text{cc-pwCVDZ}}\right). \tag{5.7}$$

It should be noted that both the cardinal numbers of the basis sets and the exponents employed in the CBS extrapolations could, in principle, be further optimized.[47] Although different exponents have been proposed in the literature for the extrapolation of correlation energies,[272,386] the present formulation adopts a distinct partitioning scheme in which the HF and MP2-F12 valence correlation components are extrapolated together, while the post-MP2-F12 term is treated separately. This unconventional choice, employing exponents of 3 and 5 for the MP2 and post-MP2 extrapolations as defined in Equations 5.4 and 5.5, provides highly accurate results.

The computational levels selected for each contribution within the PCS3 model were chosen to ensure a balanced compromise between accuracy and efficiency, particularly for molecules containing up to approximately 15 atoms. In this respect, the computational cost of the CCSD(F12*)(T+)/cc–pVDZ–F12$^+$ and MP2-F12/cc–pVQZ–F12 calculations is negligible compared with that of the corresponding CCSD(F12*)(T+)/cc–pVTZ–F12 computations. The estimation of the CV correlation at the MP2 level, as originally defined in the PCS3-A formulation, is sufficiently accurate for most properties, although it may be less reliable for atomization energies. To ensure that the evaluation of the (relatively small) CV contribution does not exceed the computational cost of the dominant valence term, the cc–pwCVDZ basis set has been adopted for the CC step, yielding remarkably consistent results.

For larger systems, comprising up to about 50 atoms, single–point CCSD(F12*)(T+) energy evaluations in conjunction with the cc-VTZ-F12 basis set can be efficiently performed thanks to the implementation of FNO, NAF, and related reduced–scaling techniques.[93–95]

Benchmark studies have demonstrated that FNO–based approximations reproduce full CC results with excellent accuracy while reducing computational time by roughly a factor of five.[95] Consequently, systems containing more than about 15 atoms can be treated using the FPCS3 protocol, in which the FNO and related approximations are applied to both the valence (explicitly correlated[95]) and CV (conventional[93]) components. The truncation error associated with the FNO threshold is mitigated by taking into account the full MP2 contribution and including the PPL correction.[94]

The density fitting approximation was employed for both the HF and post–HF steps, using the aug–cc–pV(n+1)Z–RI–JK[387] and aug–cc–pwCV(n+1)Z–RI[388] fitting basis sets, respectively. The cc–pVnZ–OPTRI auxiliary basis sets were adopted for the computation of the CABS, consistently applied to both the cc–pVTZ–F12 and cc–pVDZ–F12$^+$ (for third–row elements) basis sets.

Slater–type f_{12} correlation factors, each fitted with six Gaussian functions, were used with exponents of 0.9, 1.0, and 1.1 for the cc–pVDZ–F12 (and cc–pVDZ–F12$^+$), cc–pVTZ–F12, and cc–pVQZ–F12 basis sets, respectively.[389]

In the FNO computations, the occupation number thresholds for the retained MP2 virtual orbitals were set to $5.0 \cdot 10^{-5}$ for conventional calculations, $1.0 \cdot 10^{-4}$ for the cc–pVDZ–F12 basis set, and $1.0 \cdot 10^{-5}$ for the cc–pVTZ–F12 basis set.

All conventional calculations were carried out with the Gaussian suite of programs,[351] whereas the F12 and FNO computations were performed using the MRCC code.[349,350]

5.1.2 Molecular Structures

The geometry scheme within the PCS model expresses each geometrical parameter (r) as the sum of a valence (V) and a CV contribution

$$r = r^{\mathrm{V}} + \Delta r^{\mathrm{CV}}_{\mathrm{MP2}} \tag{5.8}$$

where the level of theory used for the valence term depends on the molecular size, and thus on the trade-off between accuracy and computational cost, while the CV correction is computed as

$$\Delta r^{\mathrm{CV}}_{\mathrm{MP2}} = r^{\text{cc-pwCVTZ}}_{ae\text{-MP2}} - r^{\text{cc-pwCVTZ}}_{fc\text{-MP2}} \tag{5.9}$$

It is worth noticing that, while the cc-pwCVTZ basis set slightly underestimates core correlation effects,[390] the MP2 method introduces a small, opposite overcorrelation,[391] resulting in a favorable error compensation and residual deviations below 0.001 Å.

For molecules in the 5-15 atom range, the valence contribution starts from a geometry optimization performed at the CCSD(T)–F12b level of theory with the cc–pVDZ–F12 basis set. In the present scheme, the CBS extrapolation is not performed, as explicitly correlated methods significantly accelerate basis set convergence, and it has been shown that highly accurate geometries can be obtained without further extrapolation when the core-valence correction is included.[90,184]

The choice of the cc–pVDZ–F12 basis set, despite its limited size, is supported by previous benchmarks showing that CCSD(T)–F12/cc–pVDZ–F12 geometries reproduce the accuracy of conventional CCSD(T) calculations with quadruple–ζ basis sets.[161] Considering the CV correction, this result has been confirmed for molecules containing first– and second–row atoms. However, the errors increase when third–row elements are involved, as discussed in Chapter 4. The use of the cc–pVDZ–F12$^+$ basis set restores uniform accuracy across first–, second–, and third–row atoms, as it will be shown in the following.

Based on the employed levels of theory, the PCS2 scheme is obtained for molecules containing 5 - 15 atoms, and Equation 5.8 reads as

$$r_{\mathrm{PCS2}} = r^{\text{cc-pVDZ-F12}^+}_{\text{CCSD(T)-F12b}} + \Delta r^{\mathrm{CV}}_{\mathrm{MP2}}. \tag{5.10}$$

Moving to the 15–60 atom range, analytical gradients are not yet available for the reduced–cost FNO variants of CC methods, which prevents their use in geometry optimizations. Consequently, for molecules larger than about a dozen atoms, a low–cost version of the PCS model is adopted. In this scheme, the valence part of Equation 5.8 replaces the explicitly correlated CC model with the revDSD double–hybrid functional, used in conjunction with a slightly modified triple–ζ F12 basis set. The latter was obtained from the cc-pVTZ-F12 basis set by removing d functions on first–row atoms and replacing the two f functions on second– and third–row atoms with a single f function taken from the cc–pVTZ basis set. The resulting basis set, which significantly reduces computational cost without appreciable loss of accuracy, was called cc–pVTZ–F12^{-}.

The resulting scheme, termed DPCS3 (where *D* indicates the use of a double–hybrid functional in place of a CC method), follows the general philosophy of the PCS family, which aims to systematically optimize the cost–to–performance ratio at each computational step and for every molecular size. In this case, while the level of theory is reduced from CCSD(T)–F12b to the revDSD double–hybrid functional to handle larger systems, the basis set is simultaneously expanded from double– to triple–ζ quality to preserve the overall accuracy.

The inclusion of the core-valence correction $\Delta r^{\mathrm{CV}}_{\mathrm{MP2}}$ into DPCS3 yields the MDPCS3[392] scheme (denoted sd MPCS1 in Reference 182, PCS/DFT in Reference 176, PCS in Reference 186), in which the geometrical parameters are obtained according to the following expression

$$r_{\mathrm{MDPCS3}} = r^{\text{cc-pVTZ-F12}^{-}}_{\mathrm{revDSD}} + \Delta r^{\mathrm{CV}}_{\mathrm{MP2}}. \tag{5.11}$$

The same $\Delta r^{\mathrm{CV}}_{\mathrm{MP2}}$ correction ensures accurate geometries for both PCS2 and MDPCS3, without introducing any additional empirical parameters beyond those of the parent functional.[392]

A more economical variant of the MDPCS3 scheme has also been proposed,[182,191,392] in which the $\Delta r^{\mathrm{CV}}_{\mathrm{MP2}}$ correction is replaced by a one–parameter bond correction ($\Delta r^{\mathrm{CV}}_{\mathrm{B}}$). This simplification is justified because CV effects are negligible for valence and dihedral angles, while they are significant only for bond lengths, as discussed in Chapter 4. An additional correction to the valence contribution, Δr^{del}, is also introduced to account for the tendency of double–hybrid functionals to overestimate bond delocalization due to residual self–interaction errors. The resulting model, denoted BDPCS3 (where *B* refers to the one–parameter bond correction and *D* to the underlying double–hybrid optimization), is defined as

$$r_{\mathrm{BDPCS3}} = r^{\text{cc-pVTZ-F12}^{-}}_{\mathrm{revDSD}} + \Delta r^{\mathrm{del}} + \Delta r^{\mathrm{CV}}_{\mathrm{B}}. \tag{5.12}$$

The BDPCS3 scheme (referred to as BDPCS3 in Reference 182, BPCS1 in Reference 189) has been validated on a broad set of molecular systems of increasing size,[188–192] demonstrating its reliability and efficiency. In this PhD work, the model is used for comparison purposes.

Geometry optimizations at the MP2 and revDSD levels were performed with the Gaussian package,[351] whereas those required at the explicitly correlated level for the PCS2 model

were carried out using MOLPRO.[352,353] Default F12 settings were employed, considering the cc–pVDZ–F12$^+$ basis set equivalent to its standard cc–pVDZ–F12 counterpart.

5.1.3 Vibrational Frequencies and Thermodynamic Functions

The relative stability of molecular species is determined by enthalpies or free energies, which combine electronic energies with contributions from ZPVEs and entropic terms. For semi-rigid molecules, anharmonic effects can be reliably accounted for by second-order vibrational perturbation theory, provided that accurate QC models are used to compute harmonic, cubic, and semi-diagonal quartic force constants. Several studies have demonstrated that revDSD-PBEP86-D3(BJ) harmonic force fields in conjunction with triple-ζ basis sets are highly reliable[148,393] and can be routinely applied to molecules containing up to a dozen atoms. Consequently, for medium-sized systems, harmonic frequencies were obtained at the revDSD double-hybrid level in conjunction with the cc-pVTZ-F12 basis set, while anharmonic corrections were efficiently evaluated at the B3LYP-D3(BJ) (B3) level with the 6-31+G* basis set. For larger molecules, harmonic frequencies were computed at the same hybrid level, and anharmonicity was incorporated through the use of suitable scaling factors.

5.2 Results and Discussion

The validity of the proposed hierarchy was demonstrated by systematically testing PCS models on increasingly complex molecular systems, from benchmark sets to biologically and technologically relevant compounds.

This section is divided into three main parts. The first part is devoted to energetic investigations, including both benchmark studies and applications carried out with the PCS3-A and PCS3-B models, followed by the analysis of the PCS3 and FPCS3 schemes. The second part focuses on molecular geometries, with the validation of structural parameters and their application to the calculation of rotational constants. The third part covers the evaluation of harmonic vibrational frequencies and thermodynamic functions.

5.2.1 Application of PCS3-A and PCS3-B to Energetic Evaluations

In the following, the performance of the PCS3-A and PCS3-B models for thermochemical applications is first assessed using the Knizia–Adler–Werner test set.[239] The analysis is then extended to reaction barrier heights, inter- and intramolecular non-covalent interactions, as well as to conformational energies and to the relative stability of different tautomeric forms; finally, application to the characterization of the PES for the H_2S + Cl hydrogen abstraction reaction is presented.

When required, experimental SO coupling values are incorporated into the PCS3-A and PCS3-B schemes for O, OH, SH, and Cl radicals, reducing their electronic energies by 0.9, 0.8, 2.3, and 3.5 kJ mol^{-1}, respectively.[75]

5.2.1.1 Thermochemistry

The first assessment of the PCS3-A and PCS3-B variants was carried out on the KAW test set. Since the reference data for this benchmark do not include CV correlation, this contribution was omitted from the composite approaches. The resulting errors are summarized in Table 5.3 and compared with those obtained from the underlying CCSD(T) and CCSD(F12*)(T+) calculations using the jun-cc-pV(T+*d*)Z and cc-pVTZ-F12 basis sets, as well as from the junChS and the refined junChS-F12 schemes, in terms of MAX, MUE, and RMSD values.

TABLE 5.3: MAX, MUE, and RMSD values for the KAW test set with respect to the CCSD(T)/aug-cc-pV(5,6)Z reference data.[239] Results are obtained using CCSD(T) and CCSD(F12*)(T+) methods combined with the jun-cc-pV(T+*d*)Z and cc-pVTZ-F12 basis sets, as well as the junChS, junChS-F12, and PCS3-A/B CSs. All values are in kJ mol^{-1}.

	$E_{\mathrm{CCSD(T)}}$		junChS	PCS3-A	$E_{\mathrm{CCSD(F12*)(T+)}}$		junChS-F12	PCS3-B
	jun-cc-pV(T+*d*)Z	cc-pVTZ-F12			jun-cc-pV(T+*d*)Z	cc-pVTZ-F12		
Atomization energy								
MAX	69.97	53.67	10.58	8.76	12.89	9.85	10.33	8.36
MUE	30.69	25.00	4.01	2.90	4.18	3.71	2.69	2.27
RMSD	33.50	27.21	4.77	3.60	5.08	4.49	3.78	3.35
Closed-shell reaction energy								
MAX	30.13	35.68	5.15	7.57	4.98	5.96	5.31	3.61
MUE	5.39	7.32	1.54	1.56	1.38	1.04	1.06	0.83
RMSD	8.29	9.99	2.11	2.18	1.95	1.64	1.72	1.24
Open-shell reaction energy								
MAX	70.10	50.48	11.57	6.30	6.40	7.19	6.57	6.26
MUE	18.47	14.65	3.02	1.89	2.02	1.80	2.14	1.69
RMSD	22.46	17.99	4.40	2.52	2.58	2.55	2.82	2.38
Overall								
MAX	70.10	53.67	11.57	8.76	12.89	9.85	10.33	8.36
MUE	20.33	17.07	3.08	2.21	2.72	2.38	2.11	1.73
RMSD	25.48	20.90	4.03	2.93	3.67	3.32	3.05	2.63

For atomization energies, all conventional CCSD(T) approaches yield MUEs of about 25 - 30 kJ mol^{-1}, and inclusion of the CBS correction decreases the error by nearly one order of magnitude. Similar trends are observed for both closed- and open-shell reaction energies. Comparison with junChS over the entire test set shows improved performance for PCS3-A: the MUE decreases by about 30% (from 3.08 to 2.21 kJ mol^{-1}), and the corresponding RMSD (2.93 kJ mol^{-1}) lies well within the chemical accuracy threshold. Similar trends are observed for the explicitly correlated calculations, with PCS3-B showing smaller MUE and RMSD values for the overall test set compared to junChS-F12.

The benchmark results confirm the improved accuracy achieved by the PCS3-A and PCS3-B variants compared to the junChS and junChS-F12 schemes. The KAW test set represents a particularly suitable validation case, since the lack of CV correlation allows a direct assessment of the impact of the improved valence energy evaluation. This aspect is especially relevant because the treatment of CV correlation remains unchanged across the different CSs, so any observed improvement can be directly attributed to the refined description of the valence contribution.

5.2.1.2 Reaction Barriers

Having validated the PCS3-A and PCS3-B models on atomization and reaction energies, the next step was to assess the reliability for evaluating activation energies. To this end, a widely used benchmark set for reaction barrier heights is provided by the DBH24 dataset,[99,394] which reports reference values obtained with the W4 composite method.[11] This compilation includes a statistically representative selection of reactions, comprising three examples for each of the following classes: heavy-atom transfer, nucleophilic substitution, unimolecular and association reactions, and hydrogen transfer.

Table 5.4 compares the reaction barriers computed with PCS3-A to those obtained with the junChS scheme.

TABLE 5.4: Barrier heights (including SO corrections) from the DBH24 dataset computed with the junChS and PCS3-A schemes, together with reference values. All values are in kJ mol^{-1}.

	Reference[a,b]	junChS[c,d]	PCS3-A[c]
Heavy-Atom Transfer			
$H^\bullet + N_2O \longrightarrow OH^\bullet + N_2$[e]	71.67/345.05	73.35/348.32	72.58/347.69
$H^\bullet + ClH \longrightarrow HCl + H^\bullet$	75.31/75.31	72.43/72.43	72.54/72.54
$CH_3^\bullet + FCl \longrightarrow CH_3F + Cl^\bullet$[e]	28.24/251.04	29.96/252.59	29.82/253.69
Nucleophilic Substitution			
$Cl^-\cdots CH_3Cl \longrightarrow ClCH_3\cdots Cl^-$	56.11/56.11	55.48/55.48	56.57/56.57
$F^-\cdots CH_3Cl \longrightarrow FCH_3\cdots Cl^-$	14.39/123.09	14.18/121.71	14.54/122.51
$OH^- + CH_3F \longrightarrow HOCH_3 + F^-$	-10.21/73.89	-10.38/72.63	-11.99/72.59
Unimolecular and Association			
$H^\bullet + N_2 \longrightarrow HN_2^\bullet$	60.08/44.39	60.00/46.53	59.81/46.16
$H^\bullet + C_2H_4 \longrightarrow C_2H_5^\bullet$	7.20/174.68	7.95/176.52	7.81/176.80
$HCN \longleftrightarrow HNC$	201.12/137.32	200.75/139.08	201.06/137.67
Hydrogen Transfer			
$OH^\bullet + CH_4 \longrightarrow CH_3^\bullet + H_2O$[f]	28.07/82.01	27.74/83.85	27.72/81.82
$H^\bullet + OH^\bullet \longrightarrow H_2 + {}^3O$[e,f]	44.81/54.89	48.16/57.61	45.69/56.36
$H^\bullet + H_2S \longrightarrow H_2 + HS^\bullet$[e]	15.15/72.51	15.48/75.06	15.50/75.16
MAX[g]		3.36/3.32 (3.36)	2.77/2.77 (2.77)
MUE[g]		1.04/1.99 (1.52)	0.85/1.58 (1.21)
RMSD[g]		1.49/2.12 (1.83)	1.15/1.85 (1.54)

[a] Barrier heights for the forward/reverse reaction. [b] Data taken from Reference 394. [c] At revDSD/jun-cc-pV(T+*d*)Z geometries. [d] Data taken from Reference 395. [e] SO contributions included on the reverse reaction barrier. [f] SO contributions included on the forward reaction barrier. [g] Forward/reverse barrier height errors and overall values in parentheses.

The PCS3-A scheme provides a total MUE of about 1.2 kJ mol^{-1}, which is slightly lower than the value obtained with the junChS model (around 1.5 kJ mol^{-1}). This level of accuracy is comparable to that of other composite methods with similar computational demands. For instance, the W1X protocol[396] yields an average error of about 1.3 kJ mol^{-1}, while the F12 SVECV-f12 model[68] achieves a very similar accuracy, with a MUE of approximately 1.1 kJ mol^{-1}.

5.2.1.3 Non-covalent Intermolecular Interactions

The non-covalent complexes of the A14 set[89] are shown in Figure 4.1, and their interaction energies, computed with the junChS and PCS3-A schemes, are reported in Table 5.5 together with reference data.[356] The latter were obtained at the CBS CCSD(T) level of theory, including corrections for CV, Rel, and higher-order excitations up to CCSDT(Q). Error statistics are also provided.

TABLE 5.5: Interaction energies of the A14 complexes computed at the junChS and PCS3-A levels of theory using revDSD/jun-cc-pV(T+*d*)Z geometries, together with reference data and MAX, MUE, and RMSD values. All values are given in kJ mol^{-1}.

	Reference[a]	junChS[b]	PCS3-A
$H_2O\cdots H_2O$	-21.08	-21.37	-21.22
$NH_3\cdots NH_3$	-13.21	-13.34	-13.43
$HF\cdots HF$	-19.22	-19.59	-19.59
$HCN\cdots HCN$	-19.98	-19.70	-19.96
$CH_4\cdots CH_4$	-2.23	-2.22	-2.27
$CH_2O\cdots CH_2O$	-18.93	-19.43	-19.41
$C_2H_4\cdots C_2H_4$	-4.60	-4.78	-4.76
$H_2O\cdots C_2H_4$	-10.77	-11.01	-10.77
$H_2O\cdots CH_4$	-2.82	-2.79	-2.78
$H_2O\cdots NH_3$	-27.38	-27.69	-27.59
$NH_3\cdots CH_4$	-3.22	-3.24	-3.22
$NH_3\cdots C_2H_4$	-5.79	-5.99	-5.88
$HF\cdots CH_4$	-6.92	-7.14	-7.12
$C_2H_4\cdots CH_2O$	-6.79	-7.07	-7.05
MAX		0.50	0.48
MUE		0.22	0.16
RMSD		0.26	0.21

[a]Data taken from Reference 356. [b]Data from Reference 90.

The PCS3-A scheme yields a MUE of 0.16 kJ mol^{-1}, even without applying any BSSE correction. Compared to the junChS model, which provides a MUE of 0.22 kJ mol^{-1}, PCS3-A yields a slight improvement.

5.2.1.4 Conformational Analysis of Biochemical Building Blocks

After testing non-covalent intermolecular interactions, the attention was focused on conformational and tautomeric equilibria, starting from simple amino acids and progressing to more complex biochemical building blocks.

The first analysis focused on the conformational landscape of glycine, whose conformers are shown in Figure 4.2. The corresponding relative energies, evaluated using the PCS3-A scheme, are reported in Table 5.6 together with reference data.[397] The latter were obtained at the CCSD(T)-F12b level with a quadruple-ζ basis set, including CV and Rel corrections, as well as contributions from full triples and perturbative quadruples excitations. Error statistics are also provided. For comparison, revDSD/cc-pVTZ-F12 energies are also reported.

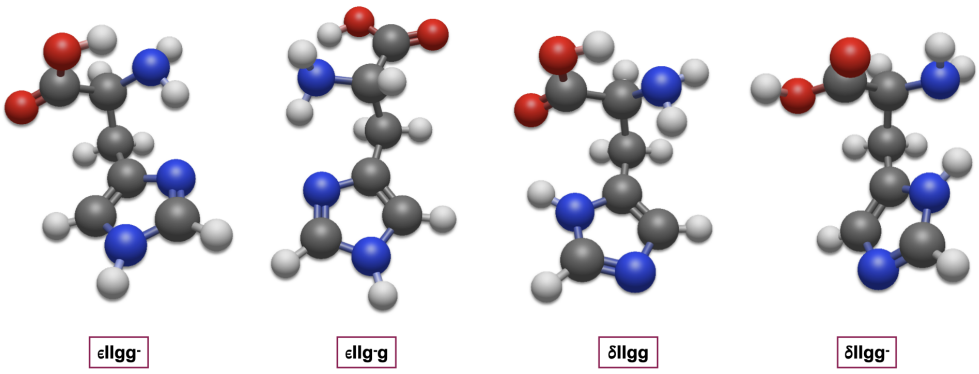
TABLE 5.6: Relative electronic energies of the glycine conformers computed at the PCS3-A and revDSD/cc-pVTZ-F12 levels, along with reference data. All values are in kJ mol^{-1}.[a]

	Reference[a]	$E^{\text{cc-pVTZ-F12}}_{\text{revDSD}}$	PCS3-A
Ip	0.0	0.00	0.00
IIn	2.68	2.49	2.66
IIIn	7.24	7.05	7.36
IVn	5.19	5.25	5.12
Vn	11.09	11.02	11.17
VIp	20.08	20.11	20.16
VIIp	24.44	24.20	24.63
VIIIn	25.36	25.51	25.40
MAX		0.24	0.19
MUE		0.12	0.08
RMSD		0.14	0.09

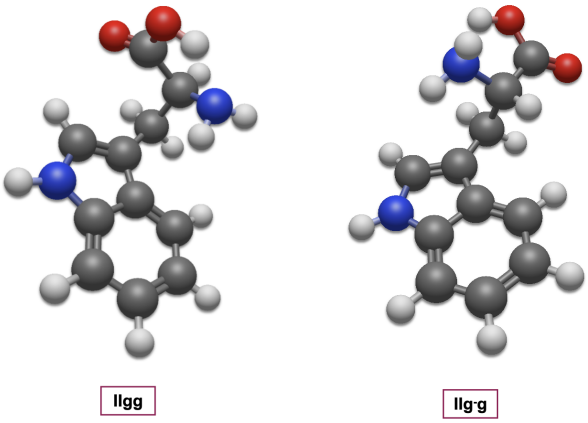
[a] Data taken from Reference 397

Several studies have reported the presence of up to eight distinct minima on the glycine conformational PES.[165] Comparison with the highly accurate reference data from Reference 397 indicates that the PCS3-A model reproduces the conformational energy landscape with excellent accuracy. This agreement is not unexpected, since conformational equilibria are usually well described even by lower-level QC methods,[165] as also confirmed by the revDSD results.

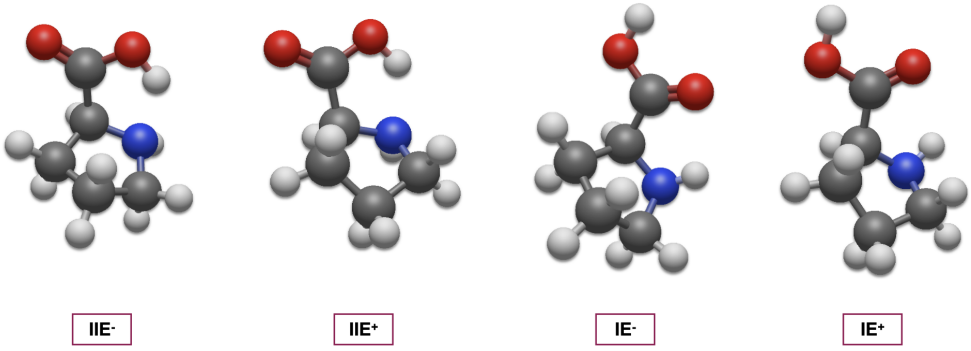
The PCS3-A model was then applied to obtain high-level reference data for the relative stabilities of more complex amino acids, characterized by distinctive structural features, including different tautomeric forms in histidine, ring puckering in proline, and condensed non-equivalent heteroaromatic structures in tryptophan. The analyzed conformers are shown in Figure 5.1 and follow the nomenclature reported in Reference 177. Their relative energies, computed at the revDSD/cc-pVTZ-F12 and PCS3-A levels of theory, are collected in Table 5.7.



(A) Histidine.



(B) Tryptophan.



(C) Proline.

FIGURE 5.1: Low-energy minima of histidine (A), tryptophan (B), and proline (C) considered in the present work.

TABLE 5.7: Relative electronic energies of histidine, tryptophan, and proline conformers obtained at the revDSD/cc-pVTZ-F12 and PCS3-A levels of theory. All values are in kJ mol^{-1}.

	$E_{\text{revDSD}}^{\text{cc-pVTZ-F12}}$	PCS3-A
Histidine		
ϵIIgg$^-$	0.00	0.00
ϵIIg$^-$g	4.73	4.95
δIIgg	1.97	1.90
δIIgg$^-$	12.02	11.90
Tryptophan		
IIgg	0.00	0.00
IIg$^-$g	4.22	4.67
Proline		
IIE$^-$	0.00	0.00
IIE$^+$	2.60	2.39
IE$^-$	7.48	6.71
IE$^+$	7.24	6.94

In this case, the results are consistent with those obtained for glycine (Table 5.6). It can therefore be concluded that, for conformational analyses, the double-hybrid functional revDSD in conjunction with the cc-pVTZ-F12 basis set provides results in close agreement with the higher-level PCS3-A model, as these properties are not strongly affected by the level of theory employed.

The accurate description of tautomeric equilibria plays a key role in understanding the behavior of nucleic acid bases, since these processes directly affect biochemical processes such as base-pairing mechanisms. The number of possible tautomers (N_T) of a given species is given by $N_T = N_S!/[N_H!(N_S - N_H)!]$, where N_S is the number of tautomeric sites and N_H the number of labile protons. In this context, the reliability of QC methods in assessing the relative stability of tautomers becomes particularly critical. To this end, the tautomeric equilibria of cytosine and guanine are investigated.

Cytosine possesses two endocyclic (N1 and N3) and two exocyclic (O7 and N8) tautomeric sites, along with two labile protons, giving rise to a total of six possible tautomers. These can be classified as keto–amino (KA and KA1), enol–amino (EA), keto–imino (KI), and enol–imino (EI and EI1) forms. Each enol and imino tautomer additionally exhibits a rotameric variant, denoted by a subscript *c*, corresponding to rotation about the O–H or N–H bond.[169,398,399] All structures are shown in Figure 4.3, ordered by decreasing stability, and the corresponding relative energies computed at the revDSD/cc-pVTZ-F12, CCSD(T)/jun-cc-pV(T+*d*)Z, junChS-F12, PCS3-A, and PCS3-B levels of theory, are reported in Table 5.8. The Table also lists MAX and MUE values with respect to PCS3-B, which is taken as the reference given its slightly better performances over PCS3-A for the KAW test set previously discussed.

TABLE 5.8: Relative electronic energies of cytosine tautomers and rotamers obtained at the revDSD/cc-pVTZ-F12, CCSD(T)/jun-cc-pV(T+*d*)Z, junChS-F12, PCS3-A, and PCS3-B levels of theory. All values are in kJ mol^{-1}.

	$E^{\text{cc-pVTZ-F12}}_{\text{revDSD}}$	$E^{\text{jun-cc-pV(T+}d\text{)Z}}_{\text{CCSD(T)}}$	junChS-F12	PCS3-A	PCS3-B
EA	0.00	0.00	0.00	0.00	0.00
EAc	3.07	2.86	2.99	3.00	3.01
KA	-1.25	4.74	3.58	3.13	3.08
KI	2.22	7.05	6.82	6.07	6.22
KIc	9.58	13.98	14.10	13.21	13.46
KA1	29.28	35.01	34.81	34.47	34.45
EI1	51.09	51.36	52.18	51.60	52.00
EI	70.33	69.61	70.58	69.78	70.33
MAX[a]	5.17	1.66	0.64	0.40	
MUE[a]	3.06	0.73	0.38	0.15	

[a] With respect to PCS3-B computations.

DFT methods tend to over-stabilize the KA tautomer,[399] a behavior also observed in this case for the revDSD/cc-pVTZ-F12 level of theory. Accurate relative stabilities are obtained only with CCSD(T)-based approaches that include both CBS and CV corrections. Experimental estimates derived from the relative intensities of infrared bands[398] indicate that, in the gas phase, the EA tautomer is the most stable, while the KA and EAc forms are less stable and nearly isoenergetic. These findings are consistent with the PCS3 results. Overall, the junChS-F12, PCS3-A, and PCS3-B models yield comparable energies, which differ within approximately 0.5 kJ mol^{-1}.

Guanine possesses four endo (N1, N3, N7, N9) and two exo (O=C and NH_2) tautomeric sites, along with three labile protons, giving rise to a total of 20 possible tautomers. Among these, ten are amino and ten are imino species; however, the latter are not considered here, as they are significantly less stable than their amino counterparts. For each of the two non-equivalent imidazole-ring structures (N7H and N9H), two keto–amino (KA) and one enol–amino (EA) tautomers can be formed. In the enol–amino series, both trans and cis conformations of the hydroxyl group (the latter denoted by a *c*) correspond to energy minima. The various species are labeled using the same two-letter code adopted for cytosine, followed by one number (for EA forms) or two numbers (for KA forms) indicating the positions of the remaining acidic hydrogen atoms. All computational methods consistently predict that only four tautomers, KA17, KA19, EA9, and EAc9, have appreciable gas-phase populations.[400] These species, along with the EA7 tautomer, labeled 1, 2, 3, 3′, and 4 in Figure 4.4, are depicted in Figure 5.2. The corresponding relative energies, computed at the revDSD/cc-pVTZ-F12, junChS-F12, PCS3-A, and PCS3-B levels of theory, are reported in Table 5.9.

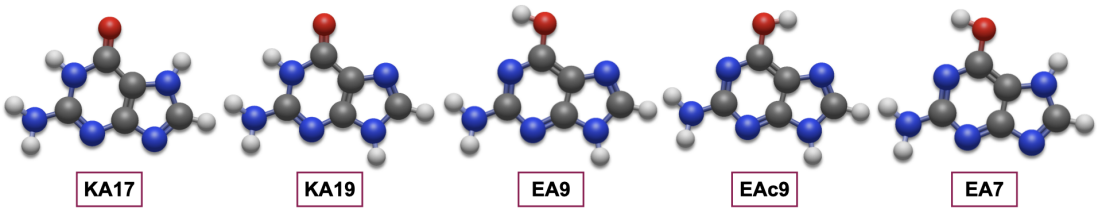


FIGURE 5.2: Structure of lower energy guanine tautomers and rotamers in order of relative stability.

TABLE 5.9: Relative electronic energies of guanine tautomers and rotamers obtained at the revDSD/cc-pVTZ-F12, junChS-F12, PCS3-A, and PCS3-B levels of theory. All values are in kJ mol^{-1}.

	$E^{\text{cc-pVTZ-F12}}_{\text{revDSD}}$	junChS-F12[a]	PCS3-A[a]	PCS3-B[a]
KA17	0.00	0.00	0.00	0.00
KA19	2.41	2.90	2.84	2.84
EA9	6.95	2.96	3.10	3.29
EAc9	8.08	4.05	4.13	4.35
EA7	19.37	14.88	15.11	15.29
MAX[a]	4.08	0.41	0.22	
MUE[a]	2.98	0.28	0.15	

[a] With respect to PCS3-B computations.

As it can be seen, the revDSD functional, despite providing the correct energy ordering, overstabilizes the KA19 tautomer by about 0.4 kJ mol^{-1} with respect to the composite methods, and most notably it overestimates the relative energy of EA9, EAc9, and EA7 by about 4 kJ mol^{-1}. On the other hand, all the CSs yield similar results with junChS-F12 and PCS3-A agreeing, on average, to PCS3-B within 0.28 kJ mol^{-1} and 0.15 kJ mol^{-1}, respectively.

It can therefore be concluded that reliable high-level methods are essential to accurately describe tautomeric equilibria, especially for species that can be detected by MW spectroscopy. Furthermore, PCS3 methods provide reliable reference data that can be used to validate lower-level computational approaches.

5.2.1.5 H_2S + Cl reactive PES

To extend the validation to reactive systems involving radical intermediates, the PCS3-A model was applied to the H_2S + Cl hydrogen abstraction reaction, a prototypical atmospheric process. As illustrated in Figure 5.3, the mechanism begins with the non-covalent interaction between hydrogen sulfide (H_2S) and chlorine (Cl), forming a pre-reactive Van der Waals complex (RW). Under atmospheric conditions, the reaction proceeds through an addition–elimination pathway. Specifically, isomerization of the RW complex via the transition state (TS) leads to the formation of a post-reactive van der Waals complex (PW), stabilized by an HS–HCl hydrogen bond. Dissociation of this complex yields the final products, HS

and HCl. The overall rate constant is governed by the interconversion between RW and PW and by the formation of RW, while decomposition of PW has a negligible effect, being considerably faster than the other elementary steps.[156,176]

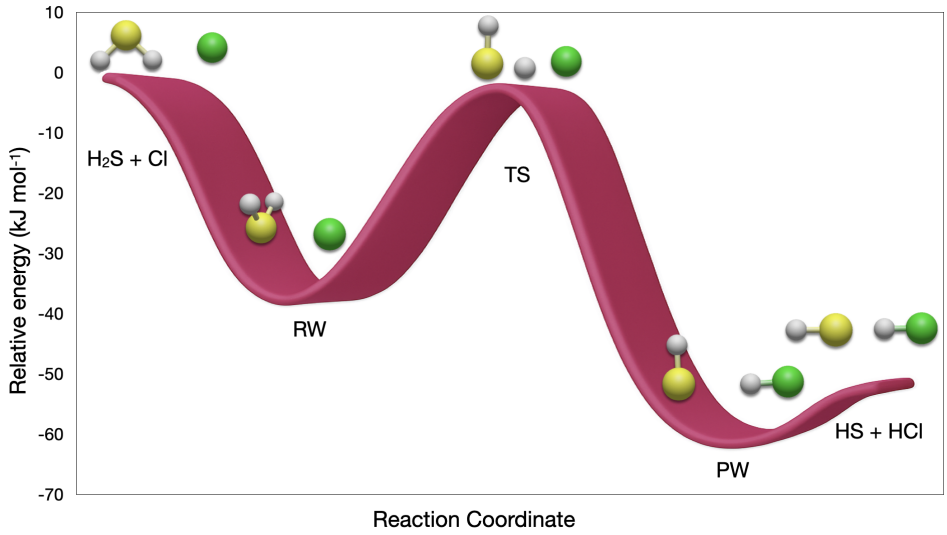


FIGURE 5.3: PES for the $H_2S + Cl \longrightarrow HS + HCl$ reaction.

The relative electronic energies of the stationary points on the reactive PES computed with the HEAT-like and PCS3-A approaches are reported in Table 5.10.

TABLE 5.10: Relative electronic energies of the stationary points for the $H_2S + Cl$ reaction, including spin-orbit corrections,[156] computed using the HEAT-like and PCS3-A schemes. All values are given in kJ mol^{-1}.

	Reactants	RW	TS	PW	Products
Level of theory	$H_2S + Cl$	$H_2S\cdots Cl$	$HS\cdots H\cdots Cl$	$HS\cdots HCl$	$HS + HCl$
HEAT-like[a]	0.00	-41.86	-0.05	-56.79	-46.25
PCS3-A	0.00	-40.07	0.36	-56.93	-46.28

[a] Data taken from Reference 156.

The PCS3-A model chemistry provides an excellent overall performance, with deviations generally within 0.5 kJ mol^{-1} from HEAT-like data for all the species but the pre-reactive complex, despite its moderate computational cost and the absence of corrections for higher-order excitations, Rel, and DBOC effects.

5.2.2 PCS3 and FPCS3 for Accurate Thermochemical Predictions

The refinement of the PCS3-B model, leading to the final PCS3 formulation, began with benchmarking against the W4-11 dataset,[83] where valence and core-valence contributions were analyzed separately. From this extensive assessment, the optimal design was identified: an extrapolation exponent of n^{-5} for the post-MP2-F12 valence term, the modification of the

standard cc-pVDZ-F12 basis set, and the improved treatment of the CV contribution. The resulting refined model was further validated against the larger W4-17 test set,[84] together with its FNO extension, FPCS3. Both models were then applied to selected molecular systems, including the reassessment of the relative stabilities of cytosine and guanine tautomers and rotamers, and the calculation of formation enthalpies for DNA and RNA nucleobases and a panel of polycyclic aromatic hydrocarbons (PAHs) of increasing size.

5.2.2.1 Atomization Energy Benchmarks

The selection of the levels of theory entering the PCS3 method was based on the W4-11 benchmark set of accurate atomization energies,[83] for which W1-F12 and W2-F12 reference data are also available.[47]

The first part of the study focused on improving the valence extrapolation strategy, with particular attention to the post-MP2-F12 component. Following several preliminary investigations and drawing on previous studies conducted within the research group,[90] the standard exponent of 3 was retained for the MP2-F12 extrapolation, consistent with the ChS schemes. For the post-MP2-F12 contribution, two exponents were examined: 3, as in the conventional PCS3-A model and its first explicitly correlated extension PCS3-B, and 5, adopted as an alternative to evaluate its influence on convergence and accuracy.

Subsequent efforts focused on improving the basis sets employed in the post-MP2-F12 valence extrapolation, which relies on double- and triple-ζ levels. To enhance accuracy without substantially increasing computational cost, additional high-angular-momentum functions were considered: an f function was added to cc-pVDZ-F12 and a g function to cc-pVTZ-F12, taken from the cc-pVTZ and cc-pVQZ basis sets, respectively.

During the tuning procedure, tests were performed by augmenting either both basis sets or only the cc-pDVZ-F12 basis. In both cases, n^{-3} and n^{-5} extrapolation procedures were investigated. The best overall performance across the W4-11 test set, based on all statistical indicators (MAX, MUE, and RMSD), was obtained with an extrapolation exponent of 5 for the high-level valence term and modification of the cc-pVDZ-F12 basis only. This configuration yielded the lowest MAX and an RMSD of 0.87 kJ mol^{-1}, only 0.01 kJ mol^{-1} higher than that obtained when also augmenting cc-pVTZ-F12 with g functions. This result is particularly significant, as it achieves nearly identical accuracy while considerably reducing the computational cost.

Finally, the impact of replacing the CCSD(T)-F12b model with the CCSD(F12*)(T+) formulation was assessed. As indicated by the results obtained for the junChS-F12 scheme employing these two explicitly correlated CC variants, the latter reduces both the MUE and RMSD by approximately 75%. Details about the different tests performed for tuning PCS3 and FPCS3 composite methods are provided in Table A.4 in Appendix A.

A more in-depth analysis of the individual PCS3 energy contributions was then carried out. A comparison of the valence energy errors obtained with the W1-F12, W2-F12, and PCS3 models is shown in Figure 5.4, where molecules containing first- and second-row atoms are

distinguished from those including third-row elements. The results for the CV contributions, computed at different levels of theory, are presented in Figure 5.5. Finally, RMSD of total electronic energies obtained from the W1-F12, W2-F12, SVECV-f12, junChS-F12, PCS3, and FPCS3 composite schemes are compared in Figure 5.6.

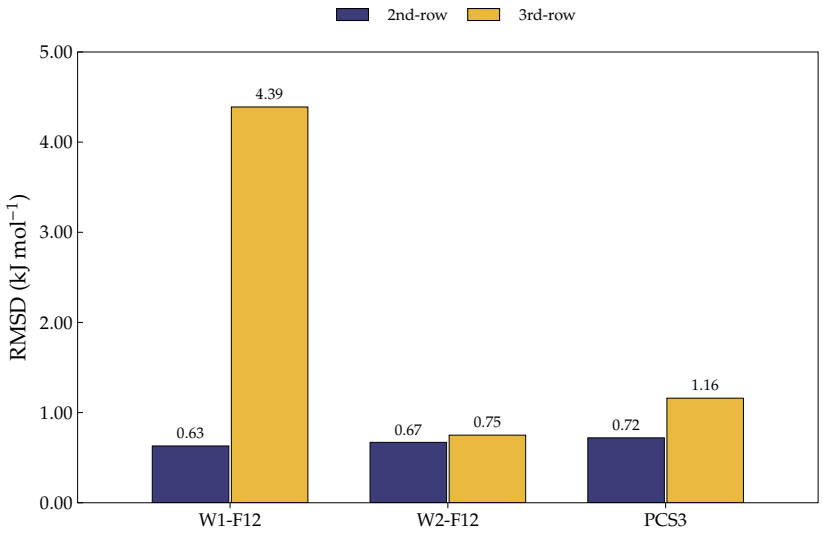


FIGURE 5.4: RMSDs, in kJ mol^{-1}, of valence energies for the W4-11 set obtained with the W1-F12, W2-F12, and PCS3 composite schemes.

Concerning the valence energies, the W1-F12 model exhibits a noticeable deterioration in accuracy on moving from molecules containing first- and second-row atoms to those containing third-row atoms. This behavior is not observed for the PCS3 scheme, which achieves consistently reliable results with a comparable computational cost, as both methods employ similar theoretical levels and basis sets. In addition, PCS3 performs similarly to the more computationally demanding W2-F12 model, which uses larger basis sets, for molecules with elements from the first two rows of the periodic table, while presenting only a marginal deterioration in the performances (about 0.4 kJ mol^{-1}) when third-row elements are involved.

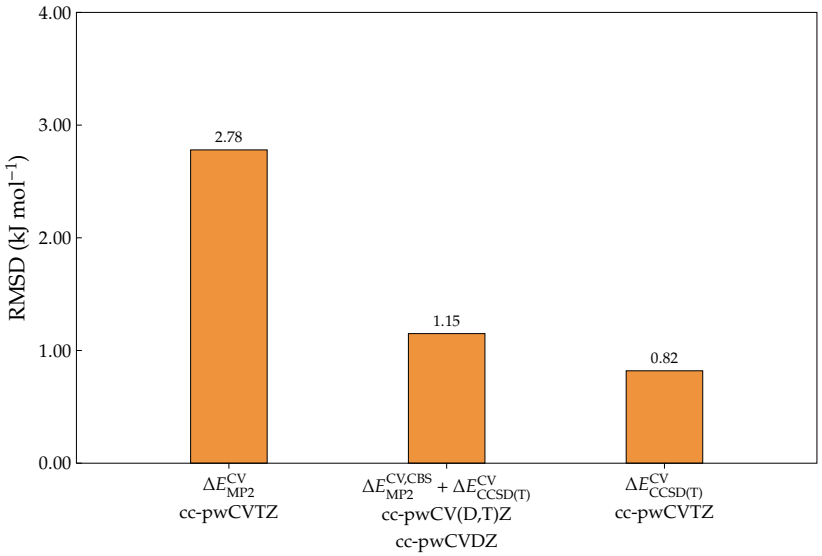


FIGURE 5.5: RMSDs, in kJ mol^{-1}, of CV energies for the W4-11 set, obtained using different models: ΔE^{CV}_{MP2} with the cc-pwCVTZ basis set, $\Delta E^{CV,CBS}_{MP2} + \Delta E^{CV}_{CCSD(T)}$ with the cc-pwCV(D,T)Z and cc-pwCVDZ basis sets, and $\Delta E^{CV}_{CCSD(T)}$ with the cc-pwCVTZ basis set.

In the case of the core-valence contribution, the MP2 estimate alone, as employed in the PCS3-A and PCS3-B schemes, leads to excessively large errors when focusing on atomization energies. These are significantly reduced when a CBS extrapolation at the MP2 level is combined with CCSD(T)/cc-pwCVDZ results (see Equations 5.6 and 5.7), yielding accuracies comparable to those obtained from a full CCSD(T) treatment with a triple-ζ basis set. This approach therefore achieves a balanced compromise between computational cost and accuracy, since the additional CC calculation is performed with a double-ζ basis without significantly increasing the overall computational effort required for the evaluation of this contribution. Details about the analysis of CV contributions evaluated at different levels of theory are summarized in Table A.5 in Appendix A.

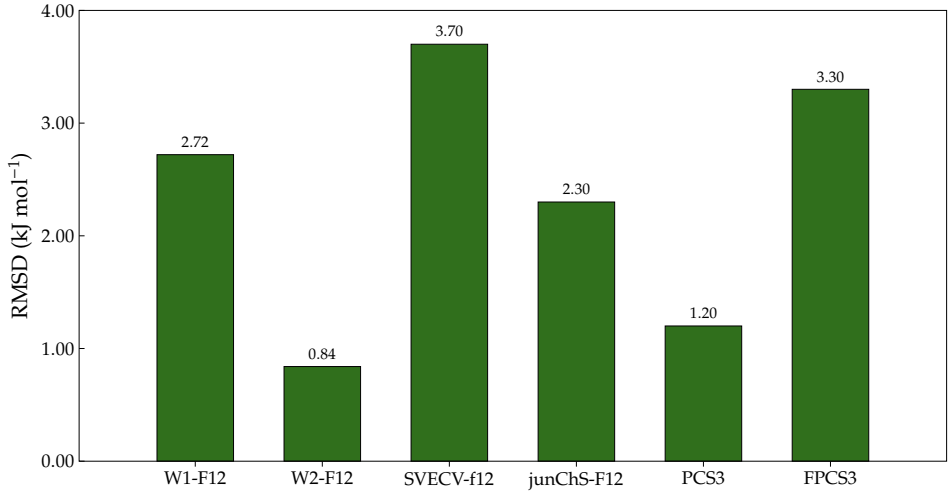


FIGURE 5.6: RMSDs, in kJ mol^{-1}, of total electronic energies for the W4-11 set, obtained with the W1-F12, W2-F12, SVECV-f12, junChS-F12, PCS3, and FPCS3 composite schemes.

As a final analysis of the W4-11 test set, the total electronic energies were examined by comparing the PCS3 model and its FNO-based extension FPCS3 with the accurate W2-F12 reference as well as with other composite schemes of comparable computational cost, i.e. W1-F12, SVECV-f12, and the refined junChS-F12 model discussed in Chapter 4. As can be appreciated from Figure 5.6, PCS3 yields smaller errors than all other methods of similar cost and achieves an accuracy close to that of the more demanding W2-F12 model. The FPCS3 scheme presents a RMSD of 3.30 kJ mol^{-1}, about 2 kJ mol^{-1} higher than that of PCS3, as expected in view of the underlying approximations. Despite this, because it is tailored for target molecules in the 15-60 atom range, FPCS3 performs remarkably well, being only approximately 0.5 kJ mol^{-1} less accurate than the more demanding W1-F12 method, and more accurate than the SVECV-f12 scheme, without the use of any empirical parameter.

A further validation of the PCS3 model was carried out using the larger W4-17 set,[84] which provides highly accurate atomization energies.[11] The complete results for the various combinations of methods and basis sets are reported in Tables A.6 and A.7 in Appendix A, while the final PCS3 results are compared in Table 5.11 with those obtained from alternative approaches of comparable computational cost; results for the FPCS3 model are also reported.

Regarding the valence contributions, the PCS3 model delivers MUE and RMSD values below 1 kJ mol^{-1} (approximately 0.8 and 1.0 kJ mol^{-1}, respectively), with a MAX deviation of about 4.0 kJ mol^{-1}. Given that the average magnitude of the higher–order excitation contribution (CCSDTQP – CCSD(T)) is around 1.2 kJ mol^{-1},[84] the accuracy achieved by PCS3 falls well within the intrinsic uncertainty of composite schemes including up to CCSD(T) terms. Thanks to the use of the cc-pVDZ-F12^{+} basis set, the residual error remains comparable for molecules containing either second- or third-row atoms (Table A.6).

The situation is more complex for the CV component, since its evaluation at the MP2 level carries an error of approximately 2.5 kJ mol^{-1}. Nevertheless, the improved approximation

described previously, which requires even lower computational cost than the valence CC correction, reduces the MUE to below 1 kJ mol^{-1}.

The improvement over the junChS-F12 scheme, discussed in Chapter 4, is evident, and even performances can be reported with respect to the alternative SVECV-f12 model.[68] The FPCS3 model provides improved results compared to SVECV-f12 for both the valence and CV contributions, and it also outperforms junChS-F12 for the CV component.

TABLE 5.11: MAX, MUE, and RMSD values obtained with PCS3, junChS-F12, SVECV-f12, and FPCS3 composite schemes for the valence and CV correlation energies of the W4-17 set, excluding post-CCSD(T) and additional corrections. Molecules yielding the MAX values are indicated in parentheses. All values are in kJ mol^{-1}.

	Valence			Core-Valence		
	MAX	MUE	RMSD	MAX	MUE	RMSD
PCS3	3.95 (C_2Cl_6)	0.77	1.00	6.26 (P_4)	0.93	1.19
junChS-F12	6.98 (P_4)	1.13	1.64	10.34 (SiF_4)	2.59	3.34
SVECV-f12	12.63 (S_4)	2.80	3.67	10.30 (C_6H_6)	3.72	4.27
FPCS3	10.69 (C_6H_6)	2.65	3.36	14.95 (C_2Cl_6)	1.85	2.71

5.2.2.2 Toward Large Molecules

Once the PCS3 and FPCS3 models were validated on benchmark datasets, their reliability and transferability were tested on large molecular systems of biological and technological relevance.

The first application concerned the reanalysis of the relative stabilities of the different tautomers and rotamers of cytosine and guanine, illustrated in Figures 4.3 and 5.2. The corresponding relative energies, computed on top of MDPCS3 geometries, are summarized in Tables 5.12 and 5.13.

TABLE 5.12: Relative electronic energies of the tautomers and rotamers of cytosine computed by PCS3 and FPCS3 methods. All the values are given in kJ mol^{-1}.

	EA	EAc	KA	KI	KIc	KA1	EI1	EI
Valence								
PCS3	0.00	2.97	3.50	6.57	13.71	34.57	51.84	70.29
FPCS3	0.00	2.94	3.48	6.61	13.73	34.61	51.84	70.29
Core-Valence								
$\Delta E_{MP2}^{CV,CBS} + \Delta E_{CCSD(T)}^{CV}$ [a]	0.00	0.05	-0.41	-0.32	-0.17	-0.06	0.24	0.15
ΔE_{MP2}^{CV} [b]	0.00	0.04	-0.26	-0.14	-0.03	0.01	0.25	0.18
Total								
PCS3[a]	0.00	3.02	3.09	6.26	13.53	34.51	52.08	70.44
PCS3[b]	0.00	3.01	3.24	6.43	13.68	34.58	52.09	70.47
FPCS3[b]	0.00	2.98	3.22	6.47	13.70	34.62	52.09	70.47

[a] $\Delta E_{MP2}^{CV,CBS} + \Delta E_{CCSD(T)}^{CV}$ with cc-pwCV(D,T)Z and cc-pwCVDZ basis sets. [b] ΔE_{MP2}^{CV} with cc-pwCVTZ basis set.

As can be appreciated by comparing the relative energies reported in Tables 5.8 and 5.9, the results delivered by PCS3 and FPCS3 are fully consistent with those from PCS3-A and PCS3-B variants, as well as from junChS-F12.

The four lowest-energy cytosine species were found to lie within approximately 7 kJ mol^{-1} from the global minimum, while a fifth tautomer presents a relative energy around 13.5 kJ mol^{-1}. The close agreement between PCS3 and FPCS3 results suggests that the FPCS3 model can be reliably used for the evaluation of relative energies, maintaining high accuracy while considerably reducing the computational cost. The CV correlation contribution can be estimated with sufficient accuracy already at the MP2 level, as tautomeric equilibria are less sensitive than atomization energies to the level of theory adopted for the evaluation of this term.

TABLE 5.13: Relative electronic energies of the tautomers and rotamers of cytosine computed by PCS3, FPCS3, and W1-F12 methods. All the values are given in kJ mol^{-1}.

	KA17	KA19	EA9	EAc9	EA7
Valence					
PCS3	0.00	2.82	3.19	4.34	15.18
FPCS3	0.00	2.85	3.12	4.38	15.20
W1-F12[a]	0.00	2.80	3.20	4.42	15.17
Core-Valence					
$\Delta E^{\mathrm{CV,CBS}}_{\mathrm{MP2}} + \Delta E^{\mathrm{CV}}_{\mathrm{CCSD(T)}}$ [b]	0.00	-0.04	0.12	0.26	0.12
$\Delta E^{\mathrm{CV}}_{\mathrm{MP2}}$ [c]	0.00	0.03	-0.01	-0.10	0.00
W1-F12[a]	0.00	0.03	0.17	0.05	0.17
Rel[a]	0.00	-0.01	0.01	0.03	0.01
DBOC[a]	0.00	-0.01	-0.02	-0.02	0.00
Total					
PCS3[b]	0.00	2.76	3.30	4.61	15.31
PCS3[c]	0.00	2.83	3.17	4.25	15.19
FPCS3[c]	0.00	2.86	3.10	4.29	15.21
W1-F12[a]	0.00	2.81	3.36	4.48	15.35

[a] Data taken from Reference 358. [b] $\Delta E^{\mathrm{CV,CBS}}_{\mathrm{MP2}} + \Delta E^{\mathrm{CV}}_{\mathrm{CCSD(T)}}$ with cc-pwCV(D,T)Z and cc-pwCVDZ basis sets. [c] $\Delta E^{\mathrm{CV}}_{\mathrm{MP2}}$ with cc-pwCVTZ basis set.

In the case of guanine (Table 5.13), the CV correlation can be accurately computed at the MP2 level, and the FPCS3 results deviate by about 0.2 kJ mol^{-1} from the corresponding PCS3 values, with both models showing excellent agreement with the W1-F12 reference data.

Summarizing, the FPCS3 scheme can be reliably employed for the calculation of tautomeric energies, providing results consistent with those obtained using the more demanding PCS3 model. This allows for a significant reduction in computational cost while preserving high accuracy. Consequently, FPCS3 can be recommended as the method of choice for studying tautomeric equilibria, since, as previously discussed, approaches based on double-hybrid functionals may yield incorrect stability orderings and are therefore not recommended for this type of analysis.

As a further application, the atomization energies and enthalpies of formation of DNA and RNA bases, and a series of PAHs were computed. These classes of molecules represent ideal test cases, as they connect the fields of biochemistry and materials science, where accurate thermochemical data are crucial for understanding stability, reactivity, and photochemical properties.

For atomization energies, the situation is more complex than for tautomeric equilibria, as the impact of the FNO approximation differs between isolated atoms and molecules, introducing a non-negligible error that increases with system size. Results from the W4-11 benchmark set (Figure 5.5) indicate that an accurate description of these quantities requires computing the CV contribution according to Equations 5.6 and 5.7, rather than relying solely on the MP2 evaluation. Furthermore, analysis of the W4-17 test set (Table 5.11) shows that FPCS3 errors are more than twice as large as those obtained with the corresponding PCS3 scheme.

The total atomization energies (TAEs) and enthalpies of formation computed both at 0 K ($\Delta H_{\mathrm{f},0K}$) and 298 K ($\Delta H^{\circ}_{\mathrm{f},298K}$) of the five nucleobases are reported in Table 5.14, together with the corresponding valence and core-valence contributions evaluated in the frame of the PSC3 and FPCS3 composite methods, and anharmonic ZPVEs. The latter ones were evaluated according to two different approaches: in the first B3 ZPVEs were scaled by 0.985,[47] while in the second a hybrid force was adopted, mixing revDSD harmonic frequencies with B3 anharmonic contributions. Scalar relativistic and diagonal Born–Oppenheimer corrections taken from the literature,[47] together with experimental fine structure SO effects,[75] were also included. For the sake of comparison, the Table also lists analogous contributions from the W2-F12 method,[47], as well as $\Delta H^{\circ}_{\mathrm{f},298K}$ values obtained in the frame of the diet-HEAT-F12 protocol[401] and also experimentally.[402] Inspection of Table 5.14 reveals that PCS3 yields energies in close agreement with their W2-F12 counterparts: the maximum differences amount to about 0.7 and 0.5 kJ mol^{-1} for valence and CV contributions, respectively, an agreement which is then mirrored in the electronic TAE (TAE$_e$). Anharmonic ZPVEs computed accordingly to the two different strategies agree with each other with an average of about 1 kJ mol^{-1}, and both are in good accord with those employed in the W2-F12 method. This agreement was expected, as the latter are obtained from harmonic B3LYP-D/def2-TZVP calculations scaled by 0.985. the agreement of each individual contribution results close values of TAE and formation enthalpies at 0 K (TAE$_0$ and ($\Delta H^{\circ}_{\mathrm{f},298K}$) computed according to PCS3 and W2-F12 approaches. Inclusion of thermal corrections, obtained from B3 computations, provides theoretical estimates of the formation enthalpies at 298 K. For these, PCS3 performs fairly well, with a MUE from W2-F12 results of 2 kJ mol^{-1}. Very notably, not only PCS3 products heats of formation that are close to the very accurate estimates of the diet-HEAT-F12 protocol, but they also agree with experimental values within the quoted uncertainties of 3.4 kJ mol^{-1}.

In contrast, the FPCS3 model gives noticeably large errors for the valence contributions, amounting to 12.43, 10.19, 13.33, 11.93, and 10.16 kJ mol^{-1} for adenine, cytosine, guanine, thymine, and uracil, respectively. The corresponding errors for the CV terms are much

smaller, with values of 0.94, 0.39, 0.60, 1.16, and 0.74 kJ mol^{-1} for the same set of molecules. This issue has been effectively addressed in local-correlation approaches through atomic-equivalent corrections.[403,404] In the present framework, the higher intrinsic accuracy of the FPCS3 model compared with local-correlation methods, together with the structural similarity of the investigated molecules, allows for a simplified one-parameter correction rather than the conventional "one-parameter-per-atom" scheme. Specifically, a scaling factor proportional to the number of non-hydrogen atoms, derived from the electronic energy difference between PCS3 and FPCS3, restores quantitative agreement with the reference data (Table 5.14).

While more elaborate strategies may be developed, this single-parameter correction can readily be implemented in two alternative ways: (i) by correcting only the FPCS3 valence plus CV energy and explicitly adding the DBOC and Rel contributions, which are non-negligible for atomization energies (unlike in tautomeric energy differences, Table 5.13); or (ii) by directly correcting the total FPCS3 valence plus CV energy, implicitly accounting for the DBOC and Rel effects in the correction parameter. The second approach is generally more effective, since DBOC and Rel corrections become increasingly expensive for larger systems. The SO contribution is always included, as it is derived from experimental fine-structure data.

Both approaches were tested, yielding a correction of –1.26 kJ mol^{-1} per non-hydrogen atom for the first method, where the average deviation of uncorrected FPCS3 electronic energies from PCS3 reference values is about 10 kJ mol^{-1} for molecules containing roughly ten non-hydrogen atoms, and –1.94 kJ mol^{-1} per non-hydrogen atom for the second approach, including in the parameter DBOC and Rel corrections.

Upon introduction of the simple parameter correction in the FPCS3 framework, its accuracy reconciles with that of both PCS3 and W2-F12 with maximum deviations of about 1.1 and 1.6 kJ mol^{-1} for the valence and TAE_0, respectively, independently of the correction strategy adopted. In this way, a quantitative agreement with both diet-HEAT-F12 predictions and experimental values is restored.

TABLE 5.14: Electronic contributions, electronic total atomization energies (TAE_e), ZPVEs, total atomization energys at 0 K (TAE_0), and enthalpies of formation at 0 K ($\Delta H_{f,0K}$) and 298 K ($\Delta H^\circ_{f,298K}$) for nucleobases at the PCS3, FPCS3, W2-F12, and diet-HEAT-F12 levels of theory. All values are in kJ mol^{-1}.

	Adenine	Cytosine	Guanine	Thymine	Uracil
Valence					
PCS3	6995.6	5864.2	7461.3	6875.4	5631.8
FPCS3[a]	6995.3	5864.3	7460.7	6876.0	5631.8
W2-F12[b]	6995.1	5863.5	7461.5[c]	6874.9	5631.3
Core-Valence					
PCS3	35.8	27.5	36.9	30.9	25.9
FPCS3	34.9	27.1	36.3	29.7	25.1
W2-F12[b]	36.1	27.8	37.4[c]	31.2	26.3
Rel[b]	-8.5	-6.9	-9.6	-7.7	-6.8
SO[d]	-1.8	-2.3	-2.7	-3.6	-3.3
DBOC[b]	1.7	1.5	1.9	1.6	1.3
TAE_e					
PCS3	7022.8	5884.0	7487.8	6896.6	5648.9
FPCS3[e]	7021.6	5883.7	7486.6	6896.0	5648.1
FPCS3[f]	7021.7	5883.7	7486.9	6896.0	5648.2
W2-F12[b]	7022.6	5883.6	7488.5[c]	6896.4	5648.8
ZPVE					
Anharm[g]	289.0	255.0	304.1	298.9	226.7
Scal. B3/6-31G*[h]	288.7	254.7	302.1	297.6	225.4
W2-F12[b,i]	288.1	253.3	301.4	296.6	224.8
TAE_0					
PCS3[j]	6733.8	5629.0	7183.7	6597.7	5422.2
FPCS3[k,m]	6732.9	5629.0	7184.5	6598.4	5422.7
FPCS3[l,m]	6733.0	5629.0	7184.8	6598.4	5422.8
W2-F12[b]	6734.5	5630.3	7187.1	6599.8	5424.0
$\Delta H^\circ_{f,0K}$					
PCS3	256.0	-44.7	53.2	-309.7	-277.7
FPCS3[k]	257.0	-44.8	52.3	-310.5	-278.2
FPCS3[l]	256.9	-44.8	52.0	-310.5	-278.3
W2-F12[b]	255.4	-45.9	49.7	-311.9	-279.5
$\Delta H^\circ_{f,298K}$					
PCS3[n]	228.2	-67.4	24.7	-334.7	-297.4
FPCS3[k,o]	229.2	-67.3	23.9	-335.4	-298.0
FPCS3[l,p]	229.1	-67.3	23.6	-335.4	-298.1
W2-F12[b]	229.8	-68.0	21.3[c]	-337.1	-299.4
diet-HEAT-F12[q]	232.0 ± 3.9	-68.8 ± 3.6		-339.3 ± 3.9	-302.1 ± 3.5
Exp.[r]	225.5 ± 3.4	-69.5 ± 3.4	-	-338.1 ± 3.4	-301.7 ± 3.4

[a] Correction of –1.26 kJ mol^{-1} per non-hydrogen atom. [b] Data taken from Reference 47. [c] Evaluated at W1-F12 level. [d] Data taken from the experimental fine structure, Reference 75. [e] FPCS3 one-parameter corrected valence energies including CV, Rel, SO, and DBOC contributions. [f] FPCS3 energies corrected by –1.94 kJ mol^{-1} per non-hydrogen atom, starting from valence, CV, and experimental SO terms. [g] revDSD/cc-pVTZ-F12 harmonic and B3/6-31G* anharmonic ZPVEs. [h] Harmonic B3/6-31G* ZPVEs scaled by 0.985.[47] [i] Harmonic B3LYP-D/Def2-TZVP ZPVEs scaled by 0.985.[47] [j] Including anharmonic ZPVEs. [k] FPCS3 TAE_e corrected with the –1.26 kJ mol^{-1} parameter. [l] FPCS3 TAE_e corrected with the –1.94 kJ mol^{-1} parameter. [m] Including scaled B3/6-31G* ZPVEs. [n] Thermal contributions at the harmonic B3/6-31G* level. [o] Thermal contributions at the harmonic B3/6-31G* level. [p] Thermal contributions at the harmonic B3LYP-D/def2-TZVP level. [q] Data taken from Reference 401. [r] Data taken from Reference 402.

The effectiveness and transportability of the one-parameter correction is demonstrated by the heats of formation for a series of PAHs of increasing size reported in Table 5.15. More precisely, five PAHs (benzene, toluene, naphthalene, anthracene, and phenanthrene) with up to 24 atoms, and a molecular motor composed of 47 atoms, whose structures are shown in Figure 5.7, were considered. The heats of formation computed using the PCS3 model are in close agreement with those from the W1-F12 scheme, and both of them are in accord with experimental determinations and with the ATcT values. In this respect, the MUE of PCS3 predictions from ATcT data amount to about 2 kJ mol^{-1} with a maximum deviation of about 2.5 kJ mol^{-1} observed for anthracene.

For the FPCS3 model, these values were determined from TAE$_e$ energies corrected by –1.94 kJ mol^{-1} per non-hydrogen atom, derived starting from valence and CV contributions combined with experimental SO terms. Although this correction was originally obtained from the nucleobases reported in Table 5.14, it constitutes a single-parameter adjustment that scales with the number of non-hydrogen atoms, treating carbon and heteroatoms on an average basis. Considering that both nucleobases and PAHs are conjugated aromatic systems, this approximation was also tested on the latter, while acknowledging that further validation and refinement of the parameter are still needed.

Despite this simplification, the FPCS3 model provides results consistent with experimental data as well as with both PCS3 and W1-F12 calculations, while maintaining a significantly lower computational cost. Moreover, its efficiency allows the accurate computation of the formation enthalpy of a molecular motor comprising nearly 50 atoms (Figure 5.7), without relying on any local correlation approximations.[130]

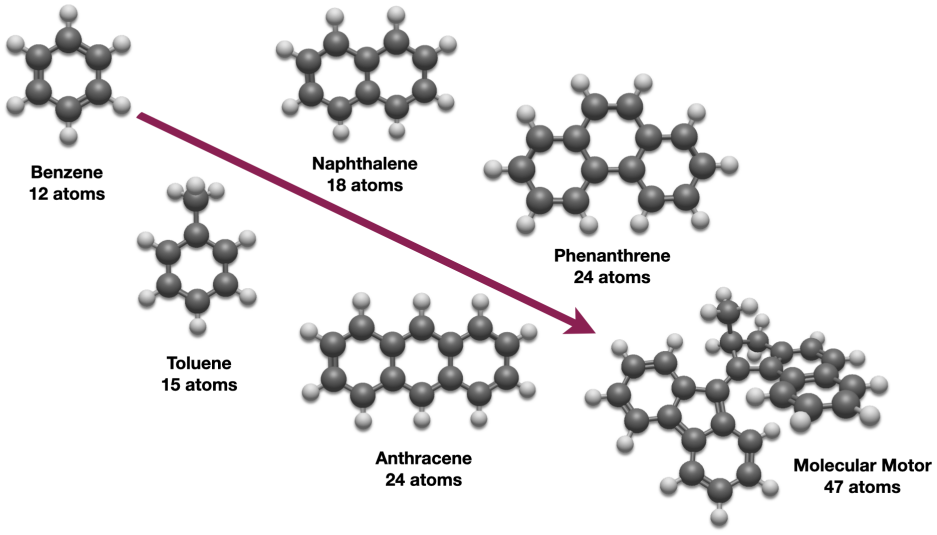


FIGURE 5.7: From benzene to extended conjugated hydrocarbons of increasing size.

TABLE 5.15: Enthalpies of formation at 298 K for conjugated hydrocarbons of increasing size, evaluated using PCS3, FPCS3, and W1-F12 methods. All values are in kJ mol^{-1}.

	Benzene	Toluene	Naphthalene	Anthracene	Phenanthrene	Molecular Motor
FPCS3	82.7	48.5	147.1	228.0	203.2	406.3
PCS3	82.2	48.0	145.7	224.3	200.7	-
W1-F12[a]	81.4	49.1	144.8	223.2	199.3	-
Exp.[b]	82.9 ± 0.9	50.1 ± 1.1	150.6 ± 1.5	229.4 ± 2.9	202.2 ± 2.3	-
ATcT[c]	83.2 ± 0.2	50.1 ± 0.3	147.6 ± 0.6	226.8 ± 1.0	203.0 ± 0.8	-

[a] Data taken from Reference 405. [b] Data taken from Reference 406. [c] Data taken from Reference 407.

5.2.3 Molecular Structures and Rotational Constants

5.2.3.1 Validation of the PCS2 and MDPCS3 Equilibrium Geometries

A first validation of the PCS2 scheme for equilibrium structures was carried out on 32 small molecules containing elements from the first three rows of the periodic table, for which accurate equilibrium semi-experimental structures are available. This initial step aimed to assess the level of theory adopted for the description of valence correlation and to analyze the impact of basis set extension up to cc-pVQZ-F12. Unless otherwise specified, reference data were taken from the SE100 database.[148] The test set included: CH_4,[142] CO_2, HCN, HNC, H_2O, NH_3, C_2H_2, C_2H_4, H_2CO, CH_2O_2 (trans formic acid), CH_2NH, BH_3NH_3, C_2H_4O (oxirane), C_2H_4NH (aziridine), CH_2F_2, HOF, CH_2CHF, cyc$-C_3H_6$, NH_2OH, BH_2OH, HNO,[359] H_2O_2,[360] SO_2, H_2S,[361] PH_3, H_2CS, CH_2PH, C_2H_4S (thiirane), H_2S_2, HClCO, CH_2Cl_2, and CH_2ClF; see Figure 5.11 for molecular structures.

For this molecular set, geometry optimizations were carried out at the CCSD(T)-F12b/cc-pVnZ-F12 level, with $n = 2, 3$, and 4. The corresponding MAX, MUE, and RMSD values were evaluated for all geometrical parameters. Starting from the CCSD(T)-F12b/cc-pVDZ-F12$^+$ geometries, the effect of basis set extension on geometrical parameters computed within the frozen core approximation was also investigated at the MP2-F12 level. The $\Delta r_{\text{MP2-F12}}$ corrections were computed as the differences between geometrical parameters obtained using the (3,2), (4,2), and (4,3) pairs of cc-pVnZ-F12 basis sets. For each level of theory, the MP2/cc-pwCVTZ core-valence correction was added. In addition, for molecules containing third-row atoms, CCSD(T)-F12b/cc-pVDZ-F12 geometries including CV correlation but excluding the additional f function were also considered. The corresponding error statistics are summarized in Table 5.16.

TABLE 5.16: MAX, MUE, and RMSD of geometrical parameters computed at the CCSD(T)-F12b/cc-pVnZ-F12 level, with n = 2, 3, and 4. The $\Delta r_{\text{MP2-F12}}$ corrections were evaluated as the differences between geometrical parameters obtained using the (3,2), (4,2), and (4,3) pairs of cc-pVnZ-F12 basis sets. For each parameter, the first row reports valence-only results, while the second row includes the MP2/cc-pwCVTZ CV correction. For molecules containing third-row atoms, statistics for CCSD(T)-F12b/cc-pVDZ-F12 geometries including CV corrections (excluding the additional f function) are shown in square brackets. All results are given relative to SE reference data. Bond lengths (d) are in Å, and valence (θ) and dihedral (φ) angles in degrees.

	$r_{\text{CCSD(T)-F12b}}$			$r^{\text{cc-pVDZ-F12}}_{\text{CCSD(T)-F12b}} + \Delta r_{\text{MP2-F12}}$		
	cc-pVDZ-F12[a]	cc-pVTZ-F12	cc-pVQZ-F12	(3,2[a])	(4,2[a])	(4,3)
2nd-row (22 molecules)						
MAX(d)	0.0044	0.0043	0.0039	0.0041	0.0039	0.0042
	0.0032	0.0038	0.0049	0.0055	0.0061	0.0036
MAX(θ,φ)	0.61	0.94	0.94	1.10	1.14	0.60
	0.62	0.83	0.83	0.99	1.03	0.61
MUE(d)	0.0012	0.0011	0.0010	0.0010	0.0010	0.0011
	0.0004	0.0004	0.0005	0.0005	0.0005	0.0004
MUE(θ,φ)	0.06	0.05	0.05	0.06	0.06	0.06
	0.04	0.05	0.06	0.06	0.06	0.05
RMSD(d)	0.0022	0.0021	0.0019	0.0021	0.0020	0.0021
	0.0011	0.0012	0.0014	0.0016	0.0017	0.0012
RMSD(θ,φ)	0.20	0.22	0.22	0.24	0.25	0.20
	0.18	0.21	0.21	0.23	0.24	0.18
3rd-row (10 molecules)						
MAX(d)	0.0050	0.0045	0.0039	0.0053	0.0053	0.0050
	0.0023	0.0032	0.0027	0.0019	0.0017	0.0021
	[0.0052]					
MAX(θ,φ)	0.18	0.17	0.25	0.19	0.21	0.18
	0.22	0.18	0.17	0.15	0.14	0.22
	[0.31]					
MUE(d)	0.0023	0.0024	0.0021	0.0026	0.0025	0.0022
	0.0006	0.0007	0.0008	0.0005	0.0005	0.0006
	[0.0014]					
MUE(θ,φ)	0.06	0.06	0.06	0.06	0.05	0.06
	0.07	0.07	0.06	0.07	0.06	0.07
	[0.14]					
RMSD(d)	0.0025	0.0026	0.0022	0.0029	0.0029	0.0026
	0.0008	0.0010	0.0011	0.0006	0.0007	0.0008
	[0.0019]					
RMSD(θ,φ)	0.08	0.08	0.08	0.08	0.08	0.08
	0.09	0.08	0.08	0.08	0.07	0.09
	[0.16]					

[a] Employing cc-pVDZ-F12 and cc-pVDZ-F12$^+$ basis sets for 2nd- and 3rd-row atoms, respectively.

All the investigated levels of theory yield errors well below the target accuracy of 1 mÅ for bond lengths and 0.1° for angles. In particular, both the MP2-F12 corrections and the CCSD(T)-F12b calculations beyond the cc-pVDZ-F12$^+$ basis set have a negligible impact within this accuracy range. Moreover, for molecules containing third-row atoms, the use of the cc-pVDZ-F12$^+$ basis set ensures errors comparable to those obtained for first- and second-row species. These results demonstrate that a double-ζ valence basis set combined

with an explicitly correlated method and a CV correction computed at the MP2/cc-pwCVTZ level provides an excellent compromise between accuracy and computational cost, without the need for larger valence basis sets or additional corrections.

A further validation of the PCS2 method was carried out by optimizing the structures of 106 molecules containing the main elements of life (H, B, C, N, O, F, P, S, and Cl), taken from the compilation of SE equilibrium geometries reported in Reference 149. In addition, five more molecules with accurate SE structures were included, namely cyclohexane,[408] piperidine,[408] H_2S,[361] o-benzyne,[409] and CH_4.[142] The resulting benchmark set comprised 111 molecules, displayed in Figure 5.11, and the corresponding statistical analysis is summarized in Table 5.17.

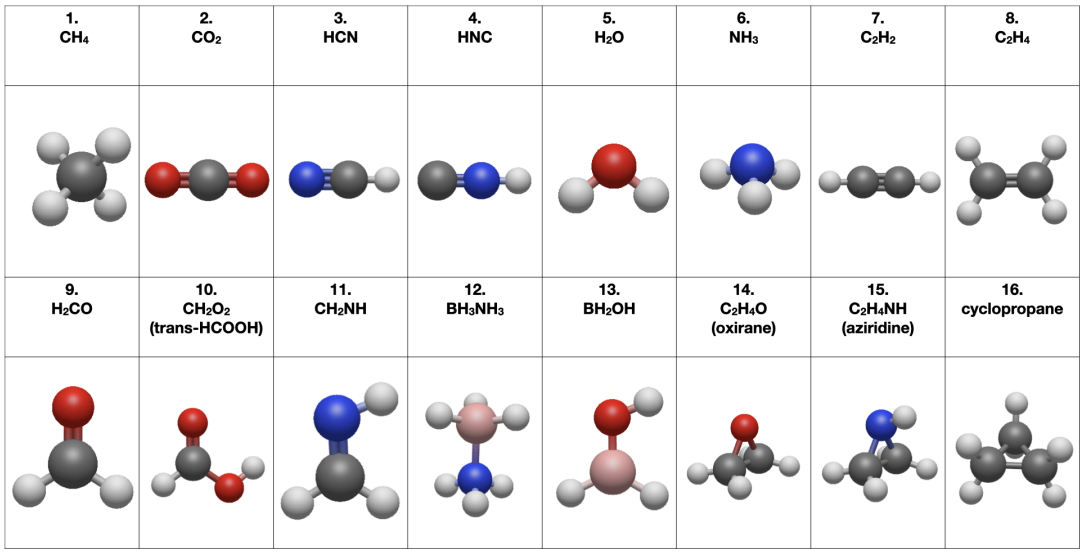
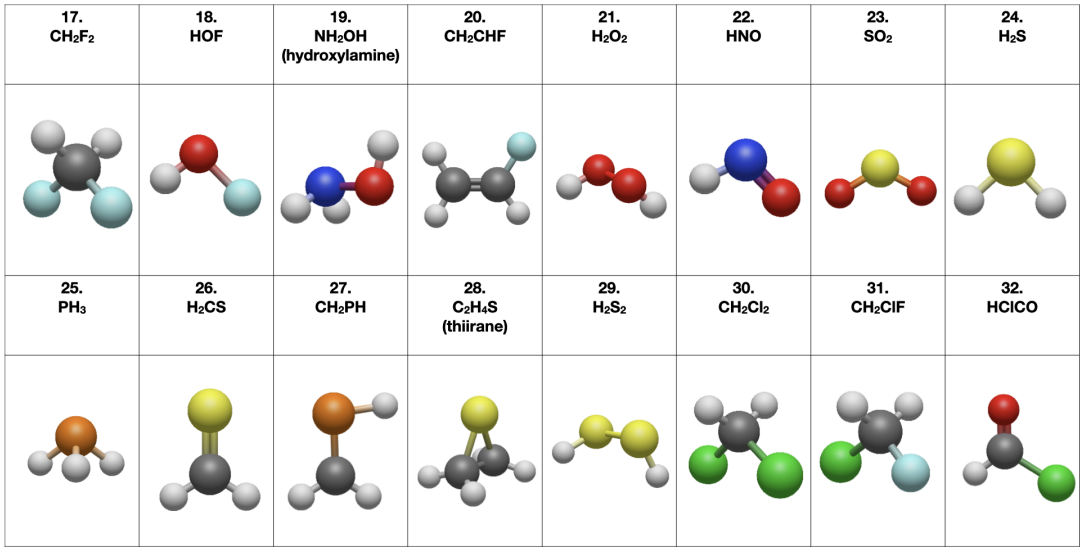


FIGURE 5.8: Collection of the 111 molecular structures included in the PCS2 validation set.



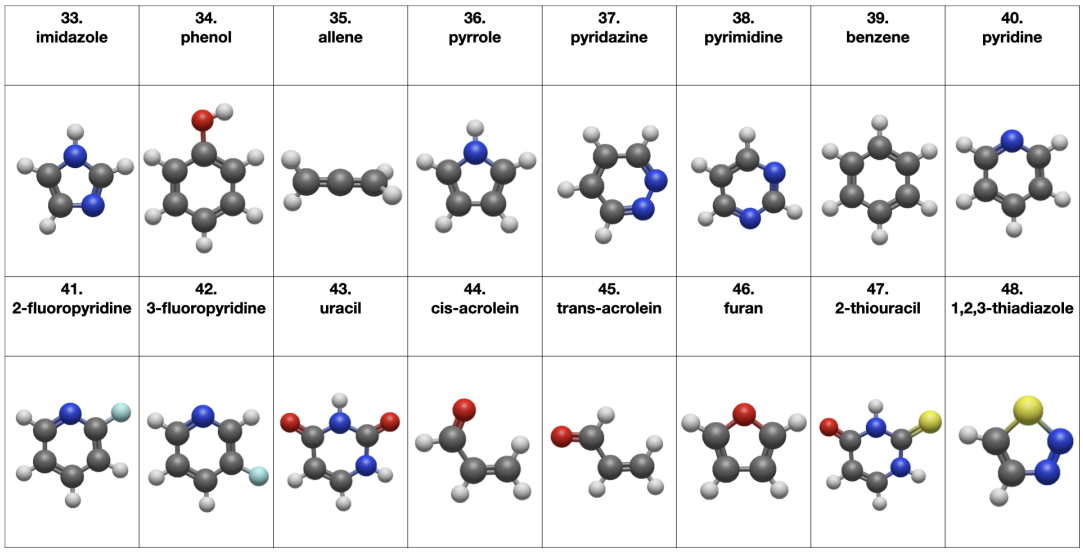
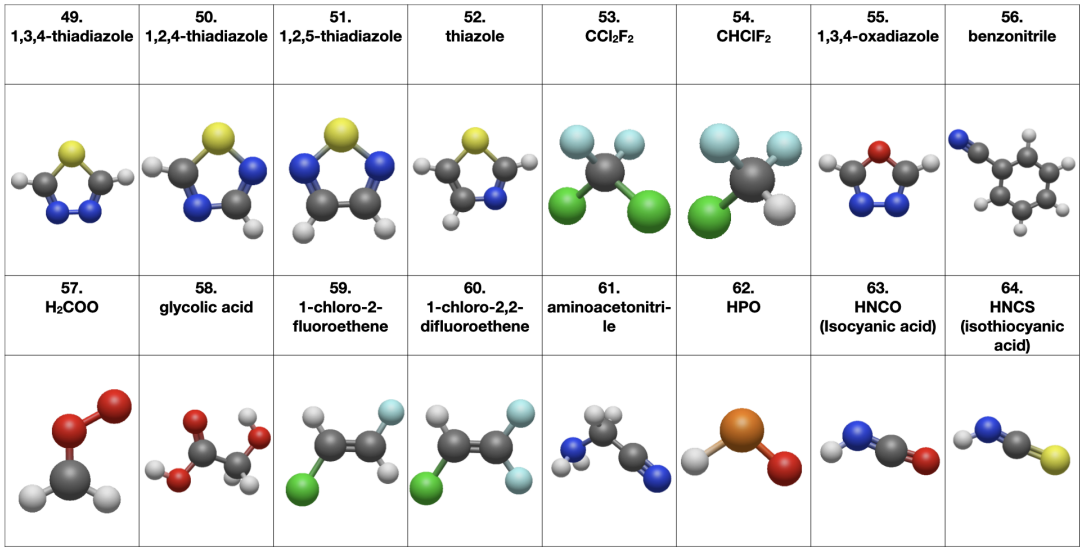


FIGURE 5.9: Collection of the 111 molecular structures included in the PCS2 validation set (continued).



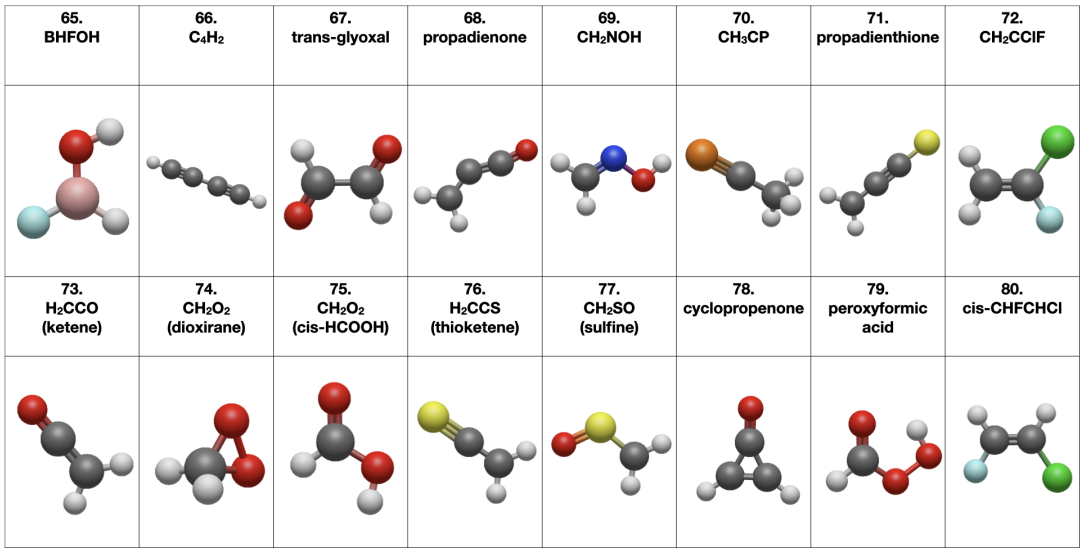


FIGURE 5.10: Collection of the 111 molecular structures included in the PCS2 validation set (continued).

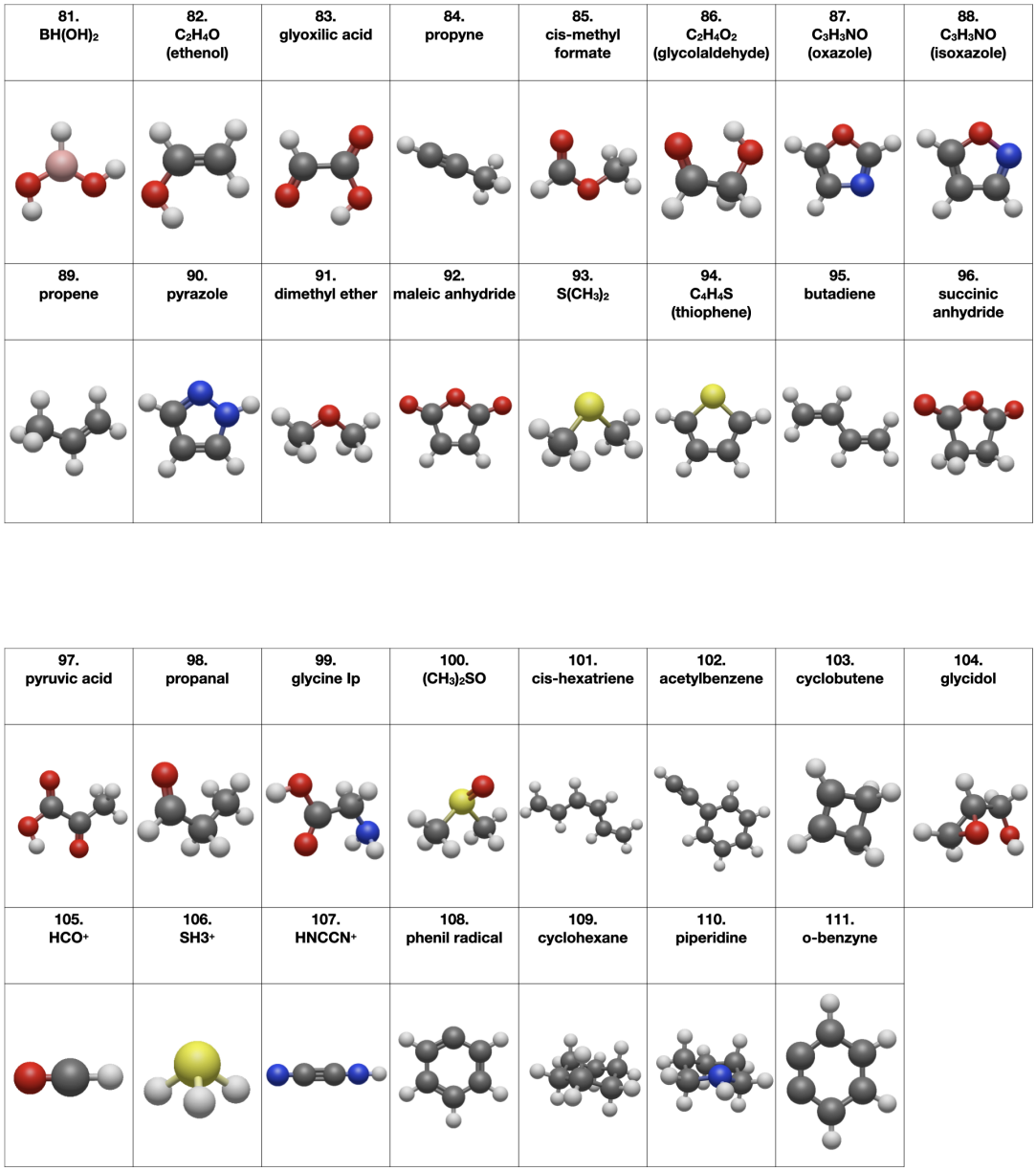


FIGURE 5.11: Collection of the 111 molecular structures included in the PCS2 validation set (continued).

TABLE 5.17: MAX, MUE, and RMSD of geometrical parameters computed using the PCS2 model for a dataset of 111 molecules. For molecules containing third-row atoms, the statistics for PCS2 results obtained without the additional f function are reported in square brackets. All results are given relative to the SE reference data. Bond lengths (d) are expressed in Å, and valence (θ) and dihedral (φ) angles in degrees.

	2nd-row (79 molecules)	3rd-row (32 molecules)	Total (111 molecules)
MAX(d)	0.0068	0.0038 [0.0052]	0.0068 [0.0068]
MAX(θ,φ)	1.30	0.69 [0.69]	1.30 [1.30]
MUE(d)	0.0008	0.0007 [0.0011]	0.0008 [0.0009]
MUE(θ,φ)	0.09	0.01 [0.01]	0.07 [0.07]
RMSD(d)	0.0011	0.0010 [0.0015]	0.0011 [0.0012]
RMSD(θ,φ)	0.16	0.15 [0.15]	1.16 [1.16]

The comparison of Tables 5.16 and 5.17 indicates that the errors remain within the same range as those obtained for the smaller 32-molecule set. Once again, replacing the cc-pVDZ-F12 basis set with its cc-pVDZ-F12$^+$ counterpart markedly improves the accuracy for molecules containing third-row atoms, yielding errors comparable to those observed for systems composed exclusively of first- and second-row elements.

After the extensive validation of the PCS2 geometric protocol, the reduced-cost MDPCS3 model was assessed by replacing the CCSD(T)-F12b/cc-pVDZ-F12$^+$ valence evaluation with the more affordable revDSD/cc-pVTZ-F12$^-$ alternative. The validation was carried out by benchmarking against the accurate SE equilibrium geometries of 54 molecules taken from the SE127 database.[149] In addition to the 32 molecules previously examined, this set includes 22 larger systems: imidazole, phenol, allene, pyrrole, pyridazine, pyrimidine, benzene, pyridine, 2-fluoropyridine, 3-fluoropyridine, uracil, cis-acrolein, trans-acrolein, furan, 2-thiouracil, 1,2,3-thiadiazole, 1,3,4-thiadiazole, 1,2,4-thiadiazole, 1,2,5-thiadiazole, thiazole, CCl_2F_2, and $CHClF_2$. The error statistics for the revDSD/cc-pVTZ-F12$^-$, MDPCS3, and PCS models are summarized in Table 5.18.

TABLE 5.18: MAX, MUE, and RMSD of geometrical parameters computed at the revDSD/cc-pVTZ-F12$^-$, MDPCS3, and PCS2 levels of theory. All results are given relative to the SE reference data. Bond lengths (d) are expressed in Å, and valence (θ) and dihedral (φ) angles in degrees.

	$r_{\text{revDSD}}^{\text{cc-pVTZ-F12}^-}$	MDPCS3	PCS2
2nd-row (36 molecules)			
MAX(d)	0.0073	0.0058	0.0046
MAX(θ,φ)	0.70	0.71	0.62
MUE(d)	0.0024	0.0017	0.0008
MUE(θ,φ)	0.11	0.13	0.07
RMSD(d)	0.0027	0.0020	0.0011
RMSD(θ,φ)	0.16	0.18	0.12
3rd-row (18 molecules)			
MAX(d)	0.0086	0.0057	0.0027
MAX(θ,φ)	0.77	0.76	0.79
MUE(d)	0.0030	0.0017	0.0007
MUE(θ,φ)	0.13	0.12	0.09
RMSD(d)	0.0039	0.0024	0.0011
RMSD(θ,φ)	0.19	0.19	0.17

The results confirm that the general accuracy of revDSD geometries, with bond length deviations of only a few mÅ. The inclusion of CV correlation has a negligible impact on valence and dihedral angles but significantly improves bond length accuracy, reducing the associated errors by roughly a factor of two. Replacing the valence description from the double-hybrid functional with the CCSD(T)-F12b/cc-pVDZ-F12$^+$ level yields a further twofold improvement. Overall, the MDPCS3 model represents an excellent compromise between computational cost and accuracy, particularly suitable for the structural characterization of increasingly large molecular systems.

5.2.3.2 Application of the PCS2 and MDPCS3 Models to the Evaluation of Rotational Constants

After validating their performance on small-to-medium molecules, the PCS2 and MDPCS3 composite geometric schemes were applied to a representative set of medium-sized gas-phase molecules of varying structural complexity, encompassing systems of both biological and technological relevance, for which experimental rotational constants are available. For such species, an unbiased validation of computed geometries is often hindered by the absence of accurate SE equilibrium structures. In these cases, structural accuracy can be indirectly evaluated through a bottom-up approach, in which theoretical equilibrium rotational constants are compared with experimental ground state measurements corrected for vibrational effects. The latter account for the transition from the experimental to the equilibrium rotational constants and are computed at the B3/6-31G* level of theory.[188] Statistical analyses were performed in terms of MAX and MUE deviations, as well as their relative counterparts, with respect to the SE experimental constants.

The first set of applications concerned the five low-energy tautomers and rotamers of cytosine[169] and guanine[400], respectively shown in Figures 4.3 and 5.2, for which experimental rotational constants are available. An exception is the highest-energy guanine tautomer, whose rotational constants were only estimated theoretically, since this species has not been detected experimentally. Cytosine and guanine, being two of the fundamental building blocks of DNA and RNA, provide a biologically meaningful application for assessing the ability of the PCSs to capture subtle energetic and structural differences among tautomers and rotamers.

For cytosine, a detailed comparison of the SE equilibrium rotational constants and those computed at the revDSD/cc-pVTZ-F12^{-}, MDPCS3, BDPCS3, CCSD(T)-F12b/cc-pVDZ-F12, and PCS2 levels of theory is reported in Table 5.19.

TABLE 5.19: Comparison of SE equilibrium rotational constants and those computed at the revDSD/cc-pVTZ-F12$^-$, MDPCS3, BDPCS3, CCSD(T)-F12b/cc-pVDZ-F12, and PCS2 levels of theory (B_a^e, B_b^e, and B_c^e) for cytosine tautomers and rotamers, including B3/6-31G* vibrational corrections. The MAX, MUE, and their relative percentage counterparts (MAX% and MUE%) with respect to the SE values are also provided. All quantities, except percentage errors, are in MHz.

		SE[a]	revDSD[b] / cc-pVTZ-F12$^-$	MDPCS3	BDPCS3[b]	CCSD(T)-F12b / cc-pVDZ-F12	PCS2	$\Delta B^{\text{vib}}_{\text{B3/6-31G*}}$[c]
	B_a^e	3979.5	3967.9	3986.4	3979.6	3962.5	3978.1	27.6
	B_b^e	2020.0	2013.9	2021.6	2019.8	2012.0	2018.7	11.0
	B_c^e	1341.1	1336.5	1341.9	1340.5	1335.3	1340.0	8.6
EA	MAX		11.6	6.9	0.6	17.0	1.4	
	MUE		7.4	3.1	0.3	10.3	1.3	
	MAX%		0.34	0.17	0.04	0.43	0.09	
	MUE%		0.31	0.11	0.02	0.42	0.06	
	B_a^e	3913.4	3902.8	3919.3	3914.7	3896.9	3911.4	23.9
	B_b^e	2038.9	2032.2	2040.4	2038.1	2030.4	2037.3	12.6
	B_c^e	1341.7	1337.1	1342.5	1341.0	1335.8	1340.4	8.8
EAc	MAX		10.6	5.9	1.3	16.5	2.0	
	MUE		7.3	2.7	0.9	10.3	1.6	
	MAX%		0.34	0.15	0.05	0.44	0.09	
	MUE%		0.31	0.09	0.04	0.42	0.08	
	B_a^e	3899.0	3888.8	3906.6	3901.5	3884.2	3899.6	27.4
	B_b^e	2034.4	2028.8	2036.6	2034.9	2027.5	2034.2	9.4
	B_c^e	1338.2	1333.7	1339.1	1337.8	1332.8	1337.4	7.9
KA	MAX		10.2	7.6	2.5	14.8	0.8	
	MUE		6.8	3.6	1.1	9.1	0.5	
	MAX%		0.34	0.19	0.06	0.41	0.06	
	MUE%		0.29	0.12	0.04	0.38	0.03	
	B_a^e	3877.6	3865.7	3882.8	3878.1	3860.9	3876.0	29.4
	B_b^e	2037.3	2030.2	2037.8	2036.4	2028.7	2035.6	11.0
	B_c^e	1335.7	1331.1	1336.4	1335.2	1329.9	1334.6	7.7
KI	MAX		11.9	5.2	0.9	16.7	1.7	
	MUE		7.9	2.2	0.6	10.4	1.5	
	MAX%		0.35	0.14	0.04	0.43	0.09	
	MUE%		0.33	0.07	0.03	0.43	0.07	
	B_a^e	3889.7	3879.6	3894.7	3891.7	3873.1	3886.9	28.4
	B_b^e	2022.8	2015.2	2023.8	2021.4	2014.5	2022.0	11.4
	B_c^e	1330.9	1326.3	1331.8	1330.4	1325.2	1330.1	7.7
KIc	MAX		10.1	5.0	2.0	16.6	2.8	
	MUE		7.4	2.3	1.3	10.2	1.5	
	MAX%		0.38	0.13	0.07	0.43	0.07	
	MUE%		0.33	0.08	0.05	0.42	0.06	

[a] SE equilibrium rotational constants obtained correcting the experimental ground state rotational constants by B3/6-31G* vibrational corrections. Experimental data (truncated to 1 decimal figure) taken from Reference 169. [b] Data taken from Reference 167. [c] Data taken from Reference 399.

Overall, the results demonstrate that PCSs are able to reproduce the rotational constants of biologically relevant polyfunctional systems with spectroscopic accuracy. The results obtained with the more affordable MDPCS3 and BDPCS3 models closely reproduce those delivered by the much more demanding PCS2 scheme. All the PCS-based approaches achieve the target accuracy of 0.1%, a level of precision that cannot be reached without explicitly accounting for CV correlation effects, as shown by the revDSD/cc-pVTZ-F12$^-$ and CCSD(T)-F12b/cc-pVDZ-F12 results. The overall error statistics further confirm the reliability of both PCS2 and BDPCS3 formulations, the latter being fully consistent with and suitable for use within the FPCS3 framework.

The equilibrium rotational constants of the five most stable guanine forms, computed at various levels of theory, are reported in Table 5.20. A comparison with SE equilibrium values includes results obtained at the revDSD/cc-pVTZ-F12$^-$, MDPCS3, *fc*-MP2/cc-pwCVTZ, and *ae*-MP2/cc-pwCVTZ levels.

TABLE 5.20: Comparison of SE equilibrium rotational constants and those computed at the revDSD/cc-pVTZ-F12$^-$, MDPCS3, *fc*-MP2/cc-pwCVTZ, and *ae*-MP2/cc-pwCVTZ levels of theory (B_a^e, B_b^e, and B_c^e) for the five most stable guanine tautomers and rotamers, including B3/6-31G* vibrational corrections. The MAX, MUE, and their relative percentage counterparts (MAX% and MUE%) with respect to the SE values are also provided. All quantities, except percentage errors, are in MHz.

		SE[a]	revDSD	MDPCS3	*fc*-MP2	*ae*-MP2	$\Delta B^{\text{vib}}_{\text{B3/6-31G*}}$
			cc-pVTZ-F12$^-$		cc-pwCVTZ	cc-pwCVTZ	
	B_a^e	1933.8	1927.0	1934.1	1921.3	1928.5	11.6
	B_b^e	1128.4	1124.5	1127.9	1128.6	1132.0	6.7
	B_c^e	713.2	710.7	713.0	711.7	714.1	4.2
KA17	MUE		4.4	0.3	4.7	3.3	
	MAX		6.8	0.4	12.5	5.3	
	MUE%		0.35	0.03	0.29	0.24	
	MAX%		0.35	0.04	0.64	0.32	
	B_a^e	1934.3	1923.6	1934.0	1922.3	1929.1	12.0
	B_b^e	1123.2	1119.1	1122.7	1122.9	1126.4	6.5
	B_c^e	711.1	708.5	710.8	709.5	711.8	4.2
KA19	MUE		4.5	0.3	4.6	3.0	
	MAX		7.0	0.5	12.0	5.2	
	MUE%		0.36	0.03	0.29	0.22	
	MAX%		0.36	0.05	0.62	0.29	
	B_a^e	1929.4	1923.2	1929.7	1921.6	1928.1	13.3
	B_b^e	1138.4	1134.7	1138.3	1137.3	1141.0	6.0
	B_c^e	716.4	713.9	716.2	714.9	717.2	4.2
EA9	MUE		4.1	0.2	3.4	1.6	
	MAX		6.2	0.3	7.8	2.6	
	MUE%		0.33	0.01	0.24	0.14	
	MAX%		0.35	0.02	0.40	0.23	
	B_a^e	1937.3	1931.2	1937.7	1930.3	1937.0	13.8
	B_b^e	1142.1	1138.5	1142.1	1141.3	1145.0	6.1
	B_c^e	719.0	716.5	718.8	717.6	719.9	4.3
EAc9	MUE		4.1	0.2	3.1	1.3	
	MAX		6.1	0.4	7.0	2.9	
	MUE%		0.33	0.02	0.21	0.13	
	MAX%		0.35	0.03	0.36	0.25	
	B_a^e	-	1912.9	1919.5	1909.9	1916.5	12.7
EA7	B_b^e	-	1137.2	1140.7	1139.8	1143.2	6.2
	B_c^e	-	715.8	714.2	716.5	4.2	

[a] SE equilibrium rotational constants obtained correcting the experimental ground state rotational constants by B3/6-31G* vibrational corrections. Experimental data (truncated to 1 decimal figure) taken from Reference 400.

Focusing on the individual components of the MDPCS3 scheme, the relative mean unsigned error (MUE%) remains consistently around 0.02%, and the corresponding maximum unsigned error (MAX%) never exceeds 0.05%. These results are comparable to those obtained

with the most sophisticated and computationally demanding wave function-based composite methods developed for small semi-rigid molecules.

Moving from the "molecular bricks of life" to naturally occurring compounds, four bicyclo[2.2.1]heptanes, namely norbornane, norbornene, norbornadiene, and norcamphor, were selected for investigation at the PCS2 and BDPCS3 levels of theory.[189] The structures of the four nor-terpenes are shown in Figure 5.12. These bicyclic hydrocarbons can be regarded as simplified models of naturally occurring monoterpenes and camphor derivatives. Their rigid frameworks and well-characterized spectra make them ideal prototypes for evaluating the performance of CSs on compact, saturated systems of biological and industrial relevance. The prefix "nor" refers to the common bicyclic backbone of bornane and camphor, where the methyl groups of the parent compounds are replaced by hydrogen atoms. These systems are rigid bicyclic hydrocarbons characterized by a single low-energy minimum, for which the corresponding rotational spectra have been experimentally measured.[410–413]

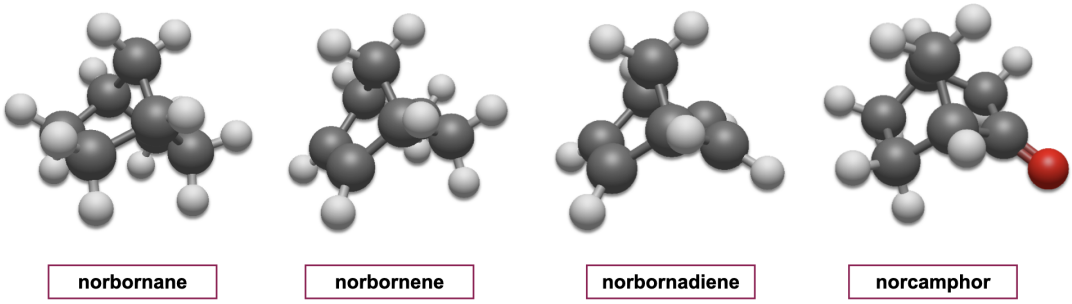


FIGURE 5.12: Structure of the norbornane, norbornene, norbornadiene, and norcamphor.

In Table 5.21, the equilibrium rotational constants computed at the revDSD/cc-pVTZ-F12$^-$, CCSD(T)-F12b/cc-pVDZ-F12, BDPCS3, and PCS2 levels of theory are compared with their SE counterparts. The latter were derived by adding B3/6-31G* vibrational and B3/cc-pVTZ-F12$^-$ electronic corrections to the experimental ground state rotational constants. As can be seen, for closed-shell molecules that do not contain heavy atoms, the electronic correction ΔB_τ^{el} is smaller than the target accuracy of the proposed computational protocol.[174] In contrast, vibrational corrections play a non-negligible role and must be included to achieve quantitatively reliable results.

TABLE 5.21: Comparison of SE equilibrium rotational constants and those computed at the revDSD/cc-pVTZ-F12$^-$, CCSD(T)-F12b/cc-pVDZ-F12, BDPCS3, and PCS2 levels of theory (B_a^e, B_b^e, and B_c^e) for the nor-compounds, including B3/6-31G* vibrational and B3/cc-pVTZ-F12$^-$ electronic corrections. The MAX, MUE, and their relative percentage counterparts (MAX% and MUE%) with respect to the SE values are also provided. All quantities, except percentage errors, are in MHz.

		SE[a]	revDSD	CCSD(T)-F12b	BDPCS3[b]	PCS2	$\Delta B^{\text{vib}}_{\text{B3/G-31G*}}$	$\Delta B^{\text{el}}_{\text{B3/cc-pVTZ-F12}^-}$
			cc-pVTZ-F12$^-$	cc-pVDZ				
Norbornane	B_a^e	3729.4	3719.6	3713.2	3733.8	3728.0	35.1	0.003
	B_b^e	3247.6	3235.1	3233.5	3246.7	3246.3	35.0	-0.007
	B_c^e	2804.8	2793.9	2792.5	2804.0	2803.5	30.0	-0.009
	MUE		11.0	14.1	2.1	1.3		
	MAX		12.5	16.1	4.4	1.4		
	MUE%		0.34	0.43	0.06	0.04		
	MAX%		0.39	0.44	0.12	0.05		
Norbornene	B_a^e	3963.0	3952.9	3946.4	3966.7	3961.9	39.2	-0.017
	B_b^e	3465.9	3452.7	3449.6	3465.4	3463.8	33.5	0.002
	B_c^e	3041.2	3031.3	3026.8	3041.7	3039.5	26.3	-0.042
	MUE		11.1	15.8	1.6	1.6		
	MAX		13.3	16.5	3.7	2.1		
	MUE%		0.32	0.45	0.04	0.05		
	MAX%		0.38	0.48	0.09	0.06		
Norbornadiene	B_a^e	4314.2	4305.6	4297.6	4319.1	4314.1	40.5	-0.077
	B_b^e	3642.3	3628.7	3623.8	3642.3	3638.6	32.0	-0.112
	B_c^e	3210.4	3199.6	3193.7	3209.9	3207.3	24.0	-0.005
	MUE		10.9	17.3	1.8	2.3		
	MAX		13.6	18.5	4.9	3.7		
	MUE%		0.30	0.47	0.04	0.07		
	MAX%		0.37	0.52	0.11	0.10		
Norcamphor	B_a^e	3465.1	3454.5	3448.0	3467.4	3462.1	31.4	-0.009
	B_b^e	2123.4	2115.4	2114.2	2123.5	2122.5	17.8	-0.006
	B_c^e	1903.2	1896.0	1894.4	1903.1	1902.0	15.6	-0.004
	MUE		8.6	11.7		0.8	1.7	
	MAX		10.6	17.1		2.3	3.0	
	MUE%		0.35	0.46	0.03	0.06		
	MAX%		0.38	0.49	0.07	0.09		

[a]SE equilibrium rotational constants obtained correcting the experimental ground state rotational constants by B3/6-31G* vibrational and B3/cc-pVTZ-F12$^-$ electronic corrections. Experimental data (truncated to 1 decimal figure) taken from References 410, 411, 412, and 413 for norbornane, norbornene, norbornadiene, and norcamphor, respectively. [b]Data taken from Reference 189.

The equilibrium rotational constants obtained with the revDSD and CCSD(T)-F12b/cc-pVDZ-F12 methods are already quite accurate, with average deviations of about 0.3 and 0.5%, respectively. However, attaining the desired accuracy of 0.1% requires the explicit inclusion of CV correlation effects. Both the BDPCS3 and PCS2 models successfully meet this requirement, delivering mean errors of approximately 0.05%. The accurate reproduction of the SE equilibrium rotational constants further confirms that the investigated schemes deliver quantitative predictions for rigid hydrocarbon backbones typical of terpene-like architectures.

To further evaluate the reliability of the proposed composite schemes, the four isomeric and tautomeric forms of vinylimidazole[193] shown in Figure 5.13 were selected as representative test cases, since accurate experimental spectroscopic data are available for comparison.[414] Beyond their spectroscopic interest, vinylimidazoles are biologically active fragments often associated with antifungal properties. The presence of a vinyl substituent at the 4- or 5-position of the imidazole ring confers distinctive conformational features and chemical reactivity. Moreover, their use as functional monomers in polymerization and copolymerization processes makes them relevant also from a technological perspective, bridging biomolecular

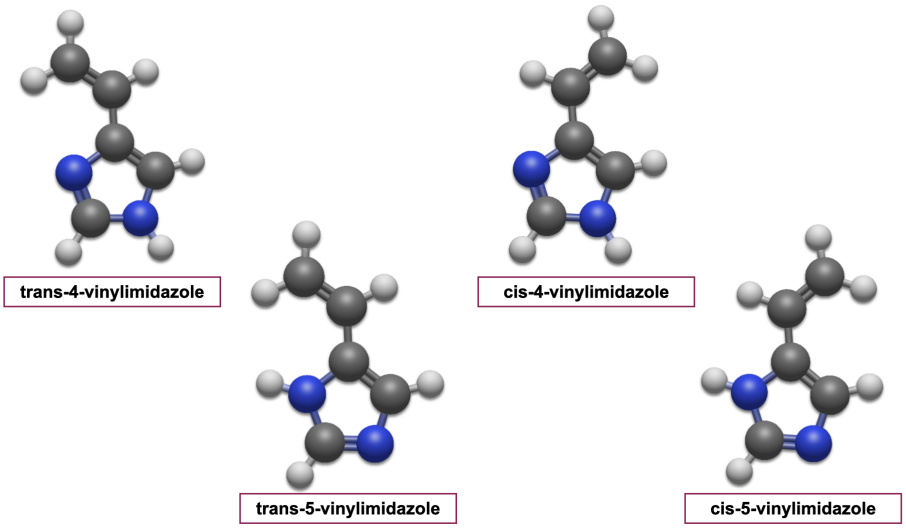


FIGURE 5.13: Cis and trans isomers of 4- and 5-vinylimidazole.

and materials-oriented applications.

The equilibrium rotational constants obtained at the revDSD/cc-pVTZ-F12^{-} and PCS2 levels of theory are compared with their corresponding semi-experimental values in Table 5.22.

TABLE 5.22: Comparison of SE equilibrium rotational constants and those computed at the revDSD/cc-pVTZ-F12$^-$ and PCS2 levels of theory (B_a^e, B_b^e, and B_c^e) for the four isomers of vinylimidazole, including B3/6-31G* vibrational corrections. The MAX, MUE, and their relative percentage counterparts (MAX% and MUE%) with respect to the SE values are also provided. All quantities, except percentage errors, are in MHz.

		SE[a]	revDSD / cc-pVTZ-F12$^-$	PCS2	$\Delta B^{\text{vib}}_{\text{B3/G-31G*}}$
	B_a^e	8091.0	8075.4	8088.9	66.6
	B_b^e	2072.2	2061.9	2070.8	12.7
	B_c^e	1649.0	1642.5	1648.7	10.3
trans-4-vinylimidazole	MUE		10.8	1.3	
	MAX		15.6	2.1	
	MUE%		0.38	0.04	
	MAX%		0.50	0.07	
	B_a^e	8194.1	8174.5	8191.5	70.2
	B_b^e	2043.3	2034.1	2042.3	11.3
	B_c^e	1635.6	1628.8	1634.7	9.2
trans-5-vinylimidazole	MUE		11.9	1.5	
	MAX		19.6	2.6	
	MUE%		0.37	0.05	
	MAX%		0.45	0.06	
	B_a^e	8101.0	8090.8	8101.0	70.9
	B_b^e	2030.7	2020.8	2029.3	11.2
	B_c^e	1624.4	1616.9	1622.8	8.4
cis-4-vinylimidazole	MUE		9.2	1.0	
	MAX		10.2	1.6	
	MUE%		0.36	0.06	
	MAX%		0.49	0.10	
	B_a^e	8170.4	8156.3	8169.5	75.0
	B_b^e	2011.6	2002.6	2010.2	10.5
	B_c^e	1614.6	1607.8	1613.3	7.2
cis-5-vinylimidazole	MUE		10.0	1.2	
	MAX		14.1	1.4	
	MUE%		0.35	0.05	
	MAX%		0.45	0.08	

[a]SE equilibrium rotational constants obtained correcting the experimental ground state rotational constants by B3/6-31G* vibrational corrections. Experimental data (truncated to 1 decimal figure) taken from Reference 414.

Also in this case, the PCS2 model provides extremely accurate results, fully consistent with the reference data.

The next group of investigated systems comprises molecular species derived from polycyclic aromatic hydrocarbons and related substituted aromatic compounds, which play a key role in several technological applications.

The first analyzed molecules are azulene, naphthalene, and quinoline,[190] whose structures are shown in Figure 5.14. These aromatic systems represent prototypical examples of PAHs and their nitrogen-substituted derivatives. They play a pivotal role in organic electronics, photophysics, and astrochemistry, where precise structural information is essential for interpreting spectroscopic signatures and designing functional molecular materials.

From a methodological viewpoint, they also constitute a valuable test of transferability from compact saturated systems to delocalized π-conjugated frameworks.

A comparison of the SE equilibrium rotational constants and those computed at the revDSD/cc-pVTZ-F12^{-}, CCSD(T)-F12b/cc-pVDZ-F12, BDPCS3, and PCS2 levels of theory is reported in Table 5.23.

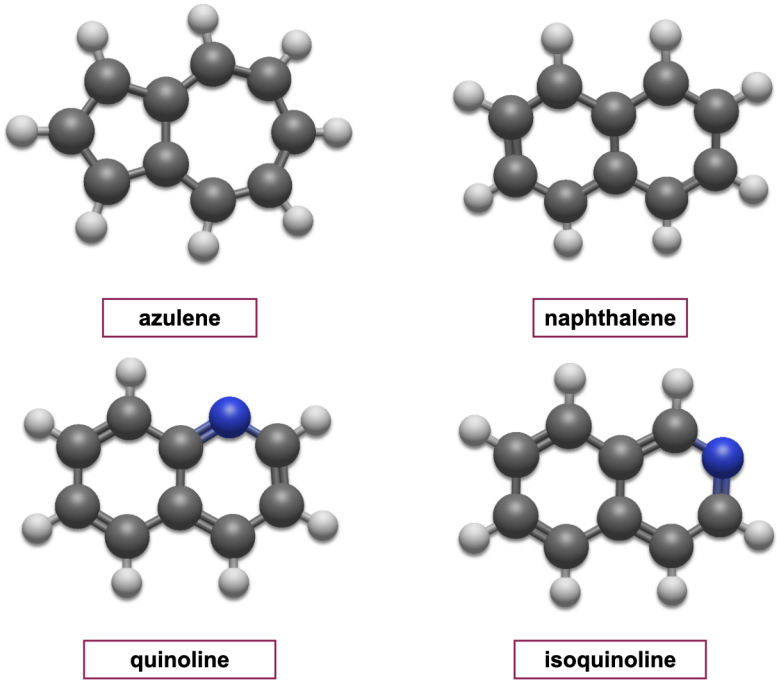


FIGURE 5.14: Structure of the azulene, naphthalene, quinoline, and isoquinoline.

TABLE 5.23: Comparison of SE equilibrium rotational constants and those computed at the revDSD/cc-pVTZ-F12$^-$, CCSD(T)-F12b/cc-pVDZ-F12, BDPCS3, and PCS2 levels of theory (B_a^e, B_b^e, and B_c^e) for azulene, naphthalene, quinoline, and isoquinoline, including B3/6-31G* vibrational corrections. The MAX, MUE, and their relative percentage counterparts (MAX% and MUE%) with respect to the SE values are also provided. All quantities, except percentage errors, are in MHz.

		SE[a]	revDSD	CCSD(T)-F12b	BDPCS3[b]	PCS2	$\Delta B^{\text{vib}}_{\text{B3/G-31G*}}$[c]
			cc-pVTZ-F12$^-$	cc-pVDZ			
Azulene	B_a^e	2844.9	2857.1	2848.5	2866.0	2860.2	20.2
	B_b^e	1250.1	1258.9	1257.4	1262.5	1262.5	7.7
	B_c^e	868.5	873.9	872.3	876.4	875.9	5.4
	MUE		3.6	7.5	1.4	0.7	
	MAX		5.0	13.6	3.9	1.9	
	MUE%		0.24	0.44	0.06	0.03	
	MAX%		0.29	0.48	0.13	0.07	
Naphthalene	B_a^e	3142.1	3132.8	3127.0	3142.3	3139.9	22.7
	B_b^e	1240.7	1237.5	1235.0	1241.0	1239.9	7.7
	B_c^e	889.5	887.1	885.3	889.7	888.9	5.6
	MUE		5.0	8.3	0.2	1.2	
	MAX		9.3	15.1	0.3	2.2	
	MUE%		0.28	0.47	0.02	0.07	
	MAX%		0.30	0.48	0.03	0.07	
Quinoline	B_a^e	3167.8	3160.0	3151.9	3169.6	3165.3	22.4
	B_b^e	1279.5	1275.9	1274.0	1279.6	1278.8	7.9
	B_c^e	911.4	908.9	907.3	911.6	910.8	5.7
	MUE		4.6	8.5	0.7	1.3	
	MAX		7.8	15.9	1.8	2.5	
	MUE%		0.26	0.46	0.03	0.07	
	MAX%		0.28	0.50	0.06	0.08	
Isoquinoline	B_a^e	3222.5	3213.0	3207.8	3222.7	3139.9	23.5
	B_b^e	1245.5	1242.5	1239.6	1246.1	1239.9	7.5
	B_c^e	898.3	896.0	984.1	898.6	888.9	5.6
	MUE		4.9	8.3	0.4	1.2	
	MAX		9.5	14.7	0.6	2.2	
	MUE%		0.26	0.47	0.03	0.07	
	MAX%		0.29	0.47	0.05	0.07	

[a] SE equilibrium rotational constants obtained correcting the experimental ground state rotational constants by B3/6-31G* vibrational and B3/cc-pVTZ-F12$^-$ electronic corrections. Experimental data (truncated to 1 decimal figure) taken from References 415, 416, and 417 for azulene, naphthalene, and quinoline and isoquinoline, respectively. [b] Data taken from Reference 190. [c] Vibrational corrections taken from Reference 188 for naphthalene, quinoline, and isoquinoline.

The PCS2 and BDPCS3 models provide highly accurate geometries for azulene, with deviations in the equilibrium rotational constants well within 0.1%. Such a level of accuracy can be achieved only when vibrational corrections are included, since their contribution amounts to about 0.5% of the corresponding equilibrium constants. Significantly larger deviations are observed for the CCSD(T)-F12b/cc-pVDZ-F12 method when CV correlation is neglected. Comparable accuracy is also obtained for naphthalene, quinoline, and isoquinoline, confirming the reliability of the proposed CSs for medium-sized aromatic systems.

The following group of investigated systems includes benzonitrile, phenylisocyanide, and fluorene,[191] whose molecular structures are depicted in Figure 5.15. These CN-substituted aromatic compounds bear technological and astrochemical relevance. Benzonitrile, for instance, was detected in interstellar molecular clouds, serving as a key tracer of aromatic chemistry in space, whereas phenylisocyanide and fluorene are relevant intermediates in

photochemical and optoelectronic applications. Their inclusion in the present set allows extending the validation of the PCSs to systems featuring polar substituents and extended π-conjugation.

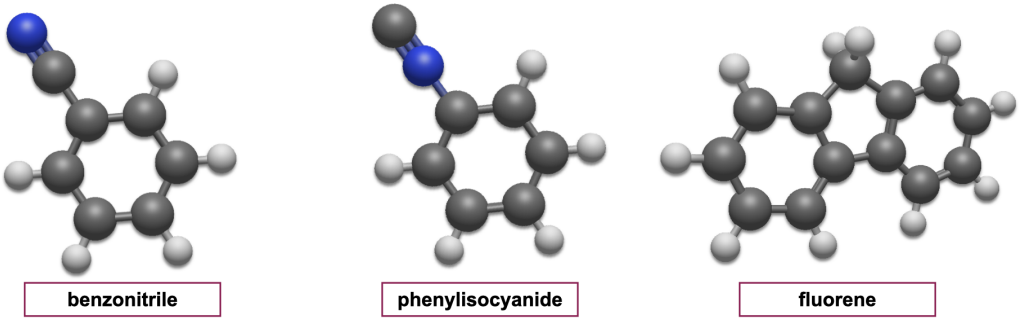


FIGURE 5.15: Structure of the benzonitrile, phenylisocyanide, and fluorene.

A comparison of the SE equilibrium rotational constants and those computed at the revDSD/cc-pVTZ-F12$^-$, BDPCS3, and PCS2 levels of theory is reported in Table 5.24.

TABLE 5.24: Comparison of SE equilibrium rotational constants and those computed at the revDSD/cc-pVTZ-F12$^-$, BDPCS3, and PCS2 levels of theory (B_a^e, B_b^e, and B_c^e) for benzonitrile, phenylisocyanide, and fluorene, including B3/6-31G* vibrational corrections. The MAX, MUE, and their relative percentage counterparts (MAX% and MUE%) with respect to the SE values are also provided. All quantities, except percentage errors, are in MHz.

		SE[a]	revDSD / cc-pVTZ-F12$^-$	BDPCS3[b]	PCS2	$\Delta B^{\text{vib}}_{\text{B3/G-31G*}}$
Benzonitrile	B_a^e	5695.7	5679.4	5695.9	5692.5	40.4
	B_b^e	1552.9	1547.4	1552.3	1551.4	6.0
	B_c^e	1220.2	1216.1	1219.8	1219.1	5.8
	MUE		8.6	0.4	1.9	
	MAX		16.3	0.6	3.2	
	MUE%		0.33	0.02	0.08	
	MAX%		0.35	0.04	0.10	
Phenylisocyanide	B_a^e	5701.0	5683.4	5699.7	5697.4	41.5
	B_b^e	1646.5	1640.9	1645.9	1644.3	6.7
	B_c^e	1277.6	1273.3	1277.1	1276.1	6.4
	MUE		9.2	0.8	2.4	
	MAX		17.6	1.3	3.6	
	MUE%		0.33	0.03	0.10	
	MAX%		0.34	0.04	0.13	
Fluorene	B_a^e	2191.7	2186.0	2192.7	2190.6	15.5
	B_b^e	590.1	588.0	589.8	589.4	3.4
	B_c^e	466.3	464.7	466.1	465.8	2.7
	MUE		3.1	0.5	0.8	
	MAX		5.7	1.0	1.1	
	MUE%		0.32	0.05	0.09	
	MAX%		0.36	0.05	0.12	

[a]SE equilibrium rotational constants obtained correcting the experimental ground state rotational constants by B3/6-31G* vibrational corrections. Experimental data (truncated to 1 decimal figure) taken from References 418, 419, and 420 for benzonitrile, phenylisocyanide, and fluorene, respectively. [b]Data taken from Reference 191.

The equilibrium rotational constants obtained at the BDPCS3 and PCS2 levels, once augmented with vibrational corrections, reproduce the experimental ground state rotational constants with relative errors below 0.1%. For fluorene, however, it should be noted that an accuracy better than 0.1% is not particularly meaningful, as the intrinsic uncertainty of the vibrational corrections computed at the B3/6-31G* level is already of this order of magnitude.

The final group of systems comprises nitroaromatic compounds relevant to explosives and their taggants. Specifically, nitrobenzene, 2-nitrophenol, and 2-nitrotoluene were considered (structures are shown in Figure 5.16). Nitroaromatic compounds constitute a class of molecules of considerable technological and forensic interest. They are central components of energetic materials, propellants, and detection taggants, and their accurate spectroscopic characterization is essential for both environmental monitoring and security-related applications. Their structural diversity, involving strong electron-withdrawing nitro groups and hydrogen-bonding substituents, provides a challenging application for assessing the robustness and transferability of the PCS formulations.

A direct comparison of the SE equilibrium rotational constants and those computed at the revDSD/cc-pVTZ-F12^{-}, BDPCS3 and PCS2 levels of theory is reported in Table 5.25.

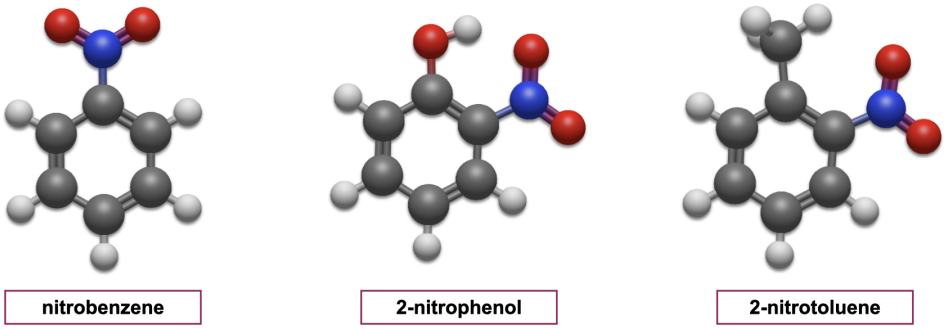


FIGURE 5.16: Structure of the nitrobenzene, 2-nitrophenol, and 2-nitrotoluene.

TABLE 5.25: Comparison of SE equilibrium rotational constants and those computed at the revDSD/cc-pVTZ-F12$^-$, BDPCS3, and PCS2 levels of theory (B_a^e, B_b^e, and B_c^e) for nitrobenzene, 2-nitrophenol and 2-nitrotoluene, including B3/6-31G* vibrational corrections. The MAX, MUE, and their relative percentage counterparts (MAX% and MUE%) with respect to the SE values are also provided. All quantities, except percentage errors, are in MHz.

		SE[a]	revDSD / cc-pVTZ-F12$^-$	BDPCS3[b]	PCS2	$\Delta B^{\text{vib}}_{\text{B3/G-31G*}}$
Nitrobenzene	B_a^e	3996.1	3984.2	3997.4	3994.4	28.0
	B_b^e	1296.5	1290.5	1294.7	1295.7	9.6
	B_c^e	979.0	974.8	977.9	978.4	6.3
	MUE		7.4	1.4	1.0	
	MAX		11.9	1.8	1.7	
	MUE%		0.40	0.09	0.06	
	MAX%		0.46	0.14	0.06	
2-Nitrophenol	B_a^e	2381.6	2375.2	2383.1	2381.3	12.9
	B_b^e	1284.9	1280.1	1284.3	1284.5	8.8
	B_c^e	834.7	831.8	834.5	834.4	4.9
	MUE		4.7	0.8	0.3	
	MAX		6.4	1.5	0.4	
	MUE%		0.33	0.04	0.03	
	MAX%		0.37	0.06	0.04	
2-Nitrotoluene	B_a^e	2282.7	2273.6	2280.7	2280.0	17.5
	B_b^e	1256.7	1252.3	1256.5	1258.3	9.5
	B_c^e	840.6	836.6	839.3	838.3	6.4
	MUE		5.8	1.2	2.2	
	MAX		9.1	2.0	2.7	
	MUE%		0.41	0.09	0.17	
	MAX%		0.48	0.15	0.27	

[a]SE equilibrium rotational constants obtained correcting the experimental ground state rotational constants by B3/6-31G* vibrational corrections. Experimental data (truncated to 1 decimal figure) taken from References 421, 422, and 423 for nitrobenzene, 2-nitrophenol, and 2-nitrotoluene, respectively. [b]Data taken from Reference 192.

For nitrobenzene and 2-nitrophenol, the BDPCS3 and PCS2 equilibrium rotational constants, corrected for vibrational effects, reproduce the SE values with excellent accuracy, typically within the 0.1% target. These results confirm that the proposed composite schemes are robust and transferable to substituted aromatic systems of technological relevance, including species where precise spectroscopic characterization is essential for detection and monitoring applications. In the case of 2-nitrotoluene, the same overall behavior is observed, although the distortions introduced by the methyl substituent are more pronounced than those caused by the hydroxyl group. Despite the unusually large deviations obtained with the PCS2 model, the BDPCS3 approach still achieves spectroscopic accuracy, confirming its reliability even in the presence of increased structural complexity.

Overall, the diverse set of molecules investigated, ranging from biologically relevant nucleobases to terpenoid frameworks, heteroaromatic compounds, and nitro-substituted systems,

demonstrates the broad applicability and transferability of the PCSs. Their consistent performance across classes of increasing structural and electronic complexity confirms that these formulations provide a unified and computationally efficient route to equilibrium geometries of spectroscopic accuracy across a wide range of molecular systems.

5.2.4 Harmonic Frequencies, Zero-point Vibrational Energies, and Thermodynamic Functions

For harmonic frequencies, considering the results of the previous analysis carried out with the junChS-F12 model[184] (Section 4.2.3 of Chapter 4), neither the CV[348] nor the CBS corrections were included. This choice allows a direct assessment of the revDSD/cc-pVTZ-F12 level of theory employed for systems in the 5-15 atom range.

The validation was performed using a set of 28 molecules for which accurate experimental or high-level theoretical reference harmonic frequencies are available, namely: H_2O,[371] HCN,[372] CO_2,[370] C_2H_2,[366–368] CO,[363] HF,[363] N_2,[363] N_2O,[373] H_2,[75] $OH^\bullet$,[75] Cl_2,[363] ClCN,[369] ClF,[364,365] CS,[363] HCl,[363] SiO,[365] NH_3,[374] CF_2,[377] H_2CO,[379] C_2H_2O,[382] CH_4,[378] HNO,[383] PH_3,[381] CCl_2,[376] H_2CS,[380] C_2H_4,[375] BH_3,[385] and HOCl.[384]

The corresponding error statistics are summarized in Table 5.26, which compares results obtained at the CCSD(T)-F12b/jun-cc-pV(T+*d*)Z level and at the revDSD double-hybrid in conjunction with the jun-cc-pV(T+*d*)Z and cc-pVTZ-F12 basis sets.

Table 5.26: MAX, MUE, and RMSD (cm^{-1}) of harmonic frequencies computed at the CCSD(T)-F12b/jun-cc-pV(T+*d*)Z, revDSD/jun-cc-pV(T+*d*)Z, and revDSD/cc-pVTZ-F12 levels of theory.

	$\omega_{\text{CCSDT(T)-F12b}}$	ω_{revDSD}	
	jun-cc-pV(T+*d*)Z	jun-cc-pV(T+*d*)Z	cc-pVTZ-F12
2nd-row (18 molecules)			
MAX	11.2	46.5	31.6
MUE	3.5	6.3	6.6
RMSD	4.3	9.4	9.3
3rd-row (10 molecules)			
MAX	12.3	22.2	17.4
MUE	3.5	8.4	8.4
RMSD	5.0	10.5	9.6

The RMSD obtained at the CCSD(T)-F12b/jun-cc-pV(T+*d*)Z level of theory amounts to 5 cm^{-1}, whereas the revDSD approach yields a value around 10 cm^{-1}. Such accuracy is generally sufficient for most spectroscopic applications and, more importantly, for the evaluation of ZPVEs and partition functions.[395,424]

Zero-point and finite-temperature contributions are typically evaluated at the DFT level within the rigid-rotor harmonic-oscillator approximation, often using scaling factors to include anharmonic effects.[425–428] The availability of resonance-free formulations for anharmonic ZPVEs,[429,430] together with black-box procedures for anharmonic resonances[431] and for the detection and treatment of hindered rotations,[432] enables an automated evaluation of anharmonic vibrational frequencies and ZPVEs. Based on these quantities, vibrational partition functions can be efficiently computed by employing the anharmonic ZPVEs and the fundamental frequencies (Δ_i) within the same expression as for the harmonic vibrational partition function

$$Q_{\text{vib}} = \frac{\exp\left(-\frac{\text{ZPVE}}{k_B T}\right)}{\prod_i \left[1 - \exp\left(-\frac{\Delta_i}{k_B T}\right)\right]}. \tag{5.13}$$

The simple perturbation theory (SPT) formulation expressed by Equation 5.13 yields analytical expressions for thermodynamic functions identical to those derived from the harmonic oscillator model and provides results in excellent agreement with accurate reference data.[433]

In this framework, the reliability of the revDSD/cc-pVTZ-F12 level of theory was assessed against experimental ZPVEs and entropies reported for 21 closed-shell and 9 open-shell semi-rigid molecules.[395]

The results, summarized in Table 5.27, confirm that the revDSD/cc-pVTZ-F12 level of theory delivers highly accurate thermodynamic properties for semi-rigid systems, with RMSDs from experimental reference data of 0.11 kJ mol^{-1} and 0.29 J (mol K)$^{-1}$ respectively for anharmonic ZPVEs and entropic contributions for closed-shell molecules, that increase to 0.58 kJ mol^{-1} and 1.21 J (mol K)$^{-1}$ for open-shell systems.

TABLE 5.27: ZPVEs (in kJ mol^{-1}) and entropies at 298 K and 1 atm (in J $(mol\,K)^{-1}$) for 21 closed-shell and 9 open-shell molecules evaluated at the revDSD/cc-pVTZ-F12 level of theory, for both harmonic and anharmonic (VPT2 and SPT) quantities.

	$ZPVE_{exp}$ [a]	$ZPVE_{harm}$ [b]	$ZPVE_{VPT2}$ [b]	S_{exp} [a]	S_{harm} [b]	S_{SPT} [b]
Closed-Shell Molecules						
HF	24.45	24.56	24.35	173.75	173.66	173.50
HCl	17.88	17.94	17.83	186.90	186.63	186.60
H_2	26.38	26.50	26.24	130.63	130.33	130.30
N_2	13.90	13.98	13.94	191.63	191.62	191.60
F_2	5.90	5.88	5.85	202.81	202.44	202.40
CO	12.90	12.95	12.92	197.66	197.65	197.70
Cl_2	3.39	3.36	3.35	223.11	222.73	222.60
CO_2	30.27	30.50	30.38	213.82	213.77	213.80
CS_2	18.21	18.23	18.18	238.02	237.78	237.80
H_2O	55.43	56.22	55.33	188.82	188.75	188.62
H_2S	39.65	40.07	39.57	205.82	205.66	205.54
HOF	36.17	36.59	36.04	226.80	226.54	226.45
HOCl	34.29	34.76	34.26	236.51	236.32	236.24
N_2O	28.47	28.66	28.50	219.97	219.85	219.90
HCN	41.66	42.09	41.69	201.85	201.63	201.60
SO_2	18.05	18.12	18.05	248.28	248.36	248.25
C_2H_2	69.33	70.12	69.38	200.92	200.46	200.40
H_2CO	69.25	70.08	69.14	218.97	218.68	218.56
NH_3	89.01	90.38	88.82	192.76	192.45	192.46
CH_4	116.35	117.82	116.13	186.23	186.15	186.08
C_2H_4	132.60	133.98	132.35	219.35	219.05	219.00
MAX		1.47	0.25		0.46	0.52
MUE		0.45	0.08		0.20	0.25
RMSD		0.65	0.11		0.24	0.29
Open-Shell Molecules						
$OH(^2\Pi)$ [c]	22.15	22.45	22.26	183.72	183.97	183.90
$SH(^2\Pi)$ [c]	15.99	16.24	16.11	195.77	197.89	197.80
$NO(^2\Pi)$ [c]	11.35	11.61	11.72	210.76	211.06	211.00
$NH_2(^2B_1)$	48.23	50.46	49.68	194.85	194.64	194.52
$HCO(^2A')$	33.87	34.31	33.67	224.66	224.31	224.20
$HO_2(^2A'')$	36.76	37.64	37.05	229.27	228.80	228.68
$CH_3(^2A'')$	77.37	78.49	77.83	194.18	194.83	193.52
t-$HOCO(^2A')$	54.85	55.19	54.38		251.49	251.71
$CH_3CO(^2A')$	111.75	1113.78	112.18	267.62	268.84	270.16
MAX		2.23	1.45		2.11 [d]	2.54 [d]
MUE		0.87	0.43		0.70 [d]	0.88 [d]
RMSD		1.14	0.58		0.93 [d]	1.21 [d]

[a] Data taken from Reference 395. [b] Data taken from Reference 176. [c] ZPVEs calculated taking into account the proper electronic state degeneracy. [d] t-HOCO open-shell molecule not included in the error analysis due to the lack of the reference S_{exp} value.

5.3 Conclusions

In this Chapter, the PCS family of composite approaches has been presented as a computationally efficient framework for high-accuracy thermochemical, structural, and rotational–vibrational spectroscopy. The PCS models build upon traditional ChS-type schemes through a refined hierarchical organization of the correlation energy and a consistent definition of theoretical levels across all computed properties. At the core of the framework lies the energetic component, where separate extrapolations of MP2 and post-MP2 contributions to the CBS limit, together with the systematic use of cc-pVnZ-F12 basis sets and optimized treatments of third-row and CV effects, yield smooth convergence and robust transferability. The resulting F12-based variant provides a reliable energetic foundation for medium-sized systems (5–15 atoms), while the FNO extension enables accurate yet affordable CC calculations for molecules up to about 50 atoms. For equilibrium geometries, F12-based CC optimizations complemented by MP2 CV corrections are employed for smaller systems, whereas larger molecules can be efficiently treated using double-hybrid functionals with MP2-level corrections or simplified bond-corrected PCS formulations. Vibrational properties are computed using hybrid or double-hybrid density functionals, with anharmonic effects treated via VPT2 for semi-rigid systems. Extensive benchmarks confirm the high and systematic accuracy of the PCS family across a broad spectrum of molecular systems, demonstrating its effectiveness as a unified and scalable framework for high-level molecular modeling.

Chapter 6

Beyond the Gold Standard: FNOs for Post-CCSD(T) Correlation

Following the development of the Pisa Composite Scheme family, the final part of this PhD research focuses on extending the protocol toward the *platinum standard* and, ultimately, spectroscopic accuracy.

As discussed in previous Chapters, post-CCSD(T) methods are generally required to reduce the intrinsic error of composite approaches below 1 kJ mol^{-1}. However, their computational scaling, $O(N_o^m N_v^{m+2})$, where m is the excitation level and N_o and N_v are the numbers of occupied and canonical virtual orbitals, respectively, makes such treatments unfeasible for medium- and large-size systems. In common basis sets, the virtual space ($N_v \gg N_o$) is the main cost driver, and the slow convergence of CC energies with respect to basis-set size further exacerbates the problem. Consequently, high-order CC models, such as CCSDT(Q) or CCSDTQ(P), although regarded as the platinum standard, remain limited to small molecules due to prohibitive computational requirements.

Several strategies have been proposed to alleviate these limitations, including (i) active-space formulations,[300,434–438] (ii) basis-set truncation or transformation techniques,[287,300,439] and (iii) wave function or integral decomposition methods.[440–444] Among these, the frozen natural orbital approximation, described in Section 2.8 of Chapter 2, has emerged as a particularly robust and widely adopted strategy for reducing the virtual space while preserving accuracy.

This Chapter presents a preliminary exploration of the FNO approximation as a means to construct a unified, threshold-based framework that achieves controllable accuracy at reduced computational cost for post-CCSD(T) correlation, thereby enabling the treatment of molecular systems larger than those currently accessible with conventional high-level methods. The approach relies on a single, sufficiently large parent basis set across all excitation levels, with the virtual space progressively compressed according to occupation thresholds, t_{NO}, which determine the number of correlated orbitals (COs) retained at each excitation rank. This occupation-based reduction guarantees size-consistent energies and provides a systematic means of balancing accuracy and efficiency. Larger thresholds discard more orbitals, lowering cost but reducing precision, whereas tighter thresholds retain nearly

complete correlation contributions with a moderate gain in computational expenses.

In addition to direct truncation, a projection-based variant of the FNO approach is also considered. In this strategy, NOs generated from a large basis set are mapped onto a smaller one and subsequently recanonicalized, retaining the same number of virtual orbitals as in the smaller basis. This procedure enables an independent assessment of two distinct effects: (i) improvement of the occupied-orbital description, whose size is fixed by the number of electrons but whose quality benefits from a larger basis, and (ii) expansion of the virtual manifold, whose size depends on t_{NO} and increases as the threshold is tightened. By allowing these two contributions to be analyzed separately, the projection-based scheme provides additional insight into the interplay between the occupied and virtual subspaces and complements the direct truncation approach.

6.1 Methods and Computational Details

All calculations were carried out with the MRCC program,[349,445] within the frozen core approximation. Correlation-consistent cc-pVnZ basis sets (n = D, T, Q, 5) were employed, and the computational study encompassed CC models beyond CCSD, up to CCSDTQ(P). Both full and perturbative treatments of triple and quadruple excitations were considered, together with perturbative pentuples. To alleviate the steep computational cost associated with these high-level methods, the FNO approximation was employed in combination with NOs generated at either the MP2 or CCSD level. In this framework, the virtual space was truncated by discarding orbitals with occupation numbers below a threshold, t_{NO}, and the retained orbitals were recanonicalized to yield the final set of COs. Neither the NAF[244] nor the density fitting approximations were employed, as both offer negligible advantages for post-CCSD(T) computations.

6.2 Benchmark Results and Discussion

The performance of the truncation strategies was assessed by analyzing the energy differences between the following pairs of CC models: CCSD(T) – CCSD, denoted as *(T)*; CCSDT – CCSD(T), denoted as *full T*; CCSDT(Q) – CCSDT, denoted as *(Q)*; CCSDT(Q)$_\Lambda$ – CCSDT, denoted as *(Q)*$_\Lambda$; CCSDT(Q)/A – CCSDT, denoted as *(Q)/A*; CCSDT(Q)/B – CCSDT, denoted as *(Q)/B*; CCSDTQ – CCSDT, denoted as *full Q*; and CCSDTQ(P) – CCSDTQ, denoted as *(P)*. MUEs, RMSDs, and MAXs were employed as statistical indicators, with normalized values obtained by dividing the absolute errors by the number of correlated orbitals (n_{CO}).

The benchmark analysis started with reference to the Karton test set,[78] which reports atomization energies for 16 molecules containing elements from the first three rows of the periodic table, computed up to the CCSDTQ level, namely BH_3, CH, CH_2, OH, H_2O, HF, AlH_3, H_2S, HCl, B_2, C_2, CO, N_2, CS, and P_2. The study was subsequently extended to more demanding cases, including the Knizia–Adler–Werner test set,[239] the DBH24 database,[99,394]

and the multi-reference subset of the W4-17 benchmark (W4-17-MR),[84] which comprises particularly challenging systems such as B_2, BN, C_2, O_3, OF, F_2O, FO_2, FOOF, OClO, Cl_2O, ClO_3, ClOO, ClOOCl, ClF_3, ClF_5, S_3, and S_4.

Table 6.1 reports the total, minimum, maximum, and average numbers of COs, together with the percentage of virtual orbitals relative to the full space, for the various basis sets, t_{NO} thresholds, and density sources (MP2 or CCSD), for the Karton, KAW, DBH24, and W4-17 MR test sets. Both direct and projection-based approaches are included, allowing the influence of the basis set and the chosen truncation strategy on the number of COs to be directly compared.

For reference, thresholds of $2.5 \cdot 10^{-4}$, $2.5 \cdot 10^{-5}$, and virtually $2.5 \cdot 10^{-6}$ for larger basis sets yield numbers of COs comparable to those obtained with the full cc-pVDZ, cc-pVTZ, and cc-pVQZ bases, respectively.

TABLE 6.1: Total, minimum, maximum, and average numbers of COs, together with the percentage of virtual orbitals relative to the full space (VO%), for each combination of basis set, t_{NO} threshold, and density source in the Karton, KAW, DBH24, and W4-17 MR test sets.[a]

Basis set	t_{NO}	Total		Minimum		Maximum		Average		VO%	
		MP2	CCSD	MP2	CCSD	MP2	CCSD	MP2	CCSD	MP2	CCSD
Karton test set											
cc-pVDZ	full[b]	503		5		28		19		100.00	
cc-pVTZ	full[b]	1179		14		71		45		100.00	
	$1.0\cdot10^{-3}$	364	383	9	9	23	23	14	15	25.83	27.26
	$5.0\cdot10^{-4}$	422	445	9	9	27	29	16	17	30.89	33.15
	$2.5\cdot10^{-4}$	531	537	11	9	37	37	20	21	41.18	41.59
	$1.0\cdot10^{-4}$	733	711	14	10	49	47	28	27	61.04	58.88
	$5.0\cdot10^{-5}$	891	841	14	14	53	51	34	32	76.98	72.20
	$2.5\cdot10^{-5}$	1001	961	14	14	57	56	39	37	86.36	82.94
	$1.0\cdot10^{-5}$	1089	1060	14	14	62	57	42	41	93.00	90.54
cc-pVQZ	full	2274		30		144		88		100.00	
	$1.0\cdot10^{-3}$	384	401	9	9	30	30	15	15	13.12	13.77
	$5.0\cdot10^{-4}$	454	464	9	9	31	30	18	18	16.43	16.77
	$2.5\cdot10^{-4}$	573	587	11	9	41	43	22	23	21.81	22.40
	$1.0\cdot10^{-4}$	825	781	16	12	51	49	32	30	34.87	32.65
	$5.0\cdot10^{-5}$	1010	944	23	16	63	60	39	36	43.83	40.46
	$2.5\cdot10^{-5}$	1208	1114	25	23	82	72	47	43	52.89	48.48
	$1.0\cdot10^{-5}$	1570	1419	30	24	93	85	60	55	70.64	63.39
cc-pV5Z	full	3890		55		255		150		100.00	
	$1.0\cdot10^{-3}$	409	424	9	9	55	55	16	16	7.60	7.93
	$5.0\cdot10^{-4}$	479	493	9	9	55	55	18	19	9.52	9.81
	$2.5\cdot10^{-4}$	604	616	11	9	55	55	23	24	12.77	13.06
	$1.0\cdot10^{-4}$	862	812	16	12	55	55	33	31	20.56	19.10
	$5.0\cdot10^{-5}$	1066	991	23	16	70	62	41	38	26.17	24.05
	$2.5\cdot10^{-5}$	1320	1199	25	23	84	78	51	46	33.11	30.00
	$1.0\cdot10^{-5}$	1800	1541	40	24	114	95	69	59	47.23	39.62
KAW test set											
cc-pVDZ	full[b]	1870		5		56		28		100.00	
cc-pVTZ	full[b]	4422		14		142		67		100.00	
	$1.0\cdot10^{-3}$	1317	1377	5	5	44	45	20	21	24.84	26.29
	$5.0\cdot10^{-4}$	1579	1679	5	5	57	60	24	25	30.90	33.39
	$2.5\cdot10^{-4}$	2004	2063	5	5	71	74	30	31	41.51	43.00
	$1.0\cdot10^{-4}$	2808	2811	9	6	96	96	43	43	62.35	62.25
	$5.0\cdot10^{-5}$	3285	3225	10	10	106	106	50	49	74.80	73.21
	$2.5\cdot10^{-5}$	3728	3603	14	14	112	112	57	55	84.92	82.09
	$1.0\cdot10^{-5}$	4098	3992	14	14	129	120	62	61	92.90	90.48
DBH24 test set											
cc-pVDZ	full[b]	1266		5		59		33		100.00	
cc-pVTZ	full[b]	3049		14		143		80		100.00	
	$1.0\cdot10^{-3}$	837	878	5	5	38	40	22	23	22.18	23.81
	$5.0\cdot10^{-4}$	1056	1107	5	5	56	58	28	29	29.59	31.69
	$2.5\cdot10^{-4}$	1340	1389	5	5	74	75	35	37	39.66	41.49
	$1.0\cdot10^{-4}$	1845	1856	9	6	86	86	49	49	59.55	59.86
W4-17 MR test set											
cc-pVDZ	full[b]	806		13		78		34		100.00	
cc-pVTZ	full[b]	1798		29		174		75		100.00	
	$1.0\cdot10^{-3}$	719	733	9	9	73	73	30	31	32.31	33.12
	$5.0\cdot10^{-4}$	827	862	9	9	78	82	35	36	38.39	40.77
	$2.5\cdot10^{-4}$	989	1013	11	9	94	94	41	42	47.75	49.93
	$1.0\cdot10^{-4}$	1497	1554	14	23	155	157	62	65	77.85	85.68

[a] All 26, 66, 38, and 24 species (atoms and molecules) in the Karton, KAW, DBH24, and W4-17 test sets, respectively, have been included in the statistics, VO% values exclude the H atom, for which FNO truncation is not performed. [b] The same values apply to projected truncated basis sets. For the Karton test set, the VO% for the projected cc-pVDZ relative to the full cc-pVTZ, cc-pVQZ, and cc-pV5Z bases are 37.57%, 18.80%, and 10.86%, respectively; for the projected cc-pVTZ relative to the full cc-pVQZ and cc-pV5Z are 49.98% and 28.86%. For the KAW, DBH24, W4-17 MR test sets, the VO% for the projected cc-pVDZ relative to the full cc-pVTZ are 37.14%, 36.55%, and 37.94%, respectively.

6.2.1 Triple Excitations

The first step toward the definition of the truncation strategy consisted of evaluating the effect of different schemes for CCSD(T) computations, for which benchmark results with very large basis sets are available, before moving on to post-CCSD(T) contributions. Figure 6.1 shows the RMSD and its counterpart normalized to the number of COs for the (T) contribution evaluated accordingly to different truncation thresholds, with respect to the CBS limit, using cc-pV(5,6)Z reference values for the atomization energies of the Karton test set. The complete statistical analysis is reported in Section A.4 of Appendix A.

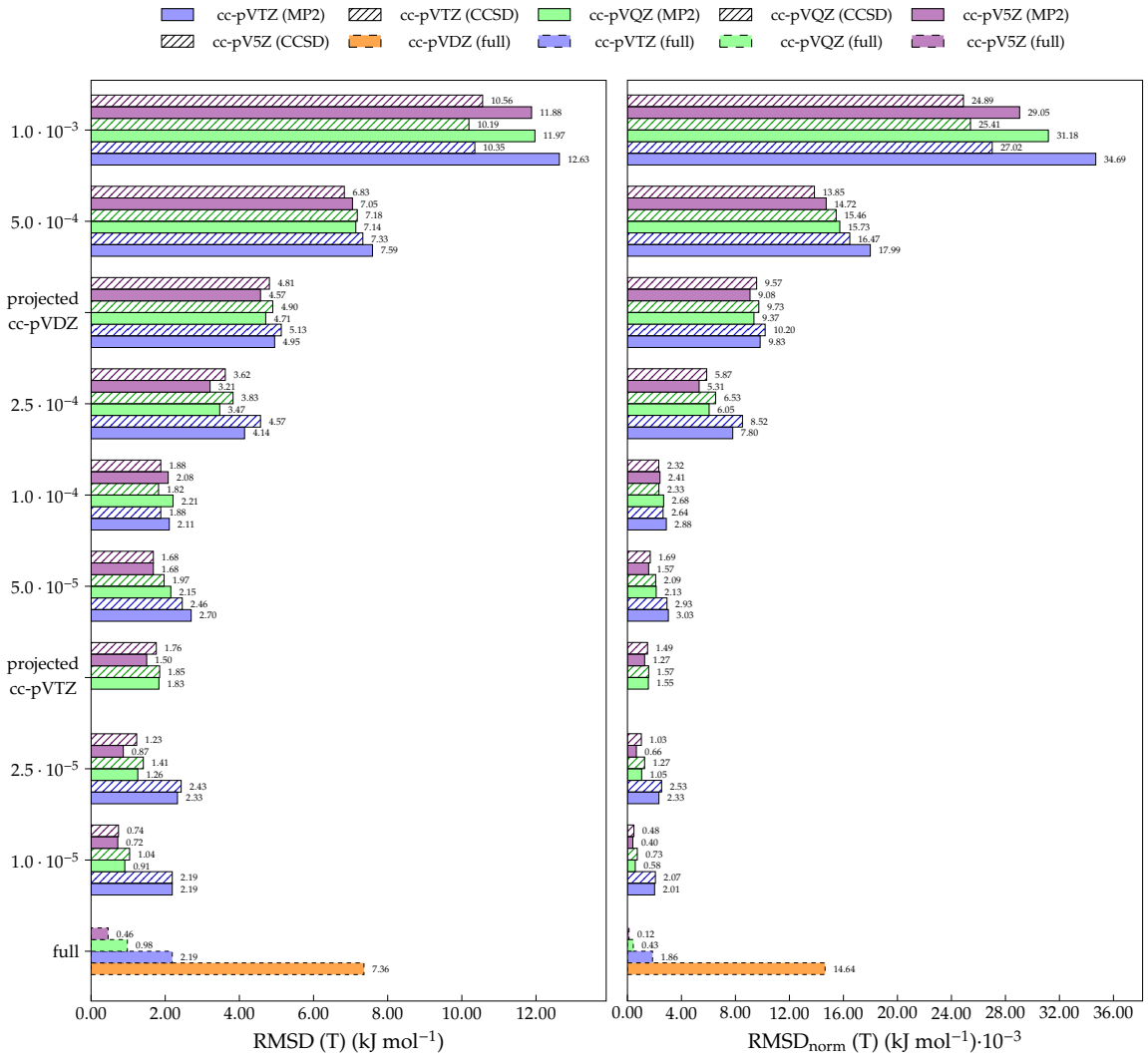


FIGURE 6.1: RMSD (left) and RMSD normalized to the number of COs ($RMSD_{norm}$, right) for the (T) contribution to the atomization energies in the Karton test set according to different FNO compression schemes, referenced to cc-pV(5,6)Z values. Results obtained with orbitals selected from MP2-NO occupations are labeled as (MP2), those from CCSD-NO occupations as (CCSD), while calculations performed in the untruncated virtual space are denoted as (full). Projected cc-pVDZ and cc-pVTZ results correspond to the same number of COs as in the respective parent basis sets.

As can be seen in Figure 6.1, at large truncation thresholds, which correspond to a strong compression of the CO space, CCSD-based NOs provide higher accuracy than those derived from MP2 densities. However, as the threshold is tightened, this advantage progressively diminishes, and MP2-based orbitals often yield comparable or slightly better results. A clear trend emerges when comparing truncated large basis sets with smaller uncompressed ones: compressed larger bases consistently outperform their smaller counterparts when the number of COs is similar. For example, compressing cc-pVnZ (n = T, Q, 5) to match the size of cc-pVDZ by selecting the NOs with the largest overlap with the cc-pVDZ orbitals and recanonicalizing reduces the RMSD by about 2 – 5 kJ mol^{-1} across all basis sets and density sources. Similarly, the projected cc-pVTZ basis yields an improvement of about 0.3 – 0.7 kJ mol^{-1} over its uncompressed counterpart. Notably, applying a $2.5 \cdot 10^{-4}$ threshold to larger basis sets delivers consistently better accuracy than the uncompressed cc-pVDZ basis while involving only a slightly larger n_{CO}. Even at $5.0 \cdot 10^{-4}$, cc-pV5Z with CCSD-derived orbitals still deliver better results than uncompressed cc-pVDZ despite comprising fewer COs.

Table 6.2 presents the statistical indicators for the (T) contribution in the KAW test set.

TABLE 6.2: MUE, RMSD, and MAX values, together with their normalized counterparts (*norm*, obtained by dividing by the total number of COs), for the (T) contribution to atomization, closed-shell, and open-shell reaction energies of the KAW test set according to different FNO compression schemes, relative to the cc-pVTZ reference. All values are in kJ mol^{-1}.

	$1.0{\cdot}10^{-3}$ cc-pVTZ		$5.0{\cdot}10^{-4}$ cc-pVTZ		cc-pVDZ	projected cc-pVDZ		$2.5{\cdot}10^{-4}$ cc-pVTZ		$1.0{\cdot}10^{-4}$ cc-pVTZ		$5.0{\cdot}10^{-5}$ cc-pVTZ		$2.5{\cdot}10^{-5}$ cc-pVTZ		$1.0{\cdot}10^{-5}$ cc-pVTZ	
	MP2	CCSD	MP2	CCSD		MP2	CCSD	MP2	CCSD	MP2	CCSD	MP2	CCSD	MP2	CCSD	MP2	CCSD
Atomization energy																	
MUE	10.75	9.39	6.29	5.80	5.99	3.12	3.43	2.11	2.10	0.86	0.90	0.75	0.67	0.33	0.44	0.08	0.09
RMSD	12.81	11.17	7.56	7.01	7.09	3.78	4.13	2.75	2.57	1.07	1.15	0.98	0.91	0.40	0.51	0.11	0.13
MAX	28.32	28.74	17.62	14.45	14.38	10.50	11.42	10.46	5.51	2.97	2.69	3.28	3.22	1.05	1.15	0.35	0.36
MUE_{norm}	0.00816	0.00682	0.00398	0.00345	0.00320	0.00167	0.00184	0.00105	0.00102	0.00031	0.00032	0.00023	0.00021	0.00009	0.00012	0.00002	0.00002
$\text{RMSD}_{\text{norm}}$	0.00973	0.00811	0.00479	0.00417	0.00379	0.00202	0.00221	0.00137	0.00125	0.00038	0.00041	0.00030	0.00028	0.00011	0.00014	0.00003	0.00003
MAX_{norm}	0.02151	0.02087	0.01116	0.00861	0.00769	0.00562	0.00611	0.00522	0.00267	0.00106	0.00096	0.00100	0.00100	0.00028	0.00032	0.00008	0.00009
Closed-shell reaction energy																	
MUE	4.66	4.09	4.51	2.61	1.49	2.11	1.74	1.07	1.19	0.98	0.78	0.92	1.03	0.26	0.36	0.12	0.14
RMSD	6.67	5.79	5.71	3.52	2.06	3.03	2.57	1.52	1.55	1.25	1.00	1.15	1.31	0.38	0.56	0.17	0.18
MAX	20.86	19.73	12.47	10.39	5.92	8.69	7.97	4.54	4.78	3.94	2.78	2.55	3.05	1.26	1.93	0.41	0.43
MUE_{norm}	0.00354	0.00297	0.00286	0.00156	0.00080	0.00113	0.00093	0.00053	0.00058	0.00035	0.00028	0.00028	0.00032	0.00007	0.00010	0.00003	0.00003
$\text{RMSD}_{\text{norm}}$	0.00507	0.00420	0.00362	0.00210	0.00110	0.00162	0.00137	0.00076	0.00075	0.00045	0.00036	0.00035	0.00041	0.00010	0.00016	0.00004	0.00004
MAX_{norm}	0.01584	0.01432	0.00790	0.00619	0.00317	0.00465	0.00426	0.00226	0.00232	0.00140	0.00099	0.00078	0.00095	0.00034	0.00054	0.00010	0.00011
Open-shell reaction energy																	
MUE	9.82	9.43	5.22	3.84	3.42	2.19	2.24	1.42	1.38	1.23	1.56	0.58	0.80	0.26	0.35	0.09	0.09
RMSD	13.45	12.80	7.63	5.41	4.55	2.86	2.85	1.99	1.79	2.13	2.38	0.75	1.03	0.33	0.44	0.11	0.12
MAX	40.51	37.81	23.04	21.79	18.49	8.28	8.42	6.59	4.72	8.89	7.71	2.44	3.48	0.93	1.26	0.28	0.39
MUE_{norm}	0.00746	0.00685	0.00331	0.00229	0.00183	0.00117	0.00120	0.00071	0.00067	0.00044	0.00055	0.00018	0.00025	0.00007	0.00010	0.00002	0.00002
$\text{RMSD}_{\text{norm}}$	0.01021	0.00930	0.00483	0.00322	0.00243	0.00153	0.00152	0.00099	0.00087	0.00076	0.00085	0.00023	0.00032	0.00009	0.00012	0.00003	0.00003
MAX_{norm}	0.03076	0.02746	0.01459	0.01298	0.00989	0.00443	0.00450	0.00329	0.00229	0.00317	0.00274	0.00074	0.00108	0.00025	0.00035	0.00007	0.00010
Overall																	
MUE	9.03	8.22	5.48	4.33	3.99	2.54	2.60	1.61	1.62	1.03	1.12	0.72	0.80	0.28	0.39	0.09	0.10
RMSD	11.99	10.93	7.22	5.77	5.35	3.29	3.36	2.24	2.09	1.60	1.71	0.95	1.06	0.37	0.50	0.12	0.14
MAX	40.51	37.81	23.04	21.79	18.49	10.50	11.42	10.46	5.51	8.89	7.71	3.28	3.48	1.26	1.93	0.41	0.43
MUE_{norm}	0.00686	0.00597	0.00347	0.00258	0.00214	0.00136	0.00139	0.00080	0.00079	0.00037	0.00040	0.00022	0.00025	0.00008	0.00011	0.00002	0.00003
$\text{RMSD}_{\text{norm}}$	0.00911	0.00793	0.00457	0.00344	0.00286	0.00176	0.00180	0.00112	0.00101	0.00057	0.00061	0.00029	0.00033	0.00010	0.00014	0.00003	0.00003
MAX_{norm}	0.03076	0.02746	0.01459	0.01298	0.00989	0.00562	0.00611	0.00522	0.00267	0.00317	0.00274	0.00100	0.00108	0.00034	0.00054	0.00010	0.00011

For the overall test set, CCSD-derived NOs yield lower or comparable RMSD and normalized RMSD values than their MP2-based counterparts for larger thresholds, whereas for tighter thresholds the trend is reversed. Projecting cc-pVTZ onto the cc-pVDZ space reduces the RMSD by approximately 2 kJ mol^{-1} on average and by up to 4 kJ mol^{-1} for atomization energies, relative to the uncompressed cc-pVDZ basis. Closed- and open-shell reaction energies exhibit similar or slightly smaller improvements. Considering the test set as a whole, the $5.0 \cdot 10^{-4}$ threshold, which corresponds to a slightly smaller n_{CO} than cc-pVDZ, produces a RMSD comparable to that obtained with cc-pVDZ when using the CCSD density. The $2.5 \cdot 10^{-4}$ threshold, associated with a slightly larger n_{CO}, provides consistently better results.

The statistical analysis of the (T) contribution, taking the DBH24 and W4-17 MR test sets as a reference, is summarized in Tables 6.3 and 6.4, respectively.

Table 6.3: MUE, RMSD, and MAX values, together with their normalized counterparts (*norm*, obtained by dividing by the total number of COs), for the (T) contribution to reaction barriers of the DBH24 test set according to different FNO compression schemes, relative to the cc-pVTZ reference. All values are in kJ mol^{-1}.

	$1.0{\cdot}10^{-3}$ cc-pVTZ		$5.0{\cdot}10^{-4}$ cc-pVTZ		cc-pVDZ	projected cc-pVDZ		$2.5{\cdot}10^{-4}$ cc-pVTZ		$1.0{\cdot}10^{-4}$ cc-pVTZ	
	MP2	CCSD	MP2	CCSD		MP2	CCSD	MP2	CCSD	MP2	CCSD
Forward											
MUE	2.56	2.17	2.44	1.84	1.41	1.20	1.32	1.00	0.76	0.79	0.49
RMSD	3.03	2.68	3.08	2.31	1.76	1.89	1.95	1.21	1.03	0.89	0.67
MAX	4.96	4.80	6.56	4.71	3.32	5.51	5.31	2.55	2.29	1.69	1.49
$\mathrm{MUE_{norm}}$	0.00305	0.00248	0.00231	0.00167	0.00112	0.00095	0.00104	0.00075	0.00055	0.00043	0.00026
$\mathrm{RMSD_{norm}}$	0.00362	0.00306	0.00291	0.00209	0.00139	0.00150	0.00154	0.00090	0.00074	0.00048	0.00036
$\mathrm{MAX_{norm}}$	0.00592	0.00547	0.00622	0.00425	0.00262	0.00435	0.00420	0.00190	0.00165	0.00092	0.00080
Reverse											
MUE	3.28	3.21	2.50	1.92	2.32	1.86	2.18	1.31	1.16	0.86	0.52
RMSD	3.70	4.59	3.04	2.23	2.79	2.46	2.85	1.51	1.67	0.97	0.73
MAX	7.55	11.71	4.79	3.50	5.38	6.54	6.47	2.52	3.65	1.80	1.83
$\mathrm{MUE_{norm}}$	0.00392	0.00365	0.00237	0.00173	0.00183	0.00147	0.00172	0.00098	0.00084	0.00047	0.00028
$\mathrm{RMSD_{norm}}$	0.00442	0.00523	0.00288	0.00201	0.00220	0.00194	0.00225	0.00113	0.00120	0.00052	0.00039
$\mathrm{MAX_{norm}}$	0.00902	0.01334	0.00454	0.00316	0.00425	0.00517	0.00511	0.00188	0.00263	0.00098	0.00099
Overall											
MUE	2.92	2.69	2.47	1.88	1.86	1.53	1.75	1.16	0.96	0.82	0.51
RMSD	3.38	3.76	3.06	2.27	2.33	2.19	2.44	1.37	1.39	0.93	0.70
MAX	7.55	11.71	6.56	4.71	5.38	6.54	6.47	2.55	3.65	1.80	1.83
$\mathrm{MUE_{norm}}$	0.00349	0.00306	0.00234	0.00170	0.00147	0.00121	0.00138	0.00086	0.00069	0.00045	0.00027
$\mathrm{RMSD_{norm}}$	0.00404	0.00428	0.00290	0.00205	0.00184	0.00173	0.00193	0.00102	0.00100	0.00050	0.00038
$\mathrm{MAX_{norm}}$	0.00902	0.01334	0.00622	0.00425	0.00425	0.00517	0.00511	0.00190	0.00263	0.00098	0.00099

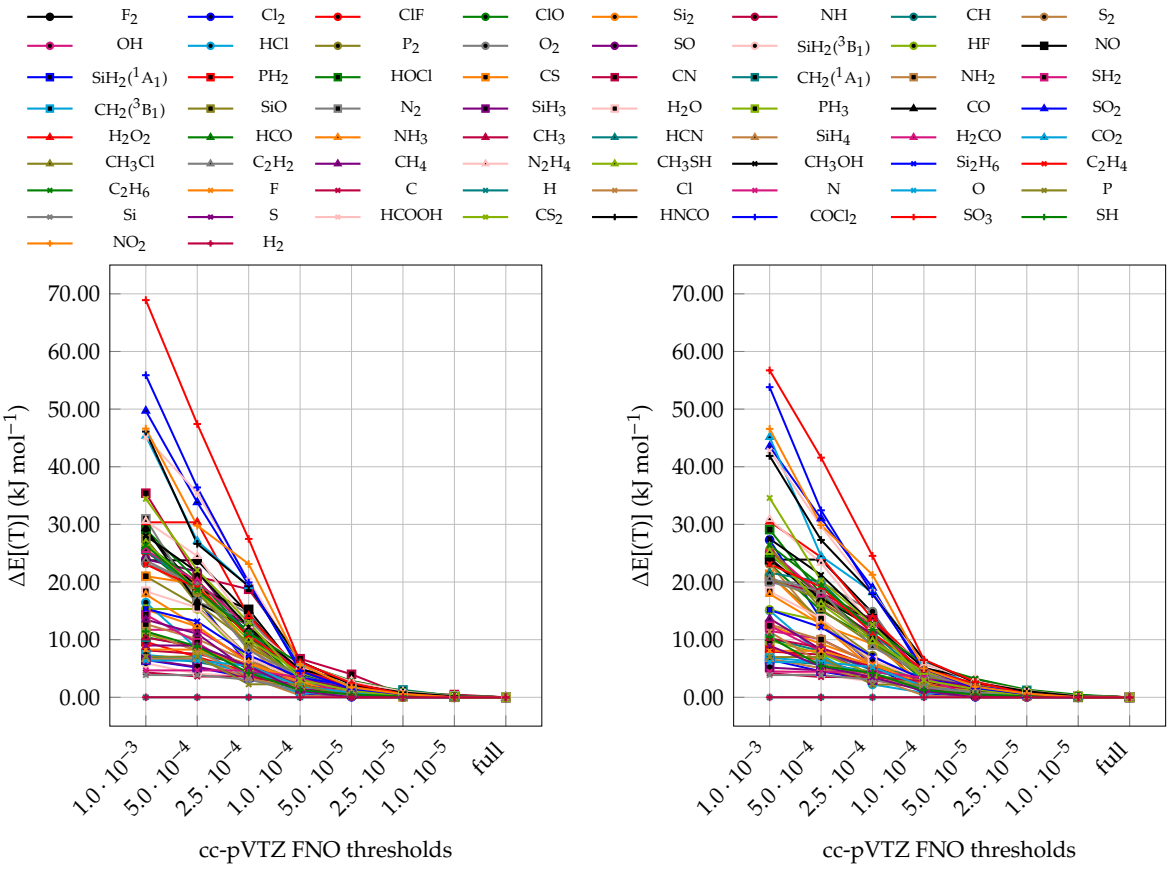
Table 6.4: MUE, RMSD, and MAX values, together with their normalized counterparts (*norm*, obtained by dividing by the total number of COs), for the (T) contribution to atomization energies of the W4-17 MR test set according to different FNO compression schemes, relative to the cc-pVTZ reference. All values are in kJ mol^{-1}.

	$1.0{\cdot}10^{-3}$ cc-pVTZ		cc-pVDZ	projected cc-pVDZ		$5.0{\cdot}10^{-4}$ cc-pVTZ		$2.5{\cdot}10^{-4}$ cc-pVTZ		$1.0{\cdot}10^{-4}$ cc-pVTZ	
	MP2	CCSD		MP2	CCSD	MP2	CCSD	MP2	CCSD	MP2	CCSD
MUE	19.92	18.26	16.78	11.46	11.80	12.13	11.65	6.45	5.74	1.32	1.42
RMSD	21.05	19.68	18.06	12.29	12.75	13.94	12.82	6.89	6.53	2.03	1.95
MAX	32.40	33.37	32.28	22.05	22.73	25.51	23.98	11.22	11.79	6.04	5.28
$\mathrm{MUE_{norm}}$	0.02770	0.02491	0.02082	0.01422	0.01464	0.01505	0.01446	0.00653	0.00567	0.00088	0.00091
$\mathrm{RMSD_{norm}}$	0.02928	0.02684	0.02240	0.01524	0.01582	0.01729	0.01591	0.00696	0.00644	0.00136	0.00125
$\mathrm{MAX_{norm}}$	0.04506	0.04553	0.04004	0.02736	0.02820	0.03165	0.02975	0.01134	0.01164	0.00403	0.00340

For the DBH24 test set, projection from larger basis sets yields RMSD values comparable to those obtained with the uncompressed cc-pVDZ reference. At the $5.0 \cdot 10^{-4}$ threshold, the overall RMSD is similar to that of the cc-pVDZ basis set when using the CCSD density. Reducing the threshold to $2.5 \cdot 10^{-4}$ further improves the RMSD, leading to better accuracy than that achieved with the cc-pVDZ basis set. The same trend observed in the previous test sets is confirmed here, with comparable n_{CO} values but improved accuracy relative to the uncompressed case.

For the W4-17 MR set, projecting cc-pVTZ onto the cc-pVDZ space reduces the RMSD by approximately 6 kJ mol^{-1} compared to the full cc-pVDZ basis. The $1.0 \cdot 10^{-3}$ threshold retains fewer COs than cc-pVDZ but leads to larger deviations, whereas the $5.0 \cdot 10^{-4}$ threshold, which includes slightly more COs, provides a RMSD improved by about 4 kJ mol^{-1} compared to the cc-pVDZ basis set. Interestingly, the projected cc-pVDZ basis performs on par with, or slightly better than, the $5.0 \cdot 10^{-4}$ threshold despite involving fewer COs, underscoring the robustness of the projection-based compression.

Figure 6.2 displays the absolute energy differences of the (T) contribution relative to the full basis set reference. Results are shown for the cc-pVTZ basis at various t_{NO} thresholds, using MP2 and CCSD densities for all species in the KAW test set.



(A) MP2 densities.

(B) CCSD densities.

FIGURE 6.2: Absolute electronic energy differences for the (T) contribution in the KAW dataset, computed with the cc-pVTZ basis set at different FNO thresholds with respect to the full cc-pVTZ results. Panels (A) and (B) correspond to calculations based on MP2 and CCSD densities, respectively.

A comparison of the two densities reveals that CCSD-based orbitals yield a smaller spread in the absolute electronic energy differences and converge more rapidly toward the uncompressed basis set result.

Moving to the full triple excitations, RMSDs and their normalized counterparts, computed with respect to the cc-pV(5,6)Z values of the Karton test set, are shown in Figure 6.3 (the full list of statistical indicators is reported in Section A.4 of Appendix A), while the statistical analyses carried out benchmarking against KAW, DBH24, and W4-17 MR test sets are detailed in Tables 6.5, 6.6, and 6.7, respectively. The observed trend closely mirrors that of the (T) contribution: compressed larger basis sets consistently outperform smaller uncompressed ones with comparable n_{CO}.

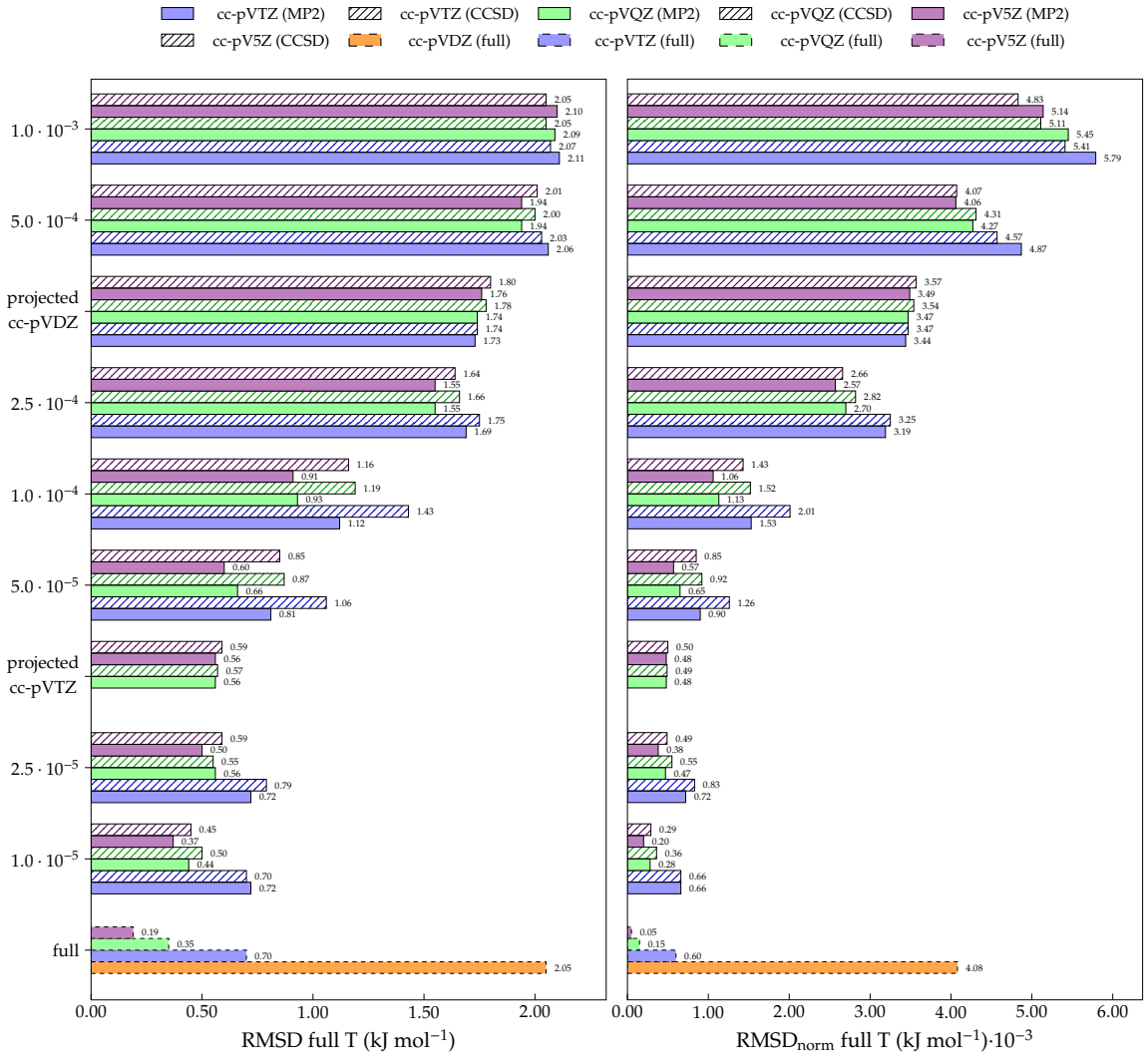


FIGURE 6.3: RMSD (left) and RMSD normalized to the number of COs (RMSD$_{\text{norm}}$, right) for the full T contribution to the atomization energies in the Karton test set according to different FNO compression schemes, referenced to cc-pV(5,6)Z values. Results obtained with orbitals selected from MP2-NO occupations are labeled as (MP2), those from CCSD-NO occupations as (CCSD), while calculations performed in the untruncated virtual space are denoted as (full). Projected cc-pVDZ and cc-pVTZ results correspond to the same number of COs as in the respective parent basis sets.

For the Karton test set, truncating all basis sets to match the size of cc-pVDZ reduces the RMSD by approximately 20% relative to cc-pVDZ itself, yielding errors below 1 kJ mol^{-1} for thresholds tighter than $5.0 \cdot 10^{-5}$. The normalized RMSD decreases almost linearly with the logarithm of the threshold and drops below 10^{-4} kJ mol^{-1} for $t_{\text{NO}} \leq 1.0 \cdot 10^{-4}$, confirming that the FNO approach enables systematic convergence even at higher excitation levels. As for perturbative triples, CCSD-derived NOs deliver slightly better accuracy at larger truncation thresholds, whereas MP2-based orbitals become competitive or marginally superior as the truncation is tightened, reflecting the increasing similarity between the two density matrices.

TABLE 6.5: MUE, RMSD, and MAX values, together with their normalized counterparts (*norm*, obtained by dividing by the total number of COs), for the full T contribution to atomization, closed-shell, and open-shell reaction energies of the KAW test set according to different FNO compression schemes, relative to the cc-pVTZ reference. All values are in kJ mol^{-1}.

	$1.0{\cdot}10^{-3}$ cc-pVTZ		$5.0{\cdot}10^{-4}$ cc-pVTZ		cc-pVDZ	projected cc-pVDZ		$2.5{\cdot}10^{-4}$ cc-pVTZ		$1.0{\cdot}10^{-4}$ cc-pVTZ		$5.0{\cdot}10^{-5}$ cc-pVTZ		$2.5{\cdot}10^{-5}$ cc-pVTZ		$1.0{\cdot}10^{-5}$ cc-pVTZ	
	MP2	CCSD	MP2	CCSD		MP2	CCSD	MP2	CCSD	MP2	CCSD	MP2	CCSD	MP2	CCSD	MP2	CCSD
Atomization energy																	
MUE	1.06	1.01	0.99	0.91	1.19	0.96	0.98	0.80	0.77	0.32	0.42	0.11	0.20	0.04	0.04	0.02	0.01
RMSD	1.32	1.27	1.22	1.13	1.43	1.11	1.13	0.97	0.92	0.41	0.53	0.23	0.29	0.08	0.06	0.04	0.02
MAX	3.01	2.91	2.72	2.85	3.58	2.55	2.69	1.83	1.65	1.16	1.23	1.38	0.80	0.40	0.21	0.18	0.05
MUE_{norm}	0.00080	0.00074	0.00063	0.00054	0.00064	0.00051	0.00052	0.00040	0.00037	0.00011	0.00015	0.00003	0.00006	0.00001	0.00001	0.00001	0.00000
$\text{RMSD}_{\text{norm}}$	0.00100	0.00092	0.00077	0.00068	0.00076	0.00059	0.00060	0.00049	0.00045	0.00015	0.00019	0.00007	0.00009	0.00002	0.00002	0.00001	0.00000
MAX_{norm}	0.00228	0.00211	0.00172	0.00170	0.00191	0.00136	0.00144	0.00091	0.00080	0.00041	0.00044	0.00042	0.00025	0.00011	0.00006	0.00004	0.00001
Closed-shell reaction energy																	
MUE	0.40	0.40	0.29	0.26	0.35	0.17	0.19	0.12	0.12	0.07	0.08	0.06	0.05	0.03	0.03	0.01	0.02
RMSD	0.50	0.52	0.39	0.34	0.44	0.23	0.28	0.15	0.16	0.08	0.09	0.08	0.07	0.05	0.05	0.02	0.03
MAX	1.37	1.63	1.09	1.01	1.29	0.66	0.97	0.34	0.45	0.17	0.25	0.20	0.18	0.13	0.14	0.05	0.07
MUE_{norm}	0.00031	0.00029	0.00019	0.00015	0.00019	0.00009	0.00010	0.00006	0.00006	0.00002	0.00003	0.00002	0.00002	0.00001	0.00001	0.00000	0.00001
$\text{RMSD}_{\text{norm}}$	0.00038	0.00038	0.00025	0.00020	0.00024	0.00012	0.00015	0.00008	0.00008	0.00003	0.00003	0.00002	0.00002	0.00001	0.00001	0.00000	0.00001
MAX_{norm}	0.00104	0.00119	0.00069	0.00060	0.00069	0.00035	0.00052	0.00017	0.00022	0.00006	0.00009	0.00006	0.00006	0.00003	0.00004	0.00001	0.00002
Open-shell reaction energy																	
MUE	0.86	0.83	0.76	0.63	0.77	0.60	0.57	0.50	0.39	0.18	0.20	0.06	0.10	0.02	0.02	0.02	0.01
RMSD	1.05	1.04	0.95	0.84	0.95	0.73	0.72	0.64	0.54	0.25	0.29	0.09	0.16	0.03	0.03	0.02	0.02
MAX	2.95	3.15	2.92	2.72	2.91	2.23	2.26	1.73	1.75	0.73	1.23	0.25	0.78	0.08	0.10	0.07	0.07
MUE_{norm}	0.00065	0.00061	0.00048	0.00037	0.00041	0.00032	0.00030	0.00025	0.00019	0.00006	0.00007	0.00002	0.00003	0.00001	0.00001	0.00000	0.00000
$\text{RMSD}_{\text{norm}}$	0.00080	0.00076	0.00060	0.00050	0.00051	0.00039	0.00038	0.00032	0.00026	0.00009	0.00010	0.00003	0.00005	0.00001	0.00001	0.00001	0.00001
MAX_{norm}	0.00224	0.00228	0.00185	0.00162	0.00156	0.00119	0.00121	0.00086	0.00085	0.00026	0.00044	0.00008	0.00024	0.00002	0.00003	0.00002	0.00002
Overall																	
MUE	0.83	0.81	0.75	0.66	0.84	0.64	0.64	0.53	0.48	0.21	0.26	0.08	0.13	0.03	0.03	0.02	0.02
RMSD	1.08	1.05	0.98	0.89	1.09	0.83	0.84	0.73	0.67	0.30	0.38	0.16	0.21	0.06	0.05	0.03	0.02
MAX	3.01	3.15	2.92	2.85	3.58	2.55	2.69	1.83	1.75	1.16	1.23	1.38	0.80	0.40	0.21	0.18	0.07
MUE_{norm}	0.00063	0.00059	0.00047	0.00039	0.00045	0.00034	0.00034	0.00027	0.00023	0.00007	0.00009	0.00002	0.00004	0.00001	0.00001	0.00000	0.00000
$\text{RMSD}_{\text{norm}}$	0.00082	0.00077	0.00062	0.00053	0.00058	0.00045	0.00045	0.00036	0.00033	0.00011	0.00014	0.00005	0.00007	0.00001	0.00001	0.00001	0.00001
MAX_{norm}	0.00228	0.00228	0.00185	0.00170	0.00191	0.00136	0.00144	0.00091	0.00085	0.00041	0.00044	0.00042	0.00025	0.00011	0.00006	0.00004	0.00002

For the KAW test set, the trends observed for full and perturbative triples are largely consistent with each other, differing mainly in two aspects. First, the gain from projecting cc-pVDZ over its uncompressed counterpart is smaller, amounting to only about 0.1 – 0.3 kJ mol^{-1} (in terms of RMSD) across all subsets of the test set. Second, by considering the whole test set, the $5.0 \cdot 10^{-4}$ threshold yields better RMSDs results than uncompressed cc-pVDZ, despite involving a smaller number of correlated orbitals. Notably, the normalized RMSD drops below 10^{-4} kJ mol^{-1} for thresholds tighter than $1.0 \cdot 10^{-4}$, confirming that the FNO approximation ensures systematic and controllable accuracy in post-CCSD(T) computations.

TABLE 6.6: MUE, RMSD, and MAX values, together with their normalized counterparts (*norm*, obtained by dividing by the total number of COs), for the full T contribution to reaction barriers of the DBH24 test set according to different FNO compression schemes, relative to the cc-pVTZ reference. All values are in kJ mol^{-1}.

	$1.0{\cdot}10^{-3}$ cc-pVTZ		$5.0{\cdot}10^{-4}$ cc-pVTZ		cc-pVDZ	projected cc-pVDZ		$2.5{\cdot}10^{-4}$ cc-pVTZ		$1.0{\cdot}10^{-4}$ cc-pVTZ	
	MP2	CCSD	MP2	CCSD		MP2	CCSD	MP2	CCSD	MP2	CCSD
Forward											
MUE	0.44	0.32	0.35	0.24	0.17	0.25	0.16	0.24	0.11	0.11	0.05
RMSD	0.57	0.39	0.44	0.29	0.20	0.31	0.22	0.28	0.15	0.13	0.06
MAX	1.10	0.73	0.93	0.47	0.42	0.66	0.57	0.47	0.26	0.22	0.10
MUE_{norm}	0.00053	0.00036	0.00034	0.00022	0.00013	0.00020	0.00013	0.00018	0.00008	0.00006	0.00002
$\text{RMSD}_{\text{norm}}$	0.00068	0.00045	0.00042	0.00026	0.00016	0.00025	0.00018	0.00021	0.00011	0.00007	0.00003
MAX_{norm}	0.00131	0.00083	0.00088	0.00043	0.00033	0.00052	0.00045	0.00035	0.00019	0.00012	0.00005
Reverse											
MUE	0.29	0.25	0.25	0.22	0.22	0.18	0.22	0.18	0.18	0.11	0.07
RMSD	0.36	0.32	0.30	0.28	0.34	0.21	0.24	0.22	0.21	0.13	0.08
MAX	0.89	0.72	0.72	0.51	1.00	0.41	0.42	0.42	0.41	0.26	0.14
MUE_{norm}	0.00035	0.00029	0.00024	0.00020	0.00018	0.00014	0.00017	0.00013	0.00013	0.00006	0.00004
$\text{RMSD}_{\text{norm}}$	0.00043	0.00036	0.00029	0.00025	0.00027	0.00017	0.00019	0.00017	0.00015	0.00007	0.00004
MAX_{norm}	0.00107	0.00082	0.00068	0.00046	0.00079	0.00032	0.00034	0.00031	0.00030	0.00014	0.00008
Overall											
MUE	0.37	0.29	0.30	0.23	0.20	0.22	0.19	0.21	0.14	0.11	0.06
RMSD	0.48	0.36	0.38	0.28	0.28	0.27	0.23	0.25	0.18	0.13	0.07
MAX	1.10	0.73	0.93	0.51	1.00	0.66	0.57	0.47	0.41	0.26	0.14
MUE_{norm}	0.00044	0.00033	0.00029	0.00021	0.00015	0.00017	0.00015	0.00015	0.00010	0.00006	0.00003
$\text{RMSD}_{\text{norm}}$	0.00057	0.00041	0.00036	0.00026	0.00022	0.00021	0.00018	0.00019	0.00013	0.00007	0.00004
MAX_{norm}	0.00131	0.00083	0.00088	0.00046	0.00079	0.00052	0.00045	0.00035	0.00030	0.00014	0.00008

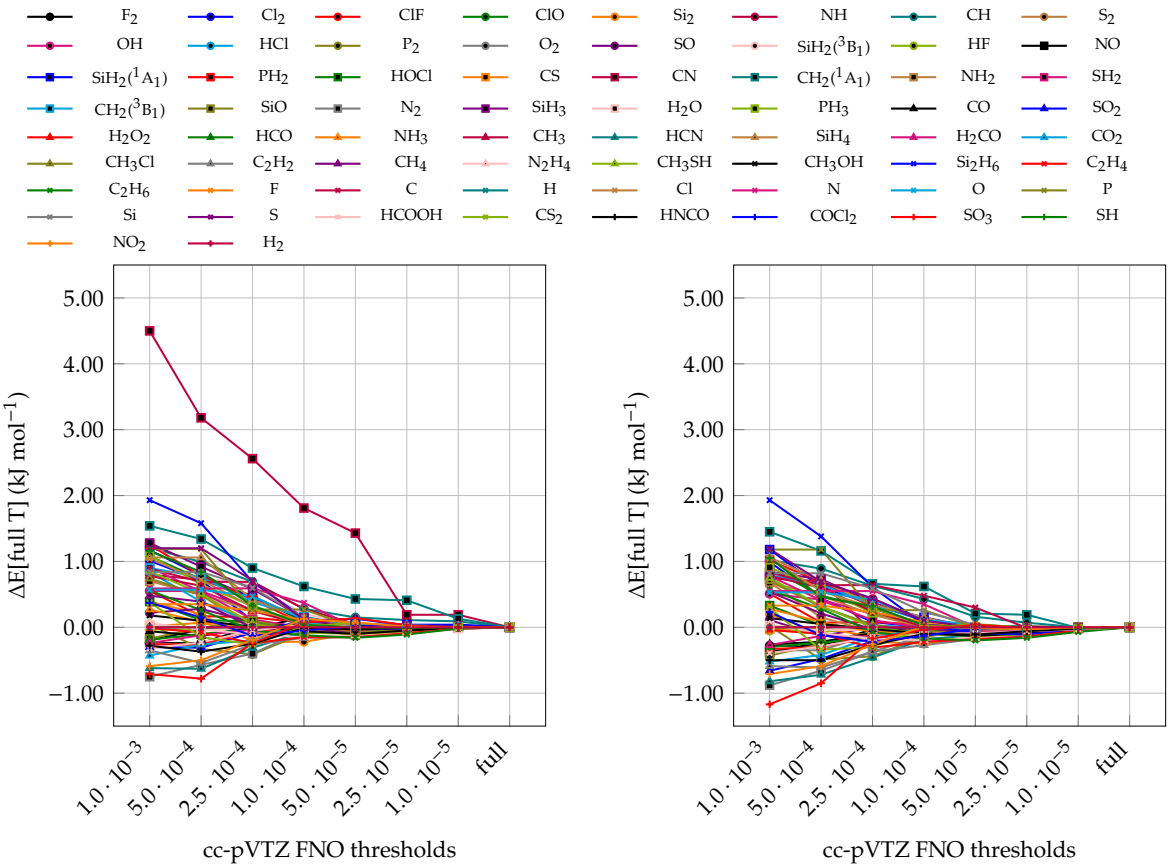
TABLE 6.7: MUE, RMSD, and MAX values, together with their normalized counterparts (*norm*, obtained by dividing by the total number of COs), for the full T contribution to atomization energies of the W4-17 MR test set according to different FNO compression schemes, relative to the cc-pVTZ reference. All values are in kJ mol^{-1}.

	$1.0{\cdot}10^{-3}$ cc-pVTZ		cc-pVDZ	projected cc-pVDZ		$5.0{\cdot}10^{-4}$ cc-pVTZ		$2.5{\cdot}10^{-4}$ cc-pVTZ		$1.0{\cdot}10^{-4}$ cc-pVTZ	
	MP2	CCSD		MP2	CCSD	MP2	CCSD	MP2	CCSD	MP2	CCSD
MUE	2.42	2.41	3.33	2.44	2.43	2.52	2.29	1.61	1.53	0.36	0.54
RMSD	2.78	2.64	3.64	2.68	2.65	2.82	2.43	1.89	1.66	0.54	0.80
MAX	5.90	5.09	7.75	5.19	5.30	5.44	3.68	4.37	2.89	1.59	2.14
MUE_{norm}	0.00337	0.00328	0.00413	0.00303	0.00302	0.00313	0.00285	0.00163	0.00151	0.00024	0.00035
$\text{RMSD}_{\text{norm}}$	0.00386	0.00361	0.00452	0.00332	0.00329	0.00350	0.00301	0.00191	0.00164	0.00036	0.00052
MAX_{norm}	0.00821	0.00694	0.00961	0.00644	0.00657	0.00675	0.00457	0.00442	0.00285	0.00106	0.00138

The same considerations made for the (T) contribution also apply to the DBH24 and W4-17 MR test sets. For W4-17 MR, two additional observations can be reported: first, projecting cc-pVTZ onto the cc-pVDZ space yields a gain of about 1 kJ mol^{-1} relative to the uncompressed

cc-pVDZ basis; second, even the use of the looser $1.0 \cdot 10^{-3}$ threshold provides a RMSD 1 kJ mol^{-1} lower than the uncompressed cc-pVDZ basis set.

Concerning the absolute energy differences of the full triple contribution, computed for the KAW test set with MP2 and CCSD densities at different t_{NO} thresholds relative to the full cc-pVTZ basis (Figure 6.4), the trend closely follows that observed for the (T) contribution, with the CCSD results exhibiting also a more systematic convergence toward the uncompressed basis set result.



(A) MP2 densities.

(B) CCSD densities.

FIGURE 6.4: Absolute electronic energy differences for the full T contribution in the KAW dataset, computed with the cc-pVTZ basis set at different FNO thresholds with respect to the full cc-pVTZ results. Panels (A) and (B) correspond to calculations based on MP2 and CCSD densities, respectively.

6.2.2 Quadruple and Higher Excitations

The effect of the different FNO truncation strategies on perturbative quadrupole excitations (Q) evaluated with reference to the Karton test set is summarized in Figure 6.5 in terms of RMSD and normalized RMSD with respect to cc-pV(5,6)Z reference data, while the complete statistical analysis can be found in Section A.4 of Appendix A. MUEs, RMSDs, and MAXs evaluated over the KAW, DBH24, and W4-17 MR benchmark sets are reported in Tables 6.8,

6.9, and 6.10, respectively, together with the corresponding values normalized to the number of COs.

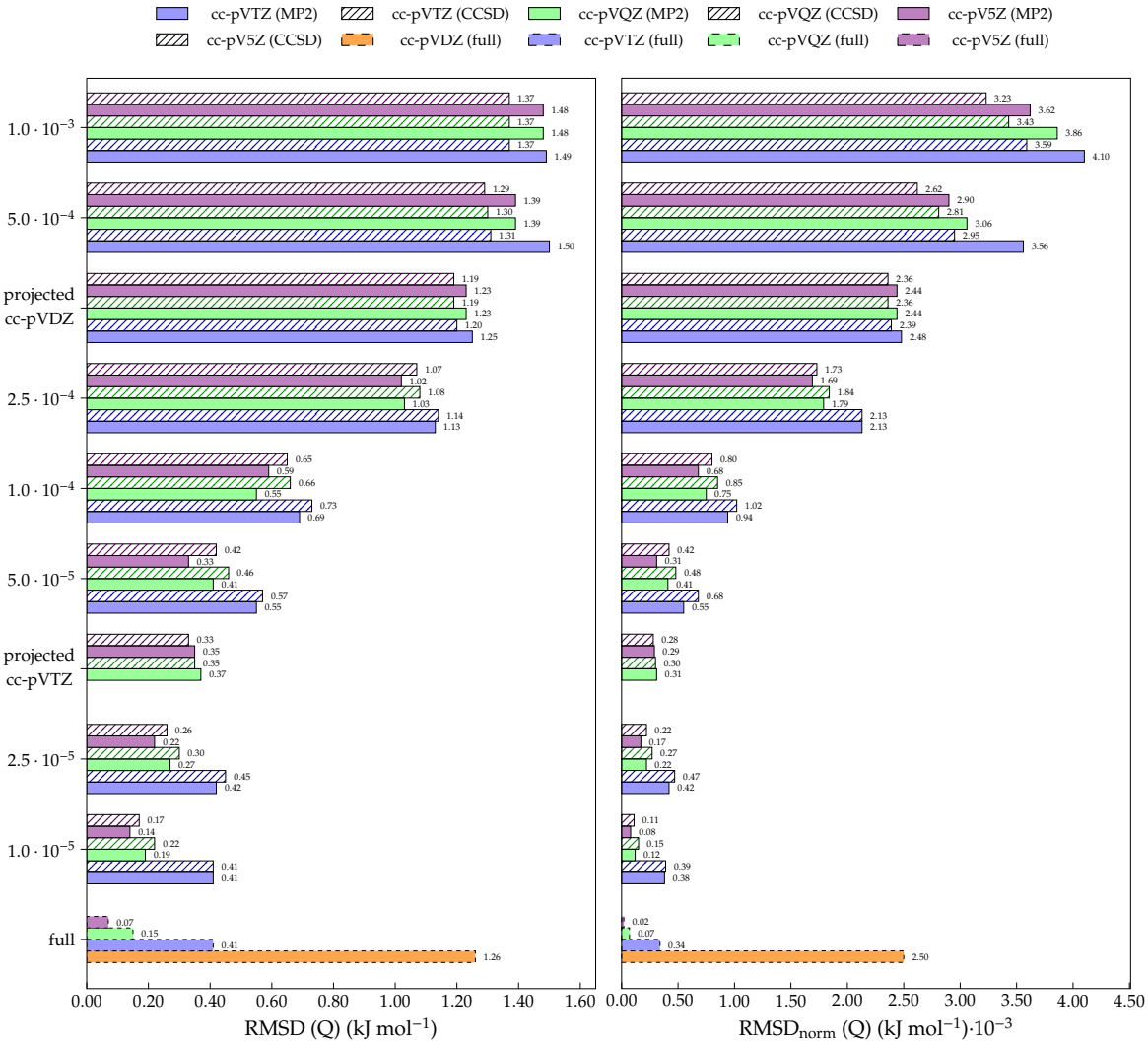


FIGURE 6.5: RMSD (left) and RMSD normalized to the number of COs ($RMSD_{norm}$, right) for the (Q) contribution to the atomization energies in the Karton test set according to different FNO compression schemes, referenced to cc-pV(5,6)Z values. Results obtained with orbitals selected from MP2-NO occupations are labeled as (MP2), those from CCSD-NO occupations as (CCSD), while calculations performed in the untruncated virtual space are denoted as (full). Projected cc-pVDZ and cc-pVTZ results correspond to the same number of COs as in the respective parent basis sets.

TABLE 6.8: MUE, RMSD, and MAX values, together with their normalized counterparts (*norm*, obtained by dividing by the total number of COs), for the (Q) contribution to atomization, closed-shell, and open-shell reaction energies of the KAW test set according to different FNO compression schemes, relative to the cc-pVTZ reference. All values are in kJ mol^{-1}.

	$1.0{\cdot}10^{-3}$ cc-pVTZ		$5.0{\cdot}10^{-4}$ cc-pVTZ		cc-pVDZ	projected cc-pVDZ		$2.5{\cdot}10^{-4}$ cc-pVTZ		$1.0{\cdot}10^{-4}$ cc-pVTZ		$5.0{\cdot}10^{-5}$ cc-pVTZ		$2.5{\cdot}10^{-5}$ cc-pVTZ		$1.0{\cdot}10^{-5}$ cc-pVTZ	
	MP2	CCSD	MP2	CCSD		MP2	CCSD	MP2	CCSD	MP2	CCSD	MP2	CCSD	MP2	CCSD	MP2	CCSD
Atomization energy																	
MUE	0.42	0.37	0.43	0.29	0.29	0.43	0.35	0.39	0.31	0.22	0.19	0.11	0.09	0.03	0.03	0.01	0.01
RMSD	0.65	0.53	0.64	0.41	0.49	0.63	0.51	0.59	0.43	0.32	0.23	0.18	0.11	0.04	0.04	0.02	0.01
MAX	2.53	1.61	2.52	1.29	1.69	2.22	1.50	2.46	1.17	1.47	0.58	1.04	0.24	0.13	0.08	0.14	0.03
MUE_{norm}	0.00032	0.00027	0.00027	0.00017	0.00015	0.00023	0.00019	0.00019	0.00015	0.00008	0.00007	0.00003	0.00003	0.00001	0.00001	0.00000	0.00000
$\text{RMSD}_{\text{norm}}$	0.00049	0.00039	0.00041	0.00024	0.00026	0.00034	0.00027	0.00029	0.00021	0.00011	0.00008	0.00006	0.00003	0.00001	0.00001	0.00001	0.00000
MAX_{norm}	0.00192	0.00117	0.00159	0.00077	0.00090	0.00118	0.00080	0.00123	0.00057	0.00052	0.00021	0.00032	0.00007	0.00004	0.00002	0.00003	0.00001
Closed-shell reaction energy																	
MUE	0.64	0.59	0.56	0.43	0.41	0.43	0.37	0.21	0.19	0.09	0.07	0.03	0.03	0.02	0.02	0.01	0.01
RMSD	0.93	0.85	0.80	0.58	0.56	0.55	0.45	0.27	0.26	0.12	0.09	0.04	0.06	0.03	0.03	0.01	0.01
MAX	2.79	2.54	2.82	1.49	1.70	1.28	1.12	0.57	0.58	0.27	0.24	0.12	0.22	0.10	0.09	0.02	0.02
MUE_{norm}	0.00049	0.00043	0.00036	0.00026	0.00022	0.00023	0.00020	0.00010	0.00009	0.00003	0.00002	0.00001	0.00001	0.00001	0.00000	0.00000	0.00000
$\text{RMSD}_{\text{norm}}$	0.00071	0.00062	0.00050	0.00035	0.00030	0.00029	0.00024	0.00014	0.00012	0.00004	0.00003	0.00001	0.00002	0.00001	0.00001	0.00000	0.00000
MAX_{norm}	0.00212	0.00184	0.00179	0.00088	0.00091	0.00068	0.00060	0.00028	0.00028	0.00010	0.00009	0.00004	0.00007	0.00003	0.00002	0.00000	0.00000
Open-shell reaction energy																	
MUE	0.92	0.78	0.70	0.40	0.43	0.48	0.37	0.28	0.25	0.17	0.15	0.07	0.08	0.03	0.02	0.01	0.01
RMSD	1.48	1.21	1.07	0.54	0.63	0.67	0.49	0.36	0.32	0.25	0.20	0.10	0.11	0.04	0.03	0.01	0.01
MAX	5.62	4.50	3.93	1.24	2.60	2.03	1.16	0.83	0.86	0.95	0.63	0.35	0.32	0.17	0.08	0.02	0.03
MUE_{norm}	0.00070	0.00057	0.00045	0.00024	0.00023	0.00025	0.00020	0.00014	0.00012	0.00006	0.00005	0.00002	0.00003	0.00001	0.00001	0.00000	0.00000
$\text{RMSD}_{\text{norm}}$	0.00112	0.00088	0.00068	0.00032	0.00033	0.00036	0.00026	0.00018	0.00016	0.00009	0.00007	0.00003	0.00003	0.00001	0.00001	0.00000	0.00000
MAX_{norm}	0.00427	0.00327	0.00249	0.00074	0.00139	0.00109	0.00062	0.00042	0.00042	0.00034	0.00022	0.00011	0.00010	0.00004	0.00002	0.00001	0.00001
Overall																	
MUE	0.66	0.58	0.57	0.37	0.37	0.45	0.36	0.31	0.26	0.17	0.15	0.08	0.07	0.03	0.02	0.01	0.01
RMSD	1.09	0.91	0.86	0.50	0.56	0.63	0.49	0.45	0.36	0.26	0.19	0.13	0.10	0.04	0.03	0.01	0.01
MAX	5.62	4.50	3.93	1.49	2.60	2.22	1.50	2.46	1.17	1.47	0.63	1.04	0.32	0.17	0.09	0.14	0.03
MUE_{norm}	0.00050	0.00042	0.00036	0.00022	0.00020	0.00024	0.00019	0.00015	0.00013	0.00006	0.00005	0.00002	0.00002	0.00001	0.00001	0.00000	0.00000
$\text{RMSD}_{\text{norm}}$	0.00083	0.00066	0.00054	0.00030	0.00030	0.00034	0.00026	0.00022	0.00017	0.00009	0.00007	0.00004	0.00003	0.00001	0.00001	0.00000	0.00000
MAX_{norm}	0.00427	0.00327	0.00249	0.00088	0.00139	0.00118	0.00080	0.00123	0.00057	0.00052	0.00022	0.00032	0.00010	0.00004	0.00002	0.00003	0.00001

TABLE 6.9: MUE, RMSD, and MAX values, together with their normalized counterparts (*norm*, obtained by dividing by the total number of COs), for the (Q) contribution to reaction barriers of the DBH24 test set according to different FNO compression schemes, relative to the cc-pVTZ reference. All values are in kJ mol^{-1}.

	$1.0{\cdot}10^{-3}$ cc-pVTZ		$5.0{\cdot}10^{-4}$ cc-pVTZ		cc-pVDZ	projected cc-pVDZ		$2.5{\cdot}10^{-4}$ cc-pVTZ		$1.0{\cdot}10^{-4}$ cc-pVTZ	
	MP2	CCSD	MP2	CCSD		MP2	CCSD	MP2	CCSD	MP2	CCSD
Forward											
MUE	0.43	0.33	0.35	0.24	0.23	0.29	0.18	0.25	0.14	0.12	0.07
RMSD	0.51	0.39	0.40	0.30	0.30	0.36	0.25	0.32	0.20	0.15	0.09
MAX	0.96	0.74	0.74	0.59	0.76	0.78	0.65	0.68	0.42	0.27	0.20
MUE_{norm}	0.00052	0.00037	0.00033	0.00022	0.00018	0.00023	0.00014	0.00018	0.00010	0.00007	0.00004
$\text{RMSD}_{\text{norm}}$	0.00060	0.00045	0.00038	0.00027	0.00024	0.00028	0.00020	0.00024	0.00015	0.00008	0.00005
MAX_{norm}	0.00115	0.00084	0.00070	0.00054	0.00060	0.00062	0.00052	0.00051	0.00030	0.00015	0.00011
Reverse											
MUE	0.59	0.46	0.48	0.38	0.32	0.40	0.29	0.32	0.21	0.15	0.09
RMSD	0.75	0.58	0.61	0.47	0.43	0.52	0.38	0.43	0.27	0.19	0.11
MAX	1.67	1.16	1.34	0.85	1.14	1.11	0.73	0.97	0.51	0.47	0.20
MUE_{norm}	0.00070	0.00053	0.00046	0.00035	0.00025	0.00032	0.00023	0.00024	0.00015	0.00008	0.00005
$\text{RMSD}_{\text{norm}}$	0.00089	0.00066	0.00057	0.00042	0.00034	0.00041	0.00030	0.00032	0.00020	0.00011	0.00006
MAX_{norm}	0.00200	0.00132	0.00127	0.00077	0.00090	0.00088	0.00058	0.00073	0.00037	0.00026	0.00011
Overall											
MUE	0.51	0.40	0.41	0.31	0.28	0.34	0.23	0.28	0.18	0.14	0.08
RMSD	0.64	0.50	0.51	0.39	0.37	0.45	0.32	0.38	0.24	0.17	0.10
MAX	1.67	1.16	1.34	0.85	1.14	1.11	0.73	0.97	0.51	0.47	0.20
MUE_{norm}	0.00061	0.00045	0.00039	0.00028	0.00022	0.00027	0.00018	0.00021	0.00013	0.00007	0.00004
$\text{RMSD}_{\text{norm}}$	0.00076	0.00056	0.00048	0.00035	0.00029	0.00035	0.00026	0.00028	0.00017	0.00009	0.00006
MAX_{norm}	0.00200	0.00132	0.00127	0.00077	0.00090	0.00088	0.00058	0.00073	0.00037	0.00026	0.00011

TABLE 6.10: MUE, RMSD, and MAX values, together with their normalized counterparts (*norm*, obtained by dividing by the total number of COs), for the (Q) contribution to atomization energies of the W4-17 MR test set according to different FNO compression schemes, relative to the cc-pVTZ reference. All values are in kJ mol^{-1}.

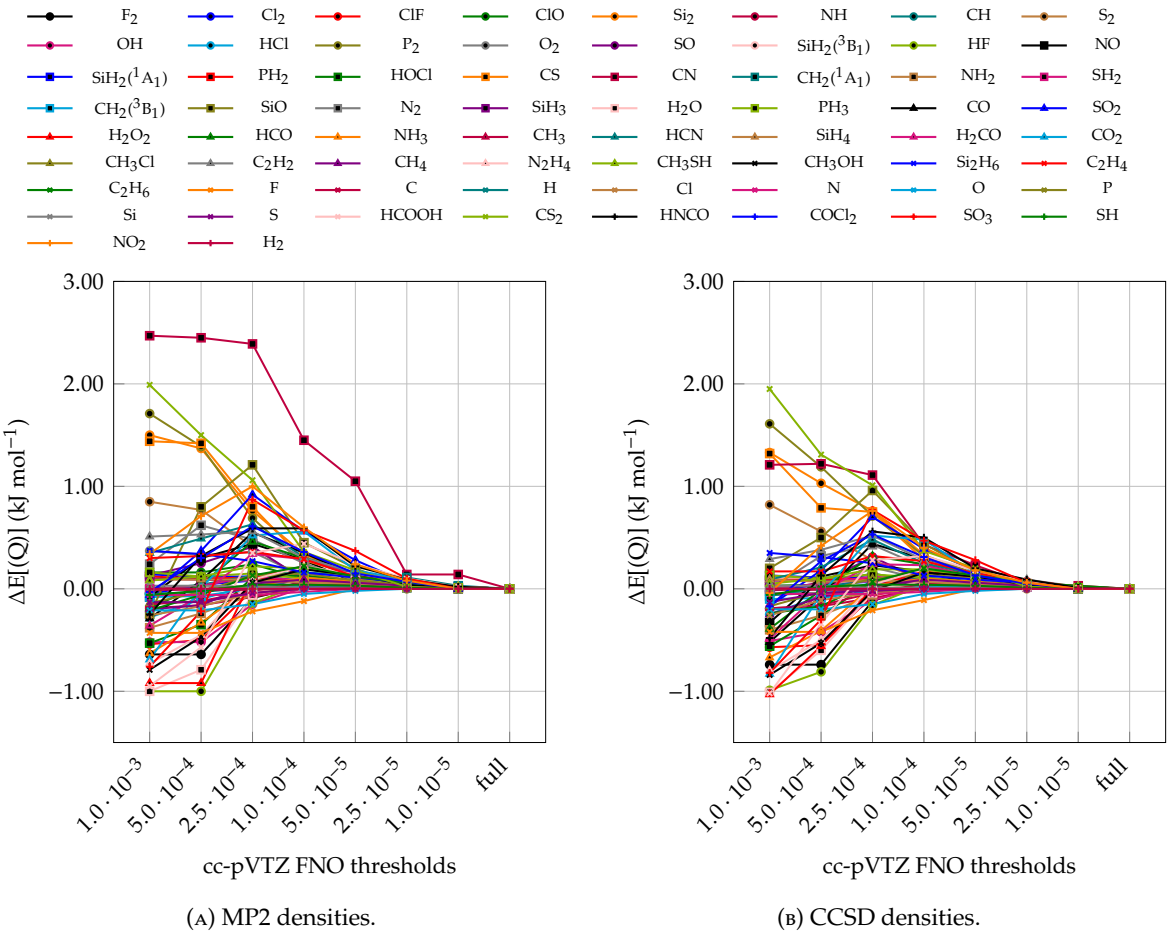
	$1.0{\cdot}10^{-3}$ cc-pVTZ		cc-pVDZ	projected cc-pVDZ		$5.0{\cdot}10^{-4}$ cc-pVTZ		$2.5{\cdot}10^{-4}$ cc-pVTZ		$1.0{\cdot}10^{-4}$ cc-pVTZ	
	MP2	CCSD		MP2	CCSD	MP2	CCSD	MP2	CCSD	MP2	CCSD
MUE	2.30	1.85	1.34	1.93	1.69	1.98	1.50	1.69	1.42	0.70	0.62
RMSD	2.88	2.36	1.92	2.34	2.13	2.45	1.93	2.02	1.68	0.78	0.73
MAX	6.06	5.40	5.54	5.27	5.46	5.72	4.39	5.34	4.08	1.57	1.86
MUE_{norm}	0.00320	0.00252	0.00166	0.00240	0.00209	0.00246	0.00186	0.00171	0.00140	0.00047	0.00040
$\text{RMSD}_{\text{norm}}$	0.00400	0.00322	0.00239	0.00290	0.00265	0.00304	0.00239	0.00204	0.00165	0.00052	0.00047
MAX_{norm}	0.00843	0.00736	0.00688	0.00654	0.00678	0.00710	0.00545	0.00540	0.00402	0.00105	0.00120

In general, similarly to the case of full triples, CCSD-based NOs generally yield slightly lower errors than those derived from MP2 for moderate thresholds (down to $t_{\text{NO}} = 2.5 \cdot 10^{-4}$), although this advantage becomes negligible as the threshold is further tightened. Furthermore, the benefit of large-basis compression is again apparent. For instance, in the Karton test set (Figure 6.5), compressing the cc-pV5Z basis at $t_{\text{NO}} = 2.5{\cdot}10^{-4}$ results in smaller RMSDs than those obtained with the full cc-pVDZ basis, even though the two bases involve a comparable n_{CO}. Normalized RMSDs for (Q) contributions fall below 10^{-4} kJ mol^{-1} when $t_{\text{NO}} \leq 5.0 \cdot 10^{-5}$, confirming that systematic and controllable accuracy can also be achieved for quadruple excitations.

For closed- and open-shell reactions in the KAW test set, the improvement upon tightening t_{NO} from $5.0 \cdot 10^{-4}$ to $1.0 \cdot 10^{-4}$ is particularly pronounced, with RMSDs reduced by more than a factor of two (Table 6.8). A comparable trend is observed for the reaction energies in the DBH24 test set (Table 6.9).

The W4-17 MR subset represents the most demanding benchmark for higher excitations. As it can be observed in Table 6.10, in this case, RMSDs for (Q) terms are between 2 and 3 kJ mol^{-1} for projected cc-pVDZ, and they remain in the range 1.68 - 2.88 kJ mol^{-1} for $t_{NO} \geq 2.5 \cdot 10^{-4}$, dropping to about 0.7 kJ mol^{-1} at $t_{NO} = 1.0 \cdot 10^{-4}$. These results demonstrate that the FNO-based truncation scheme maintains robust error control even for strongly correlated species, provided that sufficiently tight thresholds are employed.

The absolute energy differences of the (Q) contribution relative to the full cc-pVTZ basis, obtained for different t_{NO} thresholds using MP2 and CCSD densities for all species in the KAW test set are presented in Figure 6.6. As in the case of perturbative and full triples, CCSD densities generally lead to a narrower spread and faster convergence toward the uncompressed reference.



(A) MP2 densities.

(B) CCSD densities.

FIGURE 6.6: Absolute electronic energy differences for the (Q) contribution in the KAW dataset, computed with the cc-pVTZ basis set at different FNO thresholds with respect to the full cc-pVTZ results. Panels (A) and (B) correspond to calculations based on MP2 and CCSD densities, respectively.

The results obtained for the perturbative inclusion of quadrupole excitations through the $(Q)_\Lambda$ model with reference to the Karton test set are reported in Figure 6.7, while those for

the (Q)/A and (Q)/B approximations can be found in Figures A.3 and A.4, respectively, of Appendix A.

In general, the different approximations show a similar behavior, and referring to the full quadruple excitations extrapolated to the CBS limit by using cc-pV5Z and cc-pV6Z basis sets, the following conclusions can be drawn. (i) In contrast with the (Q) model (reported in Figure A.2 of Appendix A), the $(Q)_\Lambda$ approximation performs better as the threshold becomes smaller. (ii) This same trend is also observed for (Q)/A and (Q)/B. (iii) (Q)/A and (Q)/B approximations tend to perform slightly better at large thresholds, whereas $(Q)_\Lambda$ becomes preferable as the threshold is tightened.

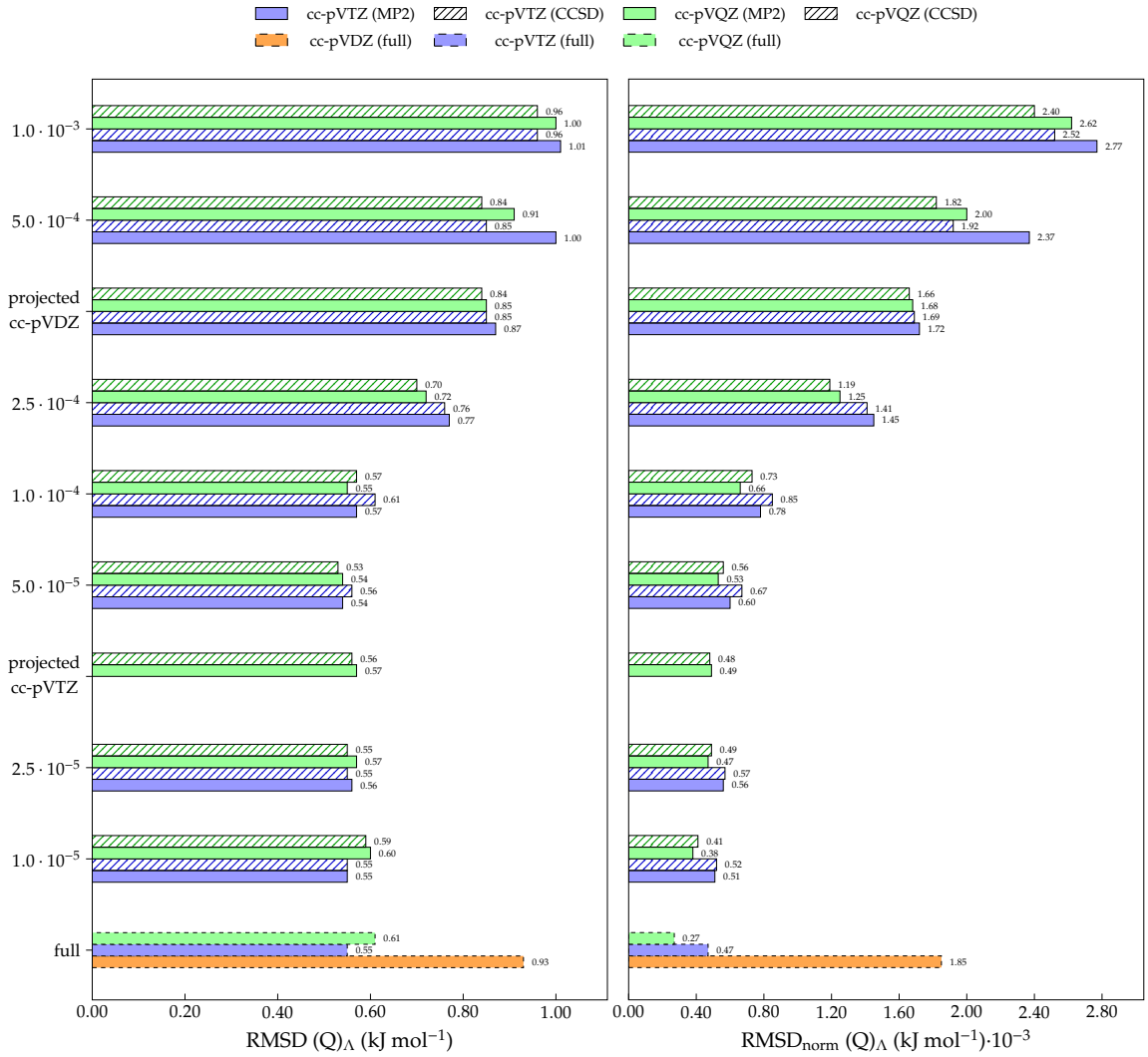


FIGURE 6.7: RMSD (left) and RMSD normalized to the number of COs ($RMSD_{norm}$, right) for the $(Q)_\Lambda$ contribution to the atomization energies in the Karton test set according to different FNO compression schemes, referenced to full Q cc-pV(5,6)Z values. Results obtained with orbitals selected from MP2-NO occupations are labeled as (MP2), those from CCSD-NO occupations as (CCSD), while calculations performed in the untruncated virtual space are denoted as (full). Projected cc-pVDZ and cc-pVTZ results correspond to the same number of COs as in the respective parent basis sets.

Concluding with the analysis of the impact of FNO truncation schemes on quadrupole excitations, the results obtained using the full CCSDTQ treatment are summarized in Figure 6.8, where RMSD and CO-normalized RMSD with respect to the Karton test set are reported. As can be seen, compared with the perturbative (Q) and its variants, the convergence behavior is significantly improved, while the general trends observed for the full-triple contribution are largely preserved.

The results obtained at the CCSDTQ level for the Karton test set are summarized in Figure 6.8.

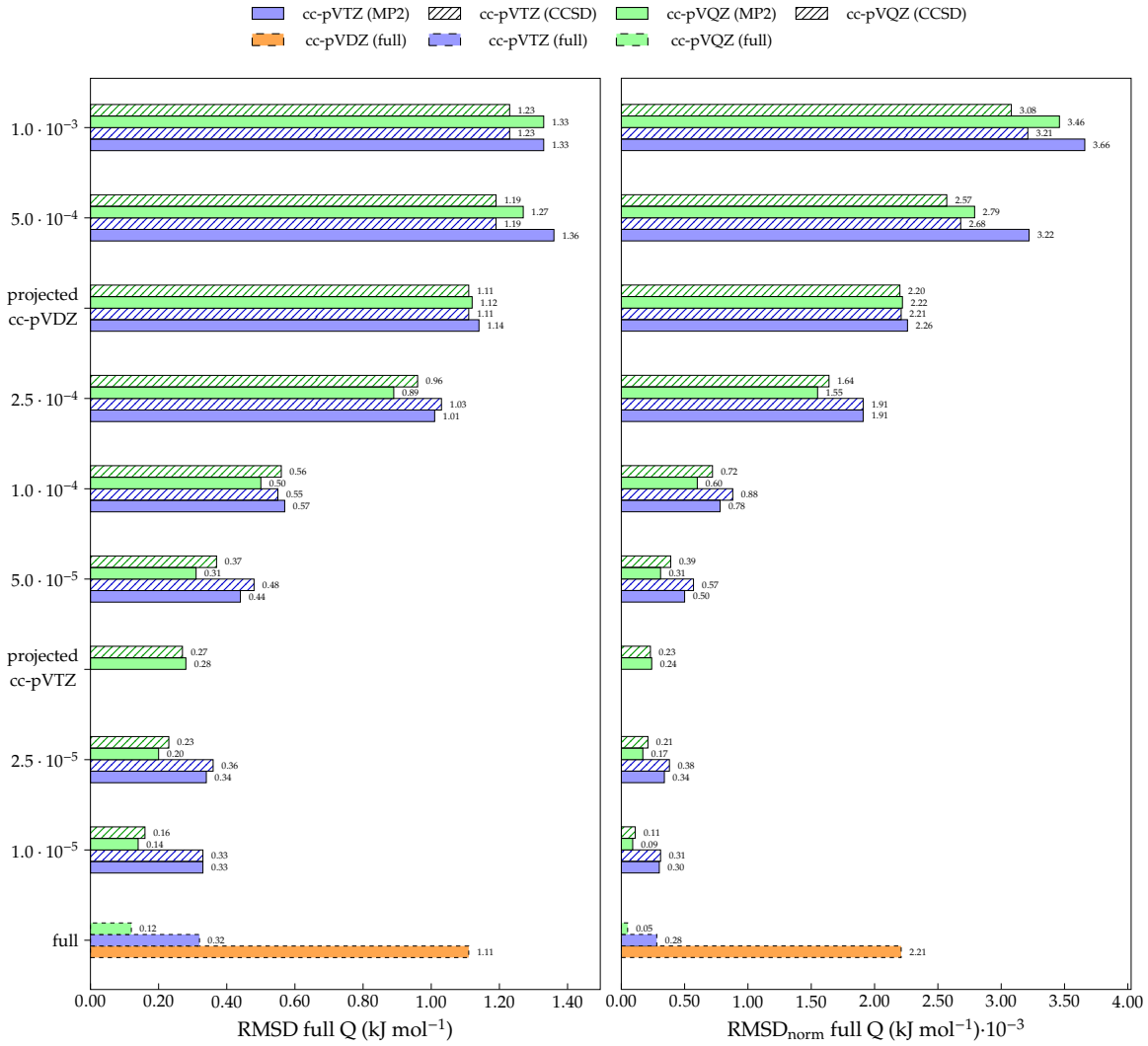


FIGURE 6.8: RMSD (left) and RMSD normalized to the number of COs ($RMSD_{norm}$, right) for the full Q contribution to the atomization energies in the Karton test set according to different FNO compression schemes, referenced to cc-pV(5,6)Z values. Results obtained with orbitals selected from MP2-NO occupations are labeled as (MP2), those from CCSD-NO occupations as (CCSD), while calculations performed in the untruncated virtual space are denoted as (full). Projected cc-pVDZ and cc-pVTZ results correspond to the same number of COs as in the respective parent basis sets.

Finally, the behavior of perturbative pentuple excitations coupled with FNOs was also investigated, even though their contribution is generally small and, actually, it is already well described using the cc-pVDZ basis set, with a RMSD of 0.09 kJ mol^{-1}. RMSDs and their corresponding values normalized by the number of COs with respect to the Karton benchmark set are shown in Figure 6.9. Since CBS values of the pentuple excitation contribution are not available for this dataset, the assessment was carried out with respect to the complete cc-pVTZ and cc-pVQZ basis sets.

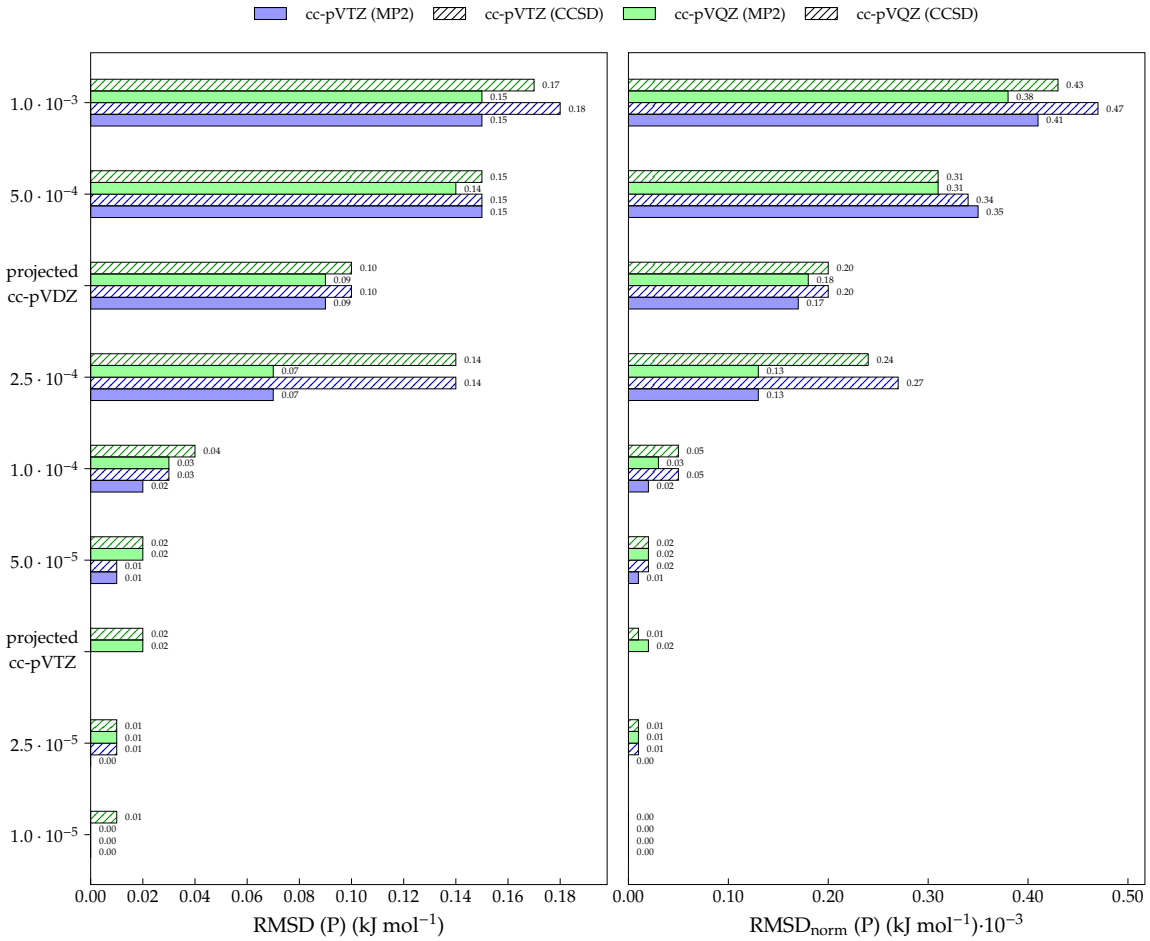


FIGURE 6.9: RMSD (left) and RMSD normalized to the number of COs ($RMSD_{norm}$, right) for the (P) contribution to the atomization energies in the Karton test set according to different FNO compression schemes, referenced to full basis set values. Results obtained with orbitals selected from MP2-NO occupations are labeled as (MP2) and those from CCSD-NO occupations as (CCSD). Projected cc-pVDZ and cc-pVTZ results correspond to the same number of COs as in the respective parent basis sets.

Overall, the benchmark results consistently indicate that, for a fixed n_{CO}, retaining the occupied space of a larger parent basis set provides higher accuracy than using smaller basis sets. This key finding supports the rationale behind the single basis set, progressively compressed strategy.

6.2.3 Trade-Off Between Accuracy and Computational Cost

Cost-effective strategies for estimating post-CCSD(T) contributions from relatively small basis sets were widely explored in the literature, with notable contributions from Karton and co-workers.[78,80] These studies demonstrated that basis set limit estimates for the perturbative connected quadruples (Q), can be obtained by extrapolating results from extended cc-pVDZ (4s,3p,1d) and cc-pVTZ basis sets. When combined with fully iterative connected quadruples,

Q–(Q), computed at the cc-pVDZ level, this approach provides a computationally efficient route to accurate Q contributions. Similarly, connected quintuple (P), can be reasonably approximated using cc-pVDZ (3s,2p) basis sets. Although these methods are efficient, they require switching between basis sets and, in some cases, adapting the extrapolation procedure for different excitation levels, which complicates the implementation of fully automated workflows.

An alternative strategy is to employ a single, sufficiently large parent basis set for all excitation classes, while progressively reducing the number of COs for higher excitations by increasing the occupation threshold t_{NO} applied to the natural orbitals in the virtual space. Larger thresholds eliminate more virtual orbitals, lowering the computational cost but slightly reducing accuracy, whereas smaller thresholds retain a larger fraction of the virtual space, enhancing accuracy at a moderate increase in cost. The occupied orbital space, fixed by the electron number, which is well described in a large basis set, is fully preserved, while the virtual manifold can be selectively pruned to retain only the most relevant correlation effects. This approach provides a smooth and tunable balance between accuracy and efficiency, laying the groundwork for the development of automated and unsupervised computational protocols.

By considering the outcomes obtained for full T and (Q) contributions evaluated over the Karton and KAW benchmark sets, Figures A.1 and A.5 and 6.10 – 6.13 illustrate that the occupation-based approach allows for n_{CO} values that smoothly interpolate between those associated with conventional basis sets of subsequent cardinality. For instance, starting from cc-pVQZ or cc-pV5Z and gradually increasing the threshold reduces n_{CO} toward the values characteristic of smaller bases. When the number of COs slightly exceeds that of cc-pVDZ, the resulting (Q) RMSDs are already comparable to those obtained with cc-pVTZ, while an n_{CO} marginally above cc-pVTZ leads to accuracy approaching that of cc-pVQZ.

In passing, it is worth noticing that the same conclusion can be drawn for the perturbative triple contribution, whose normalized RMSDs are reported in Figures A.1 and A.5 of Appendix A.

The obtained trend is particularly appealing as this ability to interpolate continuously between basis set cardinal numbers is not achievable within conventional composite frameworks or when using canonical orbitals.

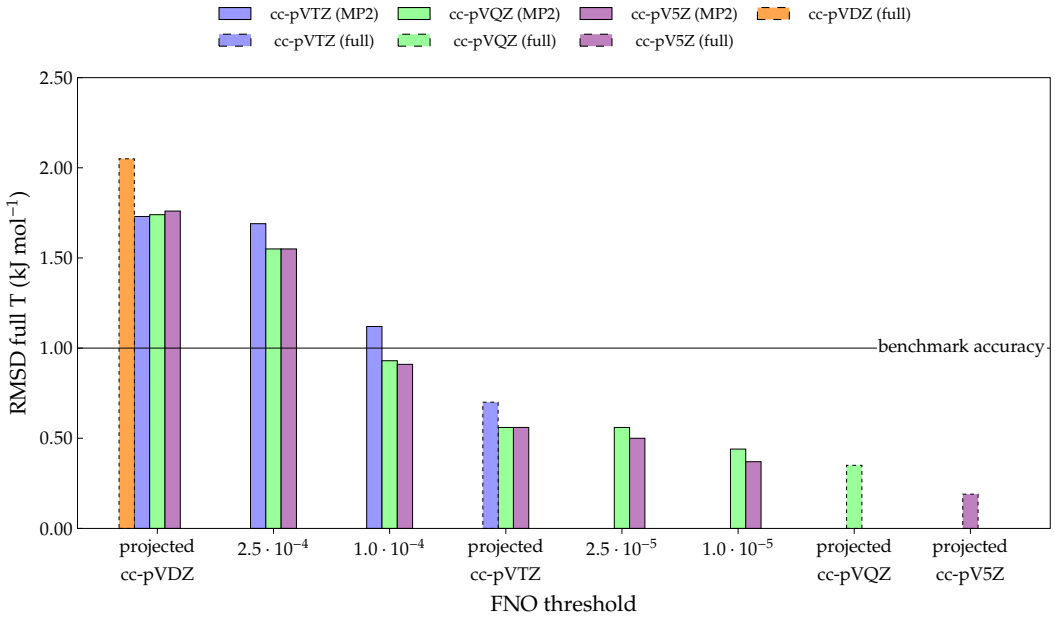


FIGURE 6.10: RMSD for the full T contribution to the atomization energies of the Karton test set, relative to the cc-pV(5,6)Z reference, as a function of the FNO truncation threshold. Results based on MP2-derived orbitals are labeled as (MP2) and calculations with the full virtual space as (full). Projected cc-pVnZ thresholds retain the same number of COs as the corresponding cc-pVnZ basis sets.

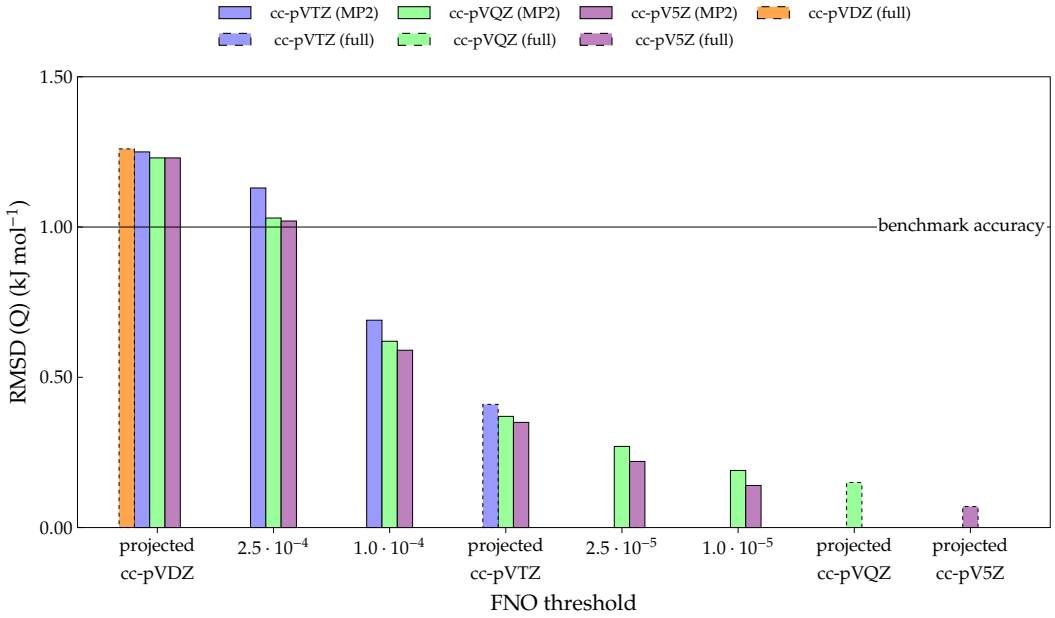


FIGURE 6.11: RMSD for the (Q) contribution to the atomization energies of the Karton test set, relative to the cc-pV(5,6)Z reference, as a function of the FNO truncation threshold. Results based on MP2-derived orbitals are labeled as (MP2) and calculations with the full virtual space as (full). Projected cc-pVnZ thresholds retain the same number of COs as the corresponding cc-pVnZ basis sets.

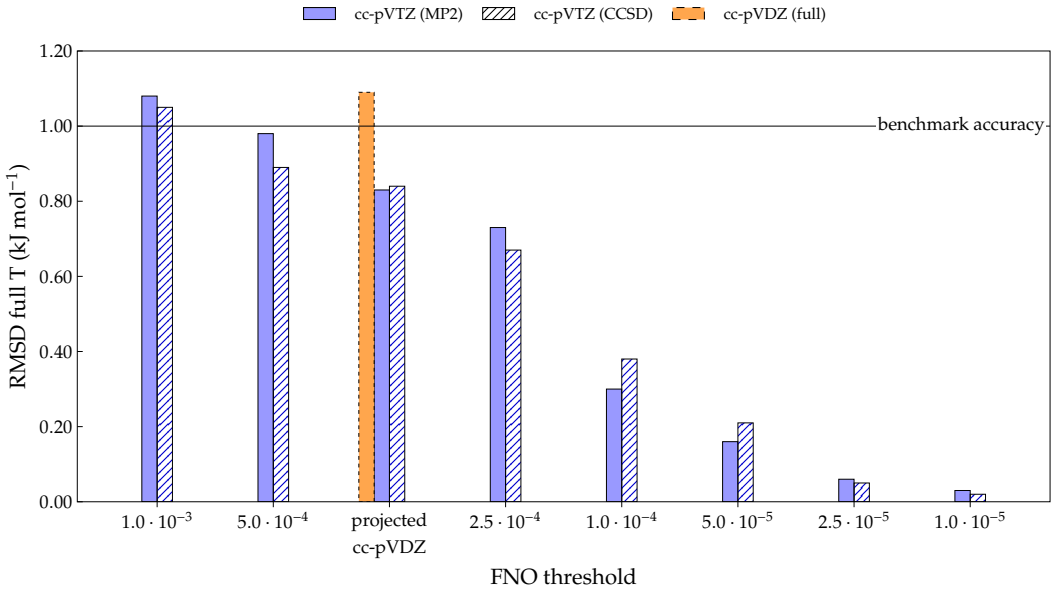


FIGURE 6.12: RMSD for the full T contribution evaluated over the KAW test set, relative to the cc-pVTZ reference, as a function of the FNO truncation threshold. Results based on MP2-derived orbitals are labeled as (MP2), those from CCSD-based occupations as (CCSD), and calculations with the full virtual space as (full). Projected cc-pVnZ thresholds retain the same number of COs as the corresponding cc-pVnZ basis sets.

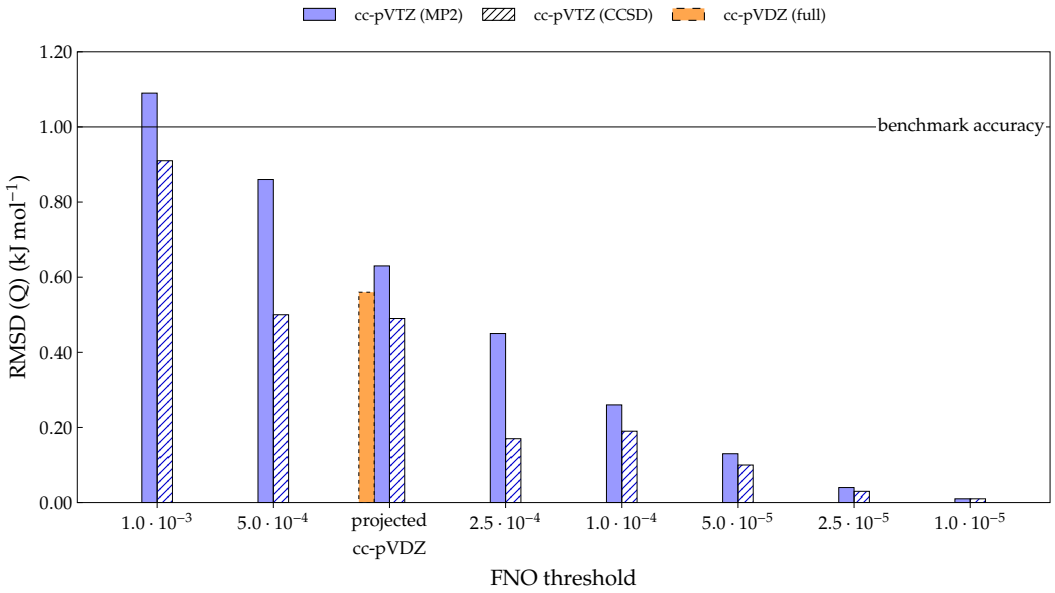


FIGURE 6.13: RMSD for the (Q) contribution evaluated over the KAW test set, relative to the cc-pVTZ reference, as a function of the FNO truncation threshold. Results based on MP2-derived orbitals are labeled as (MP2), those from CCSD-based occupations as (CCSD), and calculations with the full virtual space as (full). Projected cc-pVnZ thresholds retain the same number of COs as the corresponding cc-pVnZ basis sets.

For systems having single-reference character MP2-derived NOs are fully adequate across

all thresholds. Inspection of the benchmark carried out on the atomization energies from the Karton test set (Figures 6.10, 6.11, and A.5 of Appendix A) reveals a clear hierarchy in convergence: the (Q) contribution reaches benchmark accuracy with the fewest retained orbitals, followed by the full T contribution, while the (T) term requires the largest number of orbitals to achieve comparable accuracy. In other words, at any given threshold, lower RMSD values are consistently obtained for (Q) than for full T, and for full T than for (T). Nevertheless, for the full T and (Q) contributions, all thresholds already yield results within the chemical accuracy target, whereas for perturbative triples this level is reached only starting from the $2.5 \cdot 10^{-4}$ threshold.

In contrast, for the systems present in the KAW test set, exhibiting a more pronounced multi-reference character (Figures 6.12, 6.13, and A.1 of Appendix A), CCSD-derived NOs provide a noticeable advantage, especially at larger thresholds, where MP2-based NOs may fail to reproduce the correct correlation pattern. In this regime, the full T contribution falls below the benchmark target accuracy starting from $5.0 \cdot 10^{-4}$, with the only exception being the full cc-pVDZ RMSD. For the (Q) contribution, the benchmark accuracy is already reached at $1.0 \cdot 10^{-3}$ with CCSD-derived NOs, whereas the (T) contribution attains this level only from $5.0 \cdot 10^{-5}$ when using MP2 densities. The same hierarchy in RMSD values is maintained, with (Q) yielding the lowest deviations, followed by full T, which in turn outperforms (T).

By decoupling the choice of the parent basis set from the effective number of COs, this strategy enables fine control over the cost–accuracy trade-off and can extend platinum-standard accuracy to systems that would otherwise be computationally prohibitive. For instance, compressing the cc-pV5Z basis set with $t_{\mathrm{NO}} = 2.5 \cdot 10^{-5}$ reduces n_{CO} by nearly 70% while keeping the (T) RMSD near 1 kJ mol^{-1} in the Karton test set, thus achieving substantial savings with minimal loss of accuracy.

6.3 Conclusions

In this Chapter, a new strategy to overcome the computational limitations of high-level CC methods by combining large basis sets with the systematic compression of the virtual space through the FNO approximation has been presented and systematically assessed. NOs derived from MP2- or CCSD-level density matrices enable the controlled removal of the least-occupied virtual components while preserving size-consistency and allowing continuous control of accuracy. Extensive benchmarks indicate that large basis sets with reduced virtual spaces consistently outperform smaller uncompressed bases with a comparable number of COs, particularly for post-CCSD(T) treatments where the computational cost is dominated by the virtual space dimension. This framework allows smooth interpolation between accuracy levels typical of the cc-pVnZ hierarchy while retaining the high-quality occupied space of the larger basis set. Beyond the direct performance improvements, the decoupling of the parent basis from the virtual space size enables the development of consistent and scalable approach, in which accuracy can be tuned through the occupation threshold, providing a

route to benchmark-level accuracy at affordable computational cost, thereby extending the applicability of high-level CC methods to molecular systems of substantially larger size than previously feasible.

Chapter 7

Conclusions and Future Perspectives

The main goal of this PhD thesis has been the development and validation of composite QC methodologies capable of providing reliable thermochemical, structural, and rotational–vibrational spectroscopic properties for molecular systems of increasing size, while maintaining an optimal balance between accuracy and computational efficiency. This research contributes to the broader framework of predictive quantitative quantum chemistry, which aims to extend the reach of benchmark-level approaches, traditionally limited to small molecules, to the description of chemically relevant systems of medium or large size. To this end, the work has focused on the design and implementation of advanced composite schemes integrating state-of-the-art wave function-based methods, explicit correlation, and orbital-space compression techniques.

The thesis has pursued three closely connected lines of research. First, the junChS-F12 model has been improved through the introduction of the CCSD(F12*)(T+) formulation, replacing the F12b approximation and canonical (T) treatment, thereby providing a size-consistent description of correlation effects for energy evaluations and substantially reducing the basis set incompleteness error without increasing the computational cost. The revised protocol has been benchmarked against accurate thermochemical data, yielding MUEs within chemical accuracy for atomization and reaction energies of both closed- and open-shell systems. Importantly, the new scheme has maintained excellent performance for non-covalent interactions, conformational, and tautomeric equilibria, with average deviations around 0.1 kJ mol^{-1} relative to the original junChS-F12 approach. Beyond energetics, the model has been extended to the prediction of equilibrium geometries and harmonic frequencies. For the structural description, the addition of CV corrections at the MP2-F12 level to the explicitly correlated CCSD(T)-optimized geometry has proved essential for reaching sub-milliångström accuracy, while the effect of the CBS extrapolation has been found to be negligible for F12-based methods. For harmonic vibrational frequencies, the explicitly correlated CCSD(T) in conjunction with a partially augmented triple-ζ basis set has yielded an RMSD of about 5 cm^{-1} across a diverse test set, ensuring reliable vibrational assignments.

In a second step, building on the results obtained for the revised junChS-F12, the Pisa Composite Scheme family, a new generation of composite schemes designed to surpass the accuracy of ChS-type models while maintaining comparable computational demands, has been developed and validated. The PCS philosophy has relied on a consistent definition of

the levels of theory across different quantities, with its methodological core centered on the energetic component. In the PCS models, the valence correlation energy has been redefined with respect to the ChS schemes through a more sophisticated hierarchical organization, in which the MP2 and post-MP2 contributions have been extrapolated separately to the CBS limit. Additional improvements have included the systematic use of cc-pVnZ-F12 basis sets, together with optimized formulations for third-row elements, and a more advanced treatment of CV correlation effects. Collectively, these features have ensured a smooth and well-behaved convergence pattern, enhancing both the robustness and transferability of the model across diverse chemical environments. Starting from medium-sized molecular systems, the first energetic scheme has been developed using conventional methods for both valence and core-valence contributions. Its explicitly correlated extension has preserved the overall structure of the original formulation while replacing the conventional valence terms with their explicitly correlated counterparts, namely CCSD(F12*)(T+) and MP2-F12, thereby improving the convergence toward the CBS limit. The refined explicitly correlated PCS variant has retained the same overall structure while adopting a revised extrapolation exponent for the post-MP2-F12 component, thereby achieving a more balanced and accurate description of electronic energies within the 5–15 atom molecular size range. This final formulation has provided the high-accuracy energetic backbone of the PCS family, upon which the subsequent extension to larger molecular systems has been achieved through the introduction of the FNO approximation. This approach has markedly reduced the cost of CC correlation calculations while retaining quantitative reliability, allowing the treatment of molecules containing up to about 50 atoms without resorting to local correlation approximations. Concerning the prediction of structural properties, different strategies have been adopted depending on molecular size. For systems containing approximately 5–15 atoms, equilibrium structures have been obtained from F12-based CC optimizations complemented by MP2 CV corrections. For larger systems, accurate and computationally efficient geometries have been computed through double-hybrid functionals augmented by MP2-based CV corrections or by employing one-parameter bond-corrected PCS formulations. In both cases, explicit CBS extrapolations have been avoided. Vibrational properties have been treated through hybrid and double-hybrid density functionals, with anharmonic corrections derived in the framework of VPT2 for semi-rigid systems. Extensive benchmarks and applications to "biochemical building blocks" and technologically relevant compounds have demonstrated that the PCS protocols achieve performance comparable to more demanding composite approaches, consistently reproducing thermochemical, structural, and vibrational quantities for molecules containing elements of the first three rows, with deviations within chemical accuracy when compared to experimental or high-level theoretical data.

The final research stage has investigated the use of the FNO approximation as an efficient strategy to incorporate post-CCSD(T) correlation effects within composite schemes. By selectively compressing the virtual orbital space according to occupation thresholds, the FNO approach has introduced a continuous, user-controllable balance between computational

cost and accuracy. Unlike canonical methods, where the achievable precision is effectively discretized by the choice of the basis set cardinal number, the FNO strategy has allowed a smooth interpolation between levels of accuracy, as the number of correlated virtual orbitals can be tuned according to the targeted accuracy and available resources. Two alternative truncation procedures have been analyzed: a direct cut-off scheme, in which virtual orbitals were discarded according to their occupation numbers, and a projection-based formulation, where natural orbitals generated in a larger basis were projected onto a smaller one and subsequently recanonicalized, resulting in a virtual space whose dimension exactly matches that of the target basis set. This latter approach has enabled an independent assessment of two complementary effects: the improvement of the occupied-orbital description, enhanced by the larger basis set, and the expansion of the virtual manifold, whose size and quality are governed by the chosen occupation threshold. Together, these variants have provided a versatile framework to explore the trade-off between accuracy and efficiency in a systematically controllable manner. Extensive benchmarks have been carried out for a broad range of excitation levels, including perturbative triples, quadruples (in several approximate formulations), and pentuples, as well as full triple and quadruple excitations. The test sets have covered atomization and reaction energies for both closed- and open-shell species, as well as systems with pronounced multi-reference character. The results have demonstrated that tailored FNO compressions of post-CCSD(T) contributions preserve the accuracy of the corresponding reference energies, with deviations typically below 1 kJ mol^{-1}, while substantially reducing the computational cost. Large basis set calculations employing compressed virtual spaces have consistently outperformed conventional smaller-basis computations of comparable size, confirming the robustness and transferability of the approach. The FNO approximation has therefore provided a pragmatic and scalable route to extend the reach of post-CCSD(T) composite schemes, offering tunable and systematically improvable accuracy at a fraction of the canonical computational cost, and pushing the boundaries of feasible spectroscopic-quality QC calculations for small-to-medium molecular systems.

Overall, the research carried out has established a pathway toward reliable, scalable, and transferable QC protocols defining the most appropriate accuracy-to-cost trade-off for the accurate characterization of energetic, structural, and vibrational quantities as a function of the size and complexity of the system under investigation.

Building on the outcomes of the present thesis, a number of further developments can be envisaged. First, the energetic variant of the PCS family employing the FNO approximation, designed for molecular systems containing approximately 15–60 atoms, can be refined to achieve highly accurate estimations of heats of formation. Based on the promising preliminary results obtained with the corrective factor scaled by the number of non-hydrogen atoms, a non-empirical, transferable, and reliable strategy can be developed to ensure the generality of the approach. Such improvements will aim to consolidate this FNO-based variant as a general-purpose protocol for high-accuracy thermochemical predictions. In addition, the composite framework can be further extended to other spectroscopic molecular properties

by exploiting the flexibility of the FNO approximation, enabling accurate yet feasible calculations for systems beyond medium size. The same strategy may also be adapted to the treatment of species exhibiting pronounced multi-reference character, thus broadening the applicability of composite methods to more complex electronic structures. Further progress may involve the integration of the composite methodologies developed in this work within multi-scale approaches. Such an extension would allow the accurate evaluation of energetics and molecular properties in condensed-phase or heterogeneous environments, thereby expanding the scope of the developed protocols beyond the gas phase.

In conclusion, the methodologies developed in this work enable the accurate investigation of astrochemical, atmospheric, and biological processes, where reliable thermochemical and rotational–vibrational spectroscopic predictions are essential. Their capability to deliver near-spectroscopic accuracy for systems of increasing size makes them powerful tools for the characterization of fundamental and prototypical molecules relevant both to the chemistry of the Earth and to that of the cosmic environment, defining a coherent, scalable, and predictive framework for molecular modeling that fully embodies the aim of the Molecular Sciences for Earth and Space PhD program, aiming to extend the understanding of chemical complexity across terrestrial and extraterrestrial domains.

Appendices

Appendix A

Supplementary Data for the Evaluation of Energetic Properties

This Appendix collects all the supplementary data related to the energetic analyses discussed in Chapters 4, 5, and 6.

A.1 KAW Test Set

TABLE A.1: MAX, MUE, and RMSD for KAW test set in kJ mol^{-1}. Values computed at the $E_{\mathrm{CCSD(T)}}^{\text{jun-cc-pV(T+}d\text{)Z}} + \Delta E_{\mathrm{MP2}}^{\infty}$ level, using the jun-cc-pV(n+d)Z basis sets (n = T, Q) and the n^{-3} two-point Helgaker extrapolation formula, together with its modified variants.

	{2,3}[a]			{3,4}[b]			Varandas[c]		
	n^{-3}	n^{-4}	n^{-5}	$n^{-3}d$	n^{-4}	n^{-5}	n^{-3}	n^{-4}	n^{-5}
Atomization energy									
MAX	20.59	26.67	29.95	10.58	19.14	24.41	19.24	25.74	29.29
MUE	4.44	7.52	9.27	4.01	3.91	6.31	3.94	7.02	8.92
RMSD	6.13	9.00	10.72	4.77	5.57	7.87	5.61	8.53	10.37
Closed-shell reaction energy									
MAX	7.58	8.96	9.70	5.15	7.25	8.44	7.27	8.75	9.55
MUE	1.58	1.73	1.84	1.54	1.54	1.67	1.54	1.71	1.82
RMSD	2.33	2.83	3.16	2.11	2.24	2.62	2.25	2.74	3.09
Open-shell reaction energy									
MAX	22.01	27.93	31.13	11.57	20.60	25.73	20.70	27.03	30.48
MUE	3.54	5.09	6.09	3.02	3.27	4.43	3.29	4.82	5.88
RMSD	5.39	7.11	8.16	4.04	5.04	6.43	5.06	6.83	7.94
Overall									
MAX	22.01	27.93	31.13	11.57	20.60	25.73	20.70	27.03	30.48
MUE	3.45	5.29	6.39	3.08	3.13	4.55	3.15	4.99	6.16
RMSD	5.21	7.28	8.53	4.03	4.80	6.46	4.83	6.93	8.28

[a] n = 2, 3. [b] n = 3, 4. [c] n = 3.630, 5.335.[446,447] [d] Corresponds to the junChS scheme.[89]

TABLE A.2: MAX, MUE, and RMSD for KAW test set in kJ mol^{-1}. Values computed at the $E^{\text{jun-cc-pV(T+}d\text{)Z}}_{\text{CCSD(T)-F12b}} + \Delta E^{\infty}_{\text{MP2-F12}}$ level, using the jun-cc-pV(n+d)Z basis sets (n = T, Q) and the n^{-3} two-point Helgaker extrapolation formula, together with its modified variants.

	{2,3}[a]			{3,4}[b]			Varandas[c]		
	n^{-3}	n^{-4}	n^{-5}	n^{-3d}	n^{-4}	n^{-5}	n^{-3}	n^{-4}	n^{-5}
Atomization energy									
MAX	13.77	13.83	13.99	13.93	13.79	13.72	13.79	13.79	13.96
MUE	5.49	5.63	5.70	5.25	5.46	5.58	5.46	5.61	5.69
RMSD	6.26	6.39	6.46	6.03	6.23	6.34	6.23	6.37	6.45
Closed-shell reaction energy									
MAX	3.32	3.44	3.65	3.36	3.32	3.30	3.32	3.38	3.61
MUE	0.75	0.73	0.73	0.84	0.76	0.74	0.76	0.73	0.73
RMSD	1.17	1.18	1.20	1.20	1.17	1.18	1.17	1.18	1.19
Open-shell reaction energy									
MAX	12.54	11.90	11.55	13.67	12.69	12.14	12.68	11.99	11.62
MUE	2.87	2.89	2.90	2.83	2.86	2.88	2.86	2.89	2.90
RMSD	3.69	3.68	3.68	3.72	3.70	3.69	3.70	3.68	3.68
Overall									
MAX	13.77	13.83	13.99	13.93	13.79	13.72	13.79	13.79	13.96
MUE	3.42	3.48	3.51	3.33	3.41	3.46	3.41	3.47	3.51
RMSD	4.57	4.64	4.68	4.46	4.56	4.61	4.56	4.63	4.67

[a] n = 2, 3. [b] n = 3, 4. [c] n = 3.630, 5.335.[446,447] [d] Corresponds to the old formulation of the junChS-F12 scheme.[90]

TABLE A.3: MAX, MUE, and RMSD for KAW test set in kJ mol^{-1}. Values computed at the $E^{\text{jun-cc-pV(T+}d\text{)Z}}_{\text{CCSD(F12*)(T+)}} + \Delta E^{\infty}_{\text{MP2-F12}}$ level, using the jun-cc-pV(n+d)Z basis sets (n = T, Q) and the n^{-3} two-point Helgaker extrapolation formula, together with its modified variants.

	{2,3}[a]			{3,4}[b]			Varandas[c]		
	n^{-3}	n^{-4}	n^{-5}	n^{-3d}	n^{-4}	n^{-5}	n^{-3}	n^{-4}	n^{-5}
Atomization energy									
MAX	10.79	11.05	11.19	10.33	10.73	10.95	10.73	11.01	11.16
MUE	2.90	3.04	3.13	2.69	2.87	2.99	2.87	3.02	3.11
RMSD	3.97	4.09	4.16	3.78	3.95	4.05	3.95	4.07	4.14
Closed-shell reaction energy									
MAX	5.23	5.19	5.17	5.31	5.24	5.21	5.24	5.20	5.17
MUE	1.07	1.09	1.11	1.06	1.06	1.08	1.06	1.09	1.11
RMSD	1.69	1.69	1.69	1.72	1.69	1.69	1.69	1.69	1.69
Open-shell reaction energy									
MAX	6.09	6.12	6.14	6.57	6.08	6.11	6.08	6.12	6.14
MUE	2.05	2.02	2.00	2.14	2.06	2.03	2.06	2.02	2.00
RMSD	2.71	2.66	2.63	2.82	2.72	2.67	2.72	2.66	2.64
Overall									
MAX	10.79	11.05	11.19	10.33	10.73	10.95	10.73	11.01	11.16
MUE	2.16	2.21	2.24	2.11	2.16	2.19	2.16	2.20	2.24
RMSD	3.10	3.15	3.17	3.05	3.10	3.13	3.10	3.14	3.17

[a] n = 2, 3. [b] n = 3, 4. [c] n = 3.630, 5.335.[446,447] [d] Corresponds to the new formulation of the junChS-F12 scheme.[184]

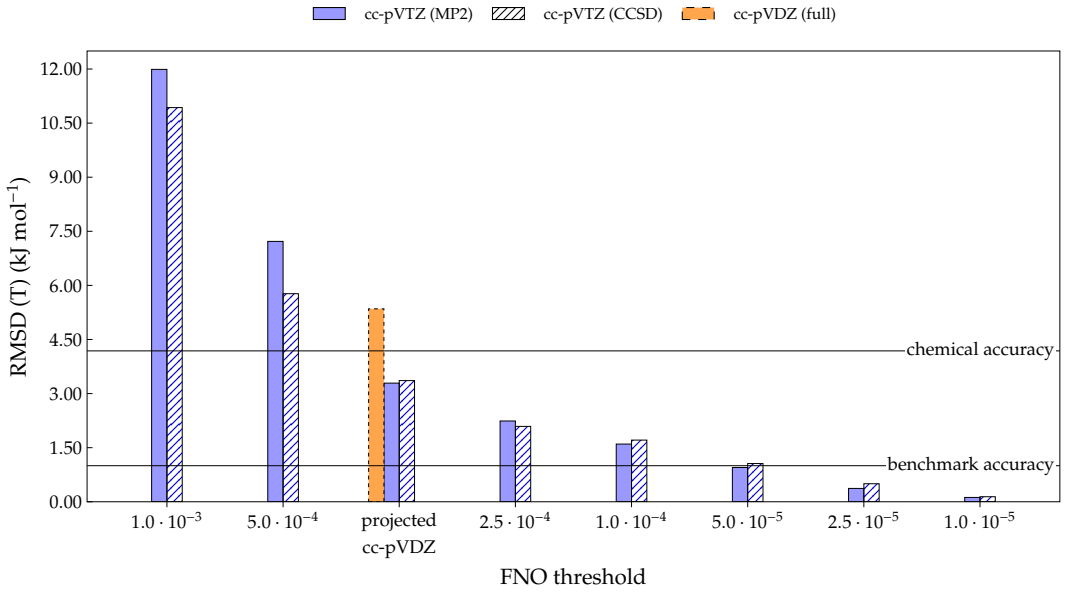


FIGURE A.1: RMSD for the (T) contribution evaluated over the KAW test set, relative to the cc-pVTZ reference, as a function of the FNO truncation threshold. Results based on MP2-derived orbitals are labeled as (MP2), those from CCSD-based occupations as (CCSD), and calculations with the full virtual space as (full). Projected cc-pVnZ thresholds retain the same number of COs as the corresponding cc-pVnZ basis sets.

A.2 W4-11 Test Set

TABLE A.4: MAX, MUE, and RMSD values delivered by different CSs for the W4-11 set,[83] considering the valence contribution. In the PCS3 and FPCS3 schemes, the extrapolation exponent for the low-level valence MP2-F12 term is fixed to 3, whereas the exponent for the high-level post-MP2-F12 valence components and for the additional basis functions are reported in curly brackets. The CC model employed in the junChS-F12 schemes is indicated in round brackets. Molecules yielding the MAX values are indicated in parentheses. All values are in kJ mol^{-1}.

	MAX	MUE	RMSD
All molecules			
W1-F12	-	1.05	2.43
W2-F12	-	0.54	0.71
PCS3 {3}	5.56 (S_4)	0.77	1.21
PCS3 {3, +f,g}	4.99 ($HC_2^{\bullet}$)	0.60	0.90
PCS3 {3, +f}	4.99 ($HC_2^{\bullet}$)	0.65	0.91
PCS3 {5}	4.02 (P_4)	0.71	0.95
PCS3 {5, +f,g}	4.11 (P_4)	0.63	0.86
PCS3 {5, +f}	3.01 (P_4)	0.68	0.87
junChS-F12 (CCSD(F12*)(T+))	6.98 (P_4)	1.00	1.42
junChS-F12 (CCSD(T)-F12b)	17.50 (P_4)	4.81	5.41
SVECV-f12	12.63 (S_4)	2.26	3.08
FPCS3 {3}	9.13 (P_4)	2.34	2.76
FPCS3 {5}	6.06 (P_4)	2.04	2.54
1st- and 2nd-row atoms			
W1-F12	-	0.42	0.63
W2-F12	-	0.50	0.67
PCS3 {3}	4.99 ($HC_2^{\bullet}$)	0.61	0.89
PCS3 {5}	1.83 ($FO_2^{\bullet}$)	0.58	0.72
junChS-F12 (CCSD(F12*)(T+))	3.68 (CF_4)	0.78	1.06
junChS-F12 (CCSD(T)-F12b)	9.44 (CH_3CO_2H)	4.67	5.11
SVECV-f12	3.50 (CH_3CO_2H)	1.58	1.81
FPCS3 {3}	5.78 (B_2H_6)	2.69	2.98
FPCS3 {5}	5.62 (CH_3CO_2H)	2.44	2.85
1st- and 3rd-row atoms			
W1-F12	-	2.51	4.39
W2-F12	-	0.59	0.75
PCS3 {3}	5.56 (S_4)	1.17	1.76
PCS3 {3, +f,g}	3.96 (P_4)	0.60	0.92
PCS3 {3, +f}	2.73 (P_4)	0.77	0.98
PCS3 {5}	4.02 (P_4)	1.05	1.36
PCS3 {5, +f,g}	4.11 (P_4)	0.76	1.14
PCS3 {5, +f}	3.01 (P_4)	0.94	1.16
junChS-F12 (CCSD(F12*)(T+))	6.98 (P_4)	1.53	2.03
junChS-F12 (CCSD(T)-F12b)	17.50 (P_4)	5.14	6.10
SVECV-f12	12.63 (S_4)	3.92	4.95
FPCS3 {3}	9.13 (P_4)	1.49	2.13
FPCS3 {5}	6.06 (P_4)	1.07	1.54

TABLE A.5: MAX, MUE, and RMSD values delivered by different levels of theory for the W4-11 set,[83] considering the core-valence contribution. Additional basis functions are reported in curly brackets. Molecules yielding the MAX values are indicated in parentheses. All values are in kJ mol^{-1}.

	MAX	MUE	RMSD
All molecules			
MP2/cc-pwCVTZ	10.34 (SiF_4)	2.12	2.78
MP2/cc-pwCVTZ {+g}	9.15 (SiF_4)	2.04	2.64
MP2/cc-pCVTZ	10.95 (SiF_4)	2.37	3.10
MP2-F12/cc-pwCVTZ	13.17 (SiF_4)	3.19	3.79
CCSD(T)/cc-pwCVTZ	3.06 (AlF_3)	0.61	0.82
CCSD(T)/cc-pwCVTZ {+g}	4.09 (P_4)	0.67	0.92
MP2/cc-pwCV(D,T)Z + CCSD(T)/cc-pwCVDZ	6.26 (P_4)	0.90	1.15
MP2/cc-pwCV(D,T)Z + CCSD(T)/cc-pwCVDZ {+f,g}	3.04 (SiF_4)	0.85	1.01
MP2/cc-pwCV(D,T)Z + CCSD(T)/cc-pwCVDZ {+f}	4.73 (SiF_4)	0.90	1.12
FNO-CCSD(T)/cc-pwCVTZ	4.77 (AlF_3)	1.40	1.76
FNO-CCSD(T)/cc-pwCVTZ {+g}	4.08 (CF_4)	1.38	1.75
MP2/cc-pwCV(D,T)Z + FNO-CCSD(T)/cc-pwCVDZ	9.94 (P_4)	1.56	2.25
MP2/cc-pwCV(D,T)Z + FNO-CCSD(T)/cc-pwCVDZ {+f,g}	4.95 (C_2)	1.22	1.51
MP2/cc-pwCV(D,T)Z + FNO-CCSD(T)/cc-pwCVDZ {+f}	4.95 (C_2)	1.47	1.85
1st- and 2nd-row atoms			
MP2/cc-pwCVTZ	4.97 (C_2N_2)	2.05	2.43
MP2/cc-pCVTZ	6.25 (propyne)	2.39	2.91
MP2-F12/cc-pwCVTZ	6.35 (C_2N_2)	2.95	3.28
CCSD(T)/cc-pwCVTZ	1.73 (C_3H_4)	0.58	0.75
MP2/cc-pwCV(D,T)Z + CCSD(T)/cc-pwCVDZ	2.55 (C_2)	0.88	0.98
FNO-CCSD(T)/cc-pwCVTZ	4.08 (CF_4)	1.53	1.88
MP2/cc-pwCV(D,T)Z + FNO-CCSD(T)/cc-pwCVDZ	4.95 (C_2)	1.21	1.50
1st- and 3rd-row atoms			
MP2/cc-pwCVTZ	10.34 (SiF_4)	2.28	3.48
MP2/cc-pwCVTZ {+g}	9.15 (SiF_4)	2.01	3.07
MP2/cc-pCVTZ	10.95 (SiF_4)	2.31	3.52
MP2-F12/cc-pwCVTZ	13.17 (SiF_4)	3.77	4.81
CCSD(T)/cc-pwCVTZ	3.06 (AlF_3)	0.68	0.99
CCSD(T)/cc-pwCVTZ {+g}	4.09 (P_4)	0.90	1.25
MP2/cc-pwCV(D,T)Z + CCSD(T)/cc-pwCVDZ	6.26 (P_4)	0.96	1.47
MP2/cc-pwCV(D,T)Z + CCSD(T)/cc-pwCVDZ {+f,g}	3.04 (SiF_4)	0.79	1.06
MP2/cc-pwCV(D,T)Z + CCSD(T)/cc-pwCVDZ {+f}	4.73 (SiF_4)	0.97	1.41
FNO-CCSD(T)/cc-pwCVTZ	4.77 (AlF_3)	1.06	1.42
FNO-CCSD(T)/cc-pwCVTZ {+g}	3.93 (SiF_4)	1.01	1.38
MP2/cc-pwCV(D,T)Z + FNO-CCSD(T)/cc-pwCVDZ	9.94 (P_4)	2.42	3.45
MP2/cc-pwCV(D,T)Z + FNO-CCSD(T)/cc-pwCVDZ {+f,g}	3.34 (SiO)	1.26	1.53
MP2/cc-pwCV(D,T)Z + FNO-CCSD(T)/cc-pwCVDZ {+f}	4.84 ($AlCl_3$)	2.09	2.49

A.3 W4-17 test set

TABLE A.6: MAX, MUE, and RMSD values delivered by different CSs for the W4-17 set,[84] considering the valence contribution. In the PCS3 and FPCS3 schemes, the extrapolation exponent for the low-level valence MP2-F12 term is fixed to 3, whereas the exponent for the high-level post-MP2-F12 valence components and for the additional basis functions are reported in curly brackets. The CC model employed in the junChS-F12 schemes is indicated in round brackets. Molecules yielding the MAX values are indicated in parentheses. All values are in kJ mol^{-1}.

	Total			non-MR			MR		
Valence	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD
All molecules									
PCS3 {3}	5.56 (S_4)	0.88	1.28	5.53 (P_4)	0.84	1.19	5.56 (S_4)	1.51	2.13
PCS3 {3, +f,g}	4.99 ($H_2C^\bullet$)	0.74	1.02	4.99 ($H_2C^\bullet$)	0.74	1.03	2.72 (S_4)	0.73	1.02
PCS3 {3, +f}	4.99 ($HC_2^\bullet$)	0.86	1.16	4.99 ($H_2C^\bullet$)	0.88	1.19	1.53 (S_4)	0.61	0.78
PCS3 {5}	4.02 (P_4)	0.76	1.00	4.02 (P_4)	0.72	0.94	3.01 (S_4)	1.24	1.48
PCS3 {5, +f,g}	4.11 (P_4)	0.65	0.87	4.11 (P_4)	0.63	0.83	2.78 (S_4)	0.87	1.19
PCS3 {5, +f}	3.95 (C_2Cl_6)	0.77	1.00	3.95 (C_2Cl_6)	0.76	0.99	1.71 ($ClO_2^\bullet$)	0.86	1.02
junChS-F12 (CCSD(F12*)(T+))	6.98 (P_4)	1.13	1.64	6.98 (P_4)	1.14	1.67	2.87 (C_2)	0.93	1.24
junChS-F12 (CCSD(T)-F12b)	17.50 (P_4)	6.09	6.90	17.50 (P_4)	6.23	7.05	9.46 (S_4)	4.59	5.08
SVECV-f12	12.63 (S_4)	2.80	3.67	11.78 ($HClO_4$)	2.68	3.48	12.63 (S_4)	4.35	5.60
FPCS3 {3}	9.13 (P_4)	2.71	3.32	9.13 (P_4)	2.78	3.40	4.12 (B_2)	2.03	2.34
FPCS3 {5}	10.69 (benzene)	2.65	3.36	10.69 (benzene)	2.77	3.47	3.34 ($ClO_3^\bullet$)	1.34	1.68
1st- and 2nd-row atoms									
PCS3 {3}	4.99 ($HC_2^\bullet$)	0.76	1.05	4.99 ($HC_2^\bullet$)	0.78	1.07	0.96 (C_2)	0.43	0.53
PCS3 {5}	2.40 (C_2F_6)	0.61	0.77	2.40 (C_2F_6)	0.60	0.76	1.44 (C_2)	0.69	0.80
junChS-F12 (CCSD(F12*)(T+))	6.80 (C_2F_6)	0.86	1.28	6.80 (C_2F_6)	0.84	1.26	2.87 (C_2)	1.03	1.44
junChS-F12 (CCSD(T)-F12b)	14.81 (n-pentane)	5.93	6.72	14.81 (n-pentane)	6.11	6.89	4.46 (O_3)	2.87	3.03
SVECV-f12	6.22 (N_2O_4)	1.78	2.06	6.22 (N_2O_4)	1.80	2.07	2.74 (O_3)	1.22	1.50
FPCS3 {3}	8.99 (benzene)	3.30	3.76	8.99 (benzene)	3.37	3.82	4.12 (B_2)	2.58	2.82
FPCS3 {5}	10.69 (benzene)	3.20	3.85	10.69 (benzene)	3.31	3.95	2.89 (B_2)	1.76	1.96
1st- and 3rd-row atoms									
PCS3 {3}	5.56 (S_4)	1.12	1.62	5.53 (P_4)	0.96	1.39	5.56 (S_4)	2.23	2.72
PCS3 {3, +f,g}	3.46 (P_4)	0.71	0.98	3.46 (P_4)	0.67	0.94	2.72 (S_4)	0.94	1.24
PCS3 {3, +f}	4.71 (C_2Cl_6)	1.04	1.35	4.71 (C_2Cl_6)	1.09	1.40	1.53 (S_4)	0.74	0.90
PCS3 {5}	4.02 (P_4)	1.03	1.31	4.02 (P_4)	0.95	1.22	3.01 (S_4)	1.60	1.80
PCS3 {5, +f,g}	4.11 (P_4)	0.73	1.02	4.11 (P_4)	0.69	0.96	2.78 (S_4)	0.99	1.39
PCS3 {5, +f}	3.95 (C_2Cl_6)	1.06	1.31	3.95 (C_2Cl_6)	1.07	1.33	1.71 ($ClO_2^\bullet$)	0.98	1.14
junChS-F12 (CCSD(F12*)(T+))	6.98 (P_4)	1.61	2.15	6.98 (P_4)	1.72	2.26	2.33 ($ClO_2^\bullet$)	0.87	1.08
junChS-F12 (CCSD(T)-F12b)	17.50 (P_4)	6.37	7.22	17.50 (P_4)	6.46	7.37	9.46 (S_4)	5.74	6.07
SVECV-f12	12.63 (S_4)	4.64	5.50	11.78 ($HClO_4$)	4.38	5.22	12.63 (S_4)	6.43	7.13
FPCS3 {3}	9.13 (P_4)	1.63	2.30	9.13 (P_4)	1.63	2.35	3.97 ($ClO_3^\bullet$)	1.67	1.95
FPCS3 {5}	6.33 (SC_4H_4)	1.64	2.19	6.33 (SC_4H_4)	1.73	2.28	3.34 ($ClO_3^\bullet$)	1.06	1.46

TABLE A.7: MAX, MUE, and RMSD values delivered by different levels of theory for the W4-17 set,[84] considering the core-valence contribution. Additional basis functions are reported in curly brackets. Molecules yielding the MAX values are indicated in parentheses. All values are in kJ mol^{-1}.

	Total			non-MR			MR		
Core-Valence	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD
All molecules									
MP2/cc-pwCVTZ	10.34 (SiF_4)	2.59	3.34	10.34 (SiF_4)	2.70	3.44	4.50 ($ClO_3^\bullet$)	1.51	1.97
MP2/cc-pwCVTZ {+g}	9.15 (SiF_4)	2.47	3.17	9.15 (SiF_4)	2.58	3.27	3.88 (S_4)	1.53	1.92
MP2/cc-pCVTZ	11.40 (benzene)	2.95	3.85	11.40 (benzene)	3.09	3.98	5.15 ($ClO_3^\bullet$)	1.59	2.14
MP2-F12/cc-pwCVTZ	15.80 (SF_6)	4.11	4.94	15.80 (SF_6)	4.15	4.97	10.62 ($ClO_3^\bullet$)	3.94	4.85
CCSD(T)/cc-pwCVTZ	3.25 (benzene)	0.75	1.01	3.25 (benzene)	0.73	0.99	2.60 (ClF_5)	1.03	1.24
CCSD(T)/cc-pwCVTZ {+g}	4.09 (P_4)	0.85	1.15	4.09 (P_4)	0.81	1.10	3.35 (S_4)	1.38	1.70
MP2/cc-pwCV(D,T)Z + CCSD(T)/cc-pwCVDZ	6.26 (P_4)	0.93	1.19	6.26 (P_4)	0.92	1.17	3.25 (S_4)	1.02	1.40
MP2/cc-pwCV(D,T)Z + CCSD(T)/cc-pwCVDZ {+f,g}	3.64 (C_2Cl_6)	0.92	1.10	3.64 (C_2Cl_6)	0.90	1.08	2.55 (C_2)	1.12	1.31
MP2/cc-pwCV(D,T)Z + CCSD(T)/cc-pwCVDZ {+f}	4.73 (SiF_4)	1.01	1.27	4.73 (SiF_4)	0.98	1.24	3.22 (ClF_5)	1.37	1.65
FNO-CCSD(T)/cc-pwCVTZ	7.03 (C_2F_6)	1.81	2.27	7.03 (C_2F_6)	1.89	2.34	3.07 (C_2)	1.07	1.30
FNO-CCSD(T)/cc-pwCVTZ {+g}	7.03 (C_2F_6)	1.74	2.20	7.03 (C_2F_6)	1.81	2.27	3.07 (C_2)	1.07	1.29
MP2/cc-pwCV(D,T)Z + FNO-CCSD(T)/cc-pwCVDZ	14.95 (C_2Cl_6)	1.85	2.71	14.95 (C_2Cl_6)	1.77	2.59	9.63 (S_4)	3.04	4.00
MP2/cc-pwCV(D,T)Z + FNO-CCSD(T)/cc-pwCVDZ {+f,g}	4.95 (C_2)	1.39	1.65	4.40 (C_2N_2)	1.36	1.59	4.95 (C_2)	1.87	2.29
MP2/cc-pwCV(D,T)Z + FNO-CCSD(T)/cc-pwCVDZ {+f}	5.25 ($HClO_4$)	1.69	2.04	5.25 ($HClO_4$)	1.63	1.96	4.95 (C_2)	2.49	2.91
1st- and 2nd-row atoms									
MP2/cc-pwCVTZ	9.02 (benzene)	2.66	3.24	9.02 (benzene)	2.76	3.32	2.93 (C_2)	1.40	1.72
MP2/cc-pCVTZ	11.40 (benzene)	3.14	3.94	11.40 (benzene)	3.27	4.04	3.21 (BN)	1.54	1.97
MP2-F12/cc-pwCVTZ	10.30 (benzene)	3.72	4.27	10.30 (benzene)	3.82	4.37	3.49 (C_2)	2.28	2.44
CCSD(T)/cc-pwCVTZ	3.25 (benzene)	0.73	0.96	3.25 (benzene)	0.74	0.98	1.62 (C_2)	0.59	0.80
MP2/cc-pwCV(D,T)Z + CCSD(T)/cc-pwCVDZ	2.91 (N_2O_4)	0.97	1.09	2.91 (N_2O_4)	0.95	1.07	2.55 (C_2)	1.29	1.46
FNO-CCSD(T)/cc-pwCVTZ	7.03 (C_2F_6)	1.95	2.42	7.03 (C_2F_6)	2.03	2.48	3.07 (C_2)	1.00	1.41
MP2/cc-pwCV(D,T)Z + FNO-CCSD(T)/cc-pwCVDZ	4.95 (C_2)	1.31	1.59	4.40 (C_2N_2)	1.27	1.50	4.95 (C_2)	2.42	2.94
1st- and 3rd-row atoms									
MP2/cc-pwCVTZ	10.34 (SiF_4)	2.46	3.50	10.34 (SiF_4)	2.59	3.66	4.50 ($ClO_3^\bullet$)	1.59	2.12
MP2/cc-pwCVTZ {+g}	9.15 (SiF_4)	2.14	3.04	9.15 (SiF_4)	2.21	3.16	3.88 (S_4)	1.62	2.04
MP2/cc-pCVTZ	10.95 (SiF_4)	2.61	3.69	10.95 (SiF_4)	2.75	3.86	5.15 ($ClO_3^\bullet$)	1.63	2.25
MP2-F12/cc-pwCVTZ	15.80 (SF_6)	4.83	5.98	15.80 (SF_6)	4.80	5.99	10.62 ($ClO_3^\bullet$)	5.05	5.94
CCSD(T)/cc-pwCVTZ	3.06 (AlF_3)	0.78	1.08	3.06 (AlF_3)	0.70	1.02	2.60 (ClF_5)	1.33	1.46
CCSD(T)/cc-pwCVTZ {+g}	4.09 (P_4)	1.07	1.44	4.09 (P_4)	0.95	1.31	3.35 (S_4)	1.91	2.10
MP2/cc-pwCV(D,T)Z + CCSD(T)/cc-pwCVDZ	6.26 (P_4)	0.85	1.35	6.26 (P_4)	0.85	1.35	3.25 (S_4)	0.83	1.36
MP2/cc-pwCV(D,T)Z + CCSD(T)/cc-pwCVDZ {+f,g}	3.64 (C_2Cl_6)	0.83	1.12	3.64 (C_2Cl_6)	0.80	1.11	2.09 (ClF_5)	1.01	1.20
MP2/cc-pwCV(D,T)Z + CCSD(T)/cc-pwCVDZ {+f}	4.73 (SiF_4)	1.08	1.55	4.73 (SiF_4)	1.04	1.51	3.22 (ClF_5)	1.42	1.76
FNO-CCSD(T)/cc-pwCVTZ	4.77 (AlF_3)	1.56	1.96	4.77 (AlF_3)	1.63	2.05	1.65 ($ClO_3^\bullet$)	1.11	1.22
FNO-CCSD(T)/cc-pwCVTZ {+g}	3.93 (SiF_4)	1.34	1.71	3.93 (SiF_4)	1.38	1.78	1.72 (S_4)	1.11	1.20
MP2/cc-pwCV(D,T)Z + FNO-CCSD(T)/cc-pwCVDZ	14.95 (C_2Cl_6)	2.83	4.02	14.95 (C_2Cl_6)	2.74	3.93	9.63 (S_4)	3.46	4.57
MP2/cc-pwCV(D,T)Z + FNO-CCSD(T)/cc-pwCVDZ {+f,g}	3.34 (SiO)	1.53	1.76	3.34 (SiO)	1.53	1.76	2.97 (ClF_5)	1.51	1.73
MP2/cc-pwCV(D,T)Z + FNO-CCSD(T)/cc-pwCVDZ {+f}	5.25 ($HClO_4$)	2.37	2.67	5.25 ($HClO_4$)	2.35	2.64	4.62 (S_4)	2.54	2.89

A.4 Karton Test Set

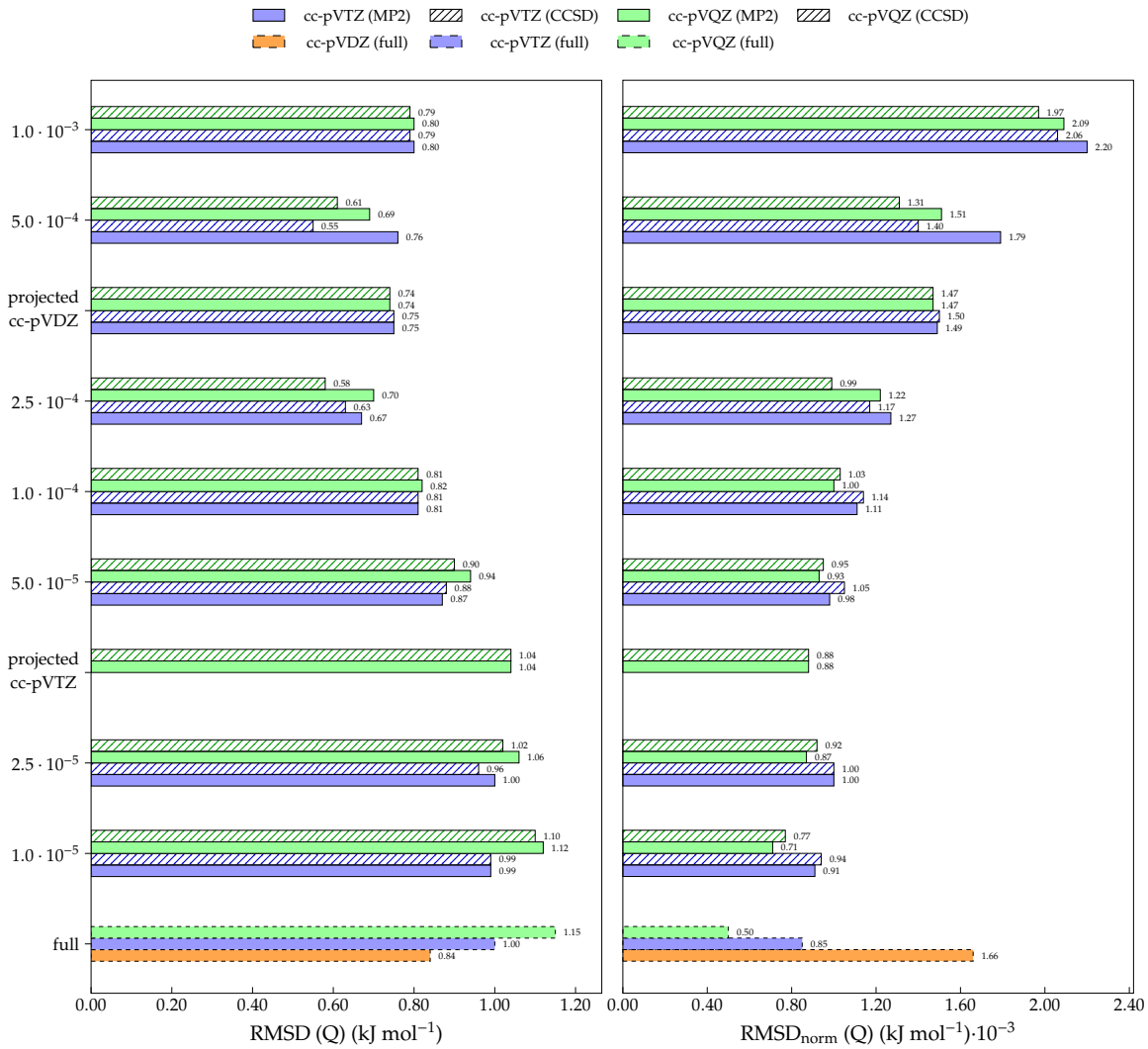


FIGURE A.2: RMSD (left) and RMSD normalized to the number of COs ($RMSD_{norm}$, right) for the (Q) contribution to the atomization energies in the Karton test set according to different FNO compression schemes, referenced to full Q cc-pV(5,6)Z values. Results obtained with orbitals selected from MP2-NO occupations are labeled as (MP2), those from CCSD-NO occupations as (CCSD), while calculations performed in the untruncated virtual space are denoted as (full). Projected cc-pVDZ and cc-pVTZ results correspond to the same number of COs as in the respective parent basis sets.

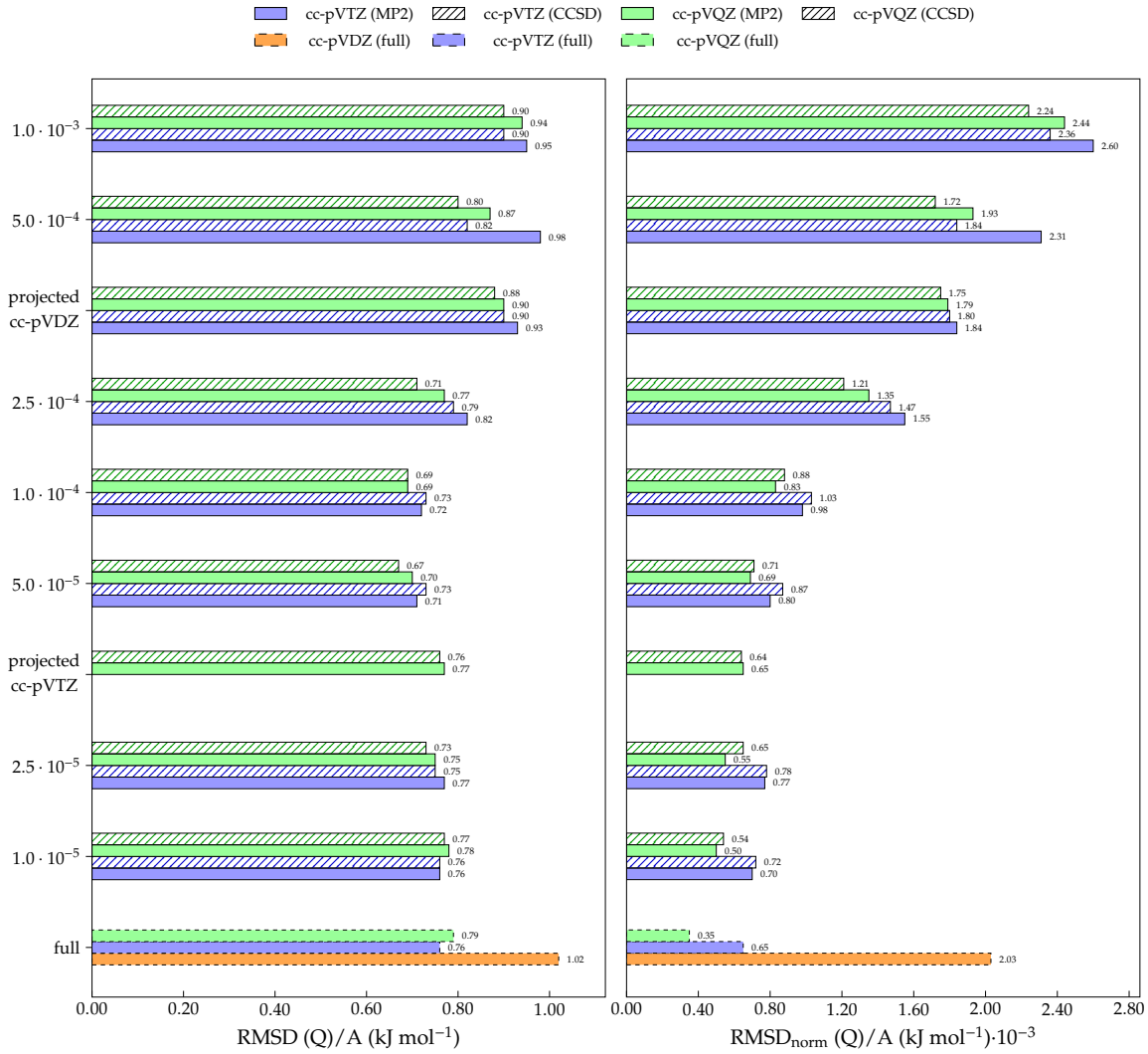


FIGURE A.3: RMSD (left) and RMSD normalized to the number of COs ($RMSD_{norm}$, right) for the (Q)/A contribution to the atomization energies in the Karton test set according to different FNO compression schemes, referenced to full Q cc-pV(5,6)Z values. Results obtained with orbitals selected from MP2-NO occupations are labeled as (MP2), those from CCSD-NO occupations as (CCSD), while calculations performed in the untruncated virtual space are denoted as (full). Projected cc-pVDZ and cc-pVTZ results correspond to the same number of COs as in the respective parent basis sets.

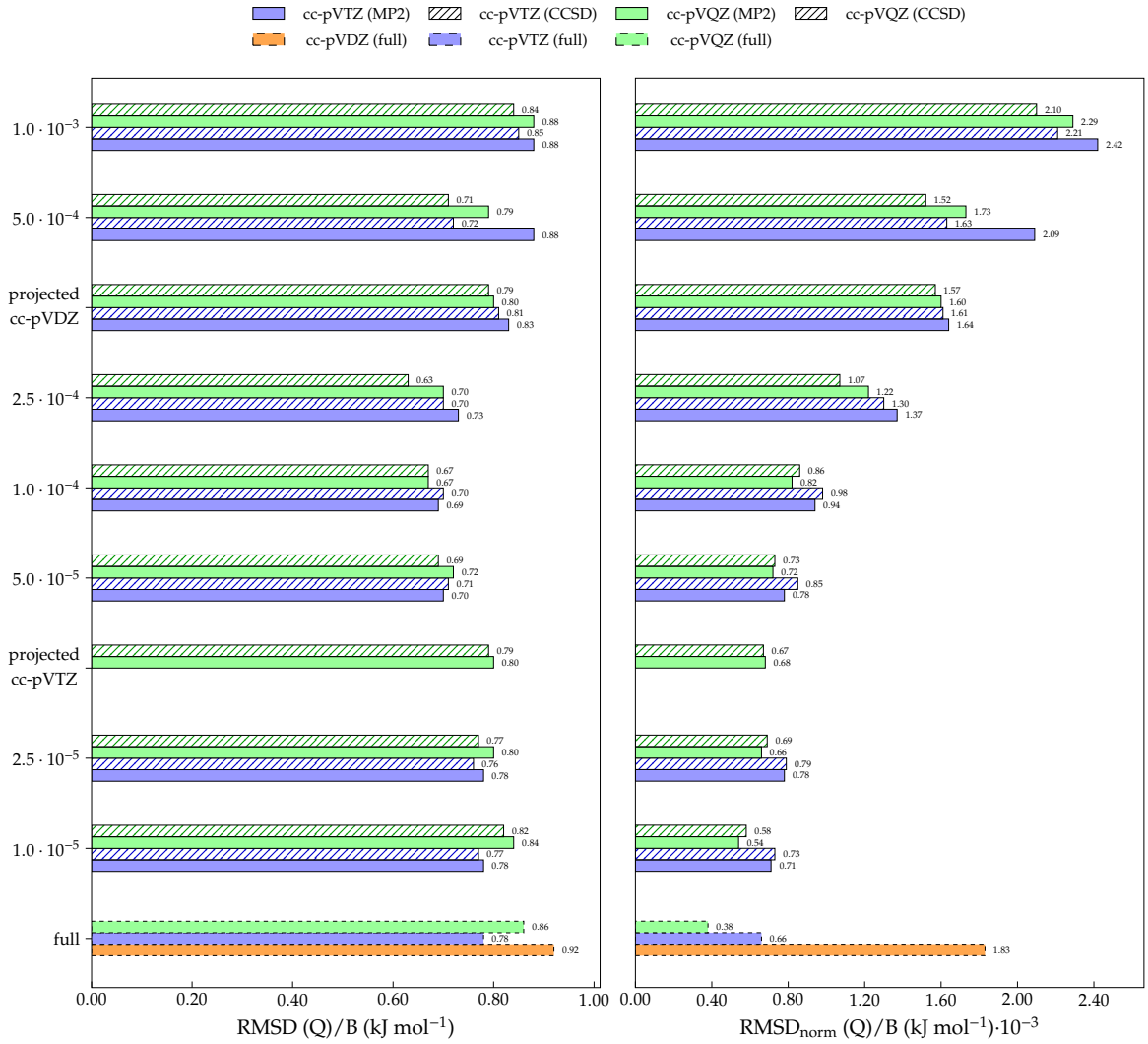


FIGURE A.4: RMSD (left) and RMSD normalized to the number of COs ($RMSD_{norm}$, right) for the (Q)/B contribution to the atomization energies in the Karton test set according to different FNO compression schemes, referenced to full Q cc-pV(5,6)Z values. Results obtained with orbitals selected from MP2-NO occupations are labeled as (MP2), those from CCSD-NO occupations as (CCSD), while calculations performed in the untruncated virtual space are denoted as (full). Projected cc-pVDZ and cc-pVTZ results correspond to the same number of COs as in the respective parent basis sets.

TABLE A.8: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the cc-pV(5,6)Z references. Results are reported for the (T), full T, (Q), and (Q)–(T) contributions obtained with the cc-pVDZ and cc-pVTZ basis sets, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	(T)			full T			(Q)			(Q)-(T)		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVDZ	full		503	6.24 0.01241	7.36 0.01464	14.03 0.02789	1.49 0.00296	2.05 0.00408	4.39 0.00872	0.75 0.00148	1.26 0.00250	3.21 0.00638	0.80 0.00159	1.02 0.00203	2.14 0.00425
cc-pVTZ	$1.0{\cdot}10^{-3}$	MP2	364	10.53 0.02892	12.63 0.03469	24.34 0.06688	1.48 0.00406	2.11 0.00579	5.04 0.01383	0.87 0.00239	1.49 0.00410	4.40 0.01208	0.80 0.00220	1.01 0.00278	2.08 0.00572
		CCSD	383	8.92 0.02330	10.35 0.02702	20.35 0.05315	1.46 0.00382	2.07 0.00541	4.92 0.01285	0.80 0.00209	1.37 0.00359	3.95 0.01032	0.86 0.00224	1.10 0.00287	2.41 0.00630
	$5.0{\cdot}10^{-4}$	MP2	422	6.62 0.01569	7.59 0.01799	14.52 0.03442	1.51 0.00358	2.06 0.00487	4.48 0.01062	0.90 0.00214	1.50 0.00356	4.52 0.01072	0.79 0.00187	0.95 0.00226	1.59 0.00376
		CCSD	445	6.25 0.01404	7.33 0.01647	15.70 0.03528	1.42 0.00320	2.03 0.00457	4.74 0.01064	0.75 0.00169	1.31 0.00295	4.11 0.00924	0.78 0.00175	0.98 0.00221	1.81 0.00407
	projected cc-pVDZ	MP2	503	4.12 0.00820	4.95 0.00983	9.45 0.01879	1.28 0.00255	1.73 0.00344	3.47 0.00690	0.76 0.00151	1.25 0.00248	3.37 0.00671	0.57 0.00114	0.70 0.00139	1.37 0.00271
		CCSD	503	4.31 0.00856	5.13 0.01020	9.58 0.01905	1.30 0.00258	1.74 0.00347	3.58 0.00712	0.72 0.00144	1.20 0.00239	3.20 0.00636	0.60 0.00120	0.72 0.00144	1.51 0.00299
	$2.5{\cdot}10^{-4}$	MP2	531	3.40 0.00640	4.14 0.00780	7.68 0.01446	1.24 0.00233	1.69 0.00319	4.04 0.00760	0.72 0.00135	1.13 0.00213	3.25 0.00613	0.55 0.00104	0.66 0.00125	1.37 0.00258
		CCSD	537	3.60 0.00670	4.57 0.00852	11.38 0.02119	1.29 0.00240	1.75 0.00325	4.23 0.00787	0.71 0.00132	1.14 0.00213	3.41 0.00635	0.58 0.00109	0.72 0.00134	1.61 0.00300
	$1.0{\cdot}10^{-4}$	MP2	733	1.76 0.00240	2.11 0.00288	4.53 0.00618	0.83 0.00113	1.12 0.00153	2.51 0.00343	0.49 0.00067	0.69 0.00094	1.78 0.00242	0.38 0.00052	0.49 0.00067	1.13 0.00154
		CCSD	711	1.54 0.00217	1.88 0.00264	3.77 0.00530	1.02 0.00143	1.43 0.00201	3.32 0.00467	0.51 0.00071	0.73 0.00102	1.88 0.00264	0.54 0.00076	0.76 0.00107	2.20 0.00309
	$5.0{\cdot}10^{-5}$	MP2	891	2.42 0.00271	2.70 0.00303	5.13 0.00575	0.60 0.00068	0.81 0.00090	1.94 0.00218	0.39 0.00043	0.55 0.00062	1.42 0.00159	0.26 0.00029	0.32 0.00036	0.78 0.00087
		CCSD	841	2.07 0.00246	2.46 0.00293	4.37 0.00520	0.78 0.00093	1.06 0.00126	2.61 0.00311	0.40 0.00048	0.57 0.00068	1.42 0.00168	0.38 0.00046	0.55 0.00065	1.42 0.00169
	$2.5{\cdot}10^{-5}$	MP2	1001	2.10 0.00210	2.33 0.00233	3.91 0.00391	0.55 0.00055	0.72 0.00072	1.63 0.00162	0.31 0.00031	0.42 0.00042	0.92 0.00092	0.28 0.00028	0.35 0.00035	0.77 0.00077
		CCSD	961	2.16 0.00225	2.43 0.00253	4.54 0.00473	0.59 0.00061	0.79 0.00083	1.73 0.00180	0.32 0.00033	0.45 0.00047	1.01 0.00105	0.30 0.00031	0.41 0.00043	1.24 0.00129
	$1.0{\cdot}10^{-5}$	MP2	1089	1.90 0.00174	2.19 0.00201	4.12 0.00378	0.53 0.00049	0.72 0.00066	1.63 0.00149	0.29 0.00026	0.41 0.00038	0.92 0.00084	0.26 0.00024	0.35 0.00032	0.81 0.00074
		CCSD	1060	1.89 0.00178	2.19 0.00207	4.18 0.00395	0.52 0.00050	0.70 0.00066	1.60 0.00151	0.28 0.00027	0.41 0.00039	0.91 0.00086	0.25 0.00023	0.32 0.00031	0.72 0.00068
	full		1179	1.87 0.00159	2.19 0.00186	4.23 0.00359	0.53 0.00045	0.70 0.00060	1.58 0.00134	0.28 0.00024	0.41 0.00034	0.91 0.00077	0.25 0.00021	0.34 0.00029	0.80 0.00068

TABLE A.9: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the cc-pV(5,6)Z references. Results are reported for the (T), full T, (Q), and (Q)–(T) contributions obtained with the cc-pVQZ basis set, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	(T)			full T			(Q)			(Q)-(T)		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVQZ	$1.0\cdot10^{-3}$	MP2	384	10.04	11.97	23.95	1.48	2.09	5.02	0.87	1.48	4.36	0.83	1.03	2.10
				0.02615	0.03118	0.06237	0.00385	0.00545	0.01308	0.00227	0.00386	0.01137	0.00215	0.00269	0.00547
		CCSD	401	8.69	10.19	19.86	1.45	2.05	4.89	0.81	1.37	3.97	0.87	1.10	2.39
				0.02167	0.02541	0.04952	0.00363	0.00511	0.01219	0.00201	0.00343	0.00989	0.00216	0.00275	0.00597
	$5.0\cdot10^{-4}$	MP2	454	6.31	7.14	11.68	1.41	1.94	4.11	0.82	1.39	4.27	0.77	0.97	1.89
				0.01390	0.01573	0.02574	0.00311	0.00427	0.00905	0.00181	0.00306	0.00940	0.00169	0.00214	0.00416
		CCSD	464	6.12	7.18	15.14	1.38	2.00	4.69	0.75	1.30	4.14	0.77	0.99	1.83
				0.01320	0.01546	0.03263	0.00298	0.00431	0.01011	0.00162	0.00281	0.00891	0.00165	0.00213	0.00394
	projected cc-pVDZ	MP2	503	3.92	4.71	8.78	1.30	1.74	3.53	0.73	1.23	3.34	0.62	0.74	1.50
				0.00779	0.00937	0.01745	0.00259	0.00347	0.00702	0.00146	0.00244	0.00663	0.00122	0.00147	0.00299
		CCSD	503	4.04	4.90	8.73	1.32	1.78	3.78	0.71	1.19	3.21	0.65	0.78	1.63
				0.00804	0.00973	0.01736	0.00263	0.00354	0.00751	0.00141	0.00236	0.00638	0.00129	0.00155	0.00325
	$2.5\cdot10^{-4}$	MP2	573	2.94	3.47	6.98	1.11	1.55	3.82	0.65	1.03	2.90	0.50	0.63	1.41
				0.00514	0.00605	0.01217	0.00194	0.00270	0.00666	0.00114	0.00179	0.00507	0.00087	0.00110	0.00247
		CCSD	587	2.86	3.83	10.84	1.21	1.66	4.14	0.65	1.08	3.41	0.57	0.74	1.52
				0.00488	0.00653	0.01846	0.00206	0.00282	0.00704	0.00110	0.00184	0.00580	0.00096	0.00125	0.00259
	$1.0\cdot10^{-4}$	MP2	825	1.88	2.21	4.32	0.68	0.93	2.37	0.42	0.62	1.68	0.31	0.39	0.92
				0.00228	0.00268	0.00524	0.00082	0.00113	0.00287	0.00051	0.00075	0.00204	0.00037	0.00047	0.00112
		CCSD	781	1.52	1.82	3.29	0.85	1.19	2.99	0.45	0.66	1.80	0.41	0.57	1.38
				0.00195	0.00233	0.00422	0.00109	0.00152	0.00383	0.00057	0.00085	0.00230	0.00053	0.00073	0.00177
	$5.0\cdot10^{-5}$	MP2	1010	1.91	2.15	4.41	0.49	0.66	1.63	0.29	0.41	1.11	0.23	0.29	0.63
				0.00189	0.00213	0.00437	0.00049	0.00065	0.00162	0.00029	0.00041	0.00110	0.00023	0.00029	0.00062
		CCSD	944	1.67	1.97	3.94	0.60	0.87	2.13	0.31	0.46	1.30	0.31	0.47	1.44
				0.00177	0.00209	0.00417	0.00064	0.00092	0.00226	0.00033	0.00048	0.00138	0.00033	0.00050	0.00152
	projected cc-pVTZ	MP2	1179	1.50	1.83	3.13	0.44	0.56	1.27	0.25	0.37	0.83	0.23	0.28	0.59
				0.00127	0.00155	0.00265	0.00037	0.00048	0.00108	0.00021	0.00031	0.00071	0.00019	0.00024	0.00050
		CCSD	1179	1.48	1.85	3.31	0.42	0.57	1.30	0.23	0.35	0.85	0.24	0.32	0.88
				0.00125	0.00157	0.00281	0.00036	0.00049	0.00110	0.00020	0.00030	0.00072	0.00020	0.00027	0.00074
	$2.5\cdot10^{-5}$	MP2	1208	1.11	1.26	2.40	0.43	0.56	1.38	0.20	0.27	0.64	0.26	0.32	0.74
				0.00092	0.00105	0.00198	0.00036	0.00047	0.00114	0.00016	0.00022	0.00053	0.00022	0.00027	0.00061
		CCSD	1114	1.23	1.41	2.91	0.44	0.62	1.50	0.21	0.30	0.78	0.25	0.36	0.99
				0.00111	0.00127	0.00261	0.00040	0.00055	0.00134	0.00019	0.00027	0.00070	0.00023	0.00032	0.00089
	$1.0\cdot10^{-5}$	MP2	1570	0.80	0.91	1.48	0.34	0.44	0.99	0.14	0.19	0.41	0.21	0.27	0.62
				0.00051	0.00058	0.00095	0.00022	0.00028	0.00063	0.00009	0.00012	0.00026	0.00014	0.00017	0.00040
		CCSD	1419	0.88	1.04	1.96	0.37	0.50	1.19	0.16	0.22	0.49	0.22	0.31	0.81
				0.00062	0.00073	0.00138	0.00026	0.00036	0.00084	0.00011	0.00015	0.00035	0.00016	0.00022	0.00057
	full		2274	0.82	0.98	1.74	0.26	0.35	0.77	0.11	0.15	0.34	0.15	0.21	0.47
				0.00036	0.00043	0.00076	0.00011	0.00015	0.00034	0.00005	0.00007	0.00015	0.00007	0.00009	0.00021

TABLE A.10: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the cc-pV(5,6)Z references. Results are reported for the (T), full T, (Q), and (Q)–(T) contributions obtained with the cc-pV5Z basis set, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	(T)			full T			(Q)			(Q)-(T)		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pV5Z	$1.0\cdot10^{-3}$	MP2	409	9.98	11.88	23.76	1.49	2.10	5.03	0.87	1.48	4.35	0.84	1.05	2.12
				0.02440	0.02905	0.05810	0.00364	0.00514	0.01229	0.00214	0.00362	0.01064	0.00206	0.00257	0.00519
		CCSD	424	8.89	10.56	19.82	1.45	2.05	4.90	0.81	1.37	3.96	0.87	1.09	2.26
				0.02096	0.02489	0.04675	0.00342	0.00483	0.01155	0.00191	0.00323	0.00934	0.00204	0.00257	0.00534
	$5.0\cdot10^{-4}$	MP2	479	6.24	7.05	11.57	1.42	1.94	4.11	0.82	1.39	4.26	0.77	0.98	1.88
				0.01303	0.01472	0.02416	0.00295	0.00406	0.00859	0.00172	0.00290	0.00889	0.00161	0.00205	0.00393
		CCSD	493	5.69	6.83	14.94	1.39	2.01	4.71	0.71	1.29	4.13	0.76	0.99	1.91
				0.01155	0.01385	0.03031	0.00282	0.00407	0.00955	0.00145	0.00262	0.00837	0.00154	0.00201	0.00387
	projected cc-pVDZ	MP2	503	3.73	4.57	8.63	1.32	1.76	3.55	0.74	1.23	3.33	0.62	0.75	1.54
				0.00741	0.00908	0.01715	0.00262	0.00349	0.00706	0.00148	0.00244	0.00662	0.00124	0.00149	0.00307
		CCSD	503	3.97	4.81	8.64	1.33	1.80	3.80	0.71	1.19	3.22	0.67	0.80	1.69
				0.00788	0.00957	0.01718	0.00265	0.00357	0.00755	0.00142	0.00236	0.00640	0.00133	0.00160	0.00336
	$2.5\cdot10^{-4}$	MP2	604	2.57	3.21	6.96	1.12	1.55	3.80	0.66	1.02	2.89	0.49	0.64	1.45
				0.00426	0.00531	0.01153	0.00186	0.00257	0.00629	0.00110	0.00169	0.00478	0.00082	0.00106	0.00240
		CCSD	616	2.71	3.62	9.86	1.20	1.64	4.02	0.65	1.07	3.30	0.55	0.73	1.57
				0.00440	0.00587	0.01600	0.00194	0.00266	0.00652	0.00105	0.00173	0.00535	0.00090	0.00119	0.00255
	$1.0\cdot10^{-4}$	MP2	862	1.80	2.08	4.12	0.66	0.91	2.32	0.40	0.59	1.64	0.29	0.38	0.94
				0.00208	0.00241	0.00478	0.00076	0.00106	0.00270	0.00047	0.00068	0.00190	0.00033	0.00044	0.00109
		CCSD	812	1.55	1.88	3.65	0.82	1.16	2.94	0.44	0.65	1.77	0.39	0.55	1.37
				0.00191	0.00232	0.00449	0.00101	0.00143	0.00363	0.00054	0.00080	0.00218	0.00048	0.00068	0.00169
	$5.0\cdot10^{-5}$	MP2	1066	1.49	1.68	2.96	0.46	0.60	1.44	0.24	0.33	0.76	0.25	0.31	0.67
				0.00140	0.00157	0.00277	0.00043	0.00057	0.00135	0.00023	0.00031	0.00072	0.00023	0.00029	0.00063
		CCSD	991	1.40	1.68	3.84	0.58	0.85	2.09	0.29	0.42	1.19	0.31	0.47	1.42
				0.00141	0.00169	0.00388	0.00058	0.00085	0.00211	0.00029	0.00042	0.00120	0.00031	0.00048	0.00143
	projected cc-pVTZ	MP2	1179	1.13	1.50	3.11	0.44	0.56	1.27	0.23	0.35	0.85	0.24	0.29	0.64
				0.00096	0.00127	0.00264	0.00037	0.00048	0.00108	0.00020	0.00029	0.00072	0.00021	0.00025	0.00054
		CCSD	1179	1.32	1.76	3.30	0.44	0.59	1.34	0.22	0.33	0.85	0.27	0.36	0.99
				0.00112	0.00149	0.00280	0.00037	0.00050	0.00113	0.00018	0.00028	0.00072	0.00023	0.00031	0.00084
	$2.5\cdot10^{-5}$	MP2	1320	0.75	0.87	1.58	0.38	0.50	1.18	0.16	0.22	0.48	0.25	0.30	0.69
				0.00057	0.00066	0.00120	0.00029	0.00038	0.00089	0.00012	0.00017	0.00037	0.00019	0.00023	0.00053
		CCSD	1199	1.07	1.23	2.11	0.43	0.59	1.40	0.19	0.26	0.65	0.26	0.36	1.01
				0.00089	0.00103	0.00176	0.00036	0.00049	0.00117	0.00016	0.00022	0.00054	0.00022	0.00030	0.00084
	$1.0\cdot10^{-5}$	MP2	1800	0.66	0.72	1.05	0.29	0.37	0.82	0.11	0.14	0.30	0.20	0.25	0.53
				0.00037	0.00040	0.00058	0.00016	0.00020	0.00046	0.00006	0.00008	0.00017	0.00011	0.00014	0.00029
		CCSD	1541	0.62	0.74	1.23	0.33	0.45	1.09	0.12	0.17	0.41	0.21	0.29	0.75
				0.00040	0.00048	0.00080	0.00021	0.00029	0.00070	0.00008	0.00011	0.00027	0.00014	0.00019	0.00049
	full		3890	0.38	0.46	0.81	0.14	0.19	0.41	0.05	0.07	0.17	0.09	0.12	0.28
				0.00010	0.00012	0.00021	0.00004	0.00005	0.00010	0.00001	0.00002	0.00004	0.00002	0.00003	0.00007

TABLE A.11: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the cc-pV(5,6)Z references. Results are reported for the full Q and Q–(T) contributions obtained with the cc-pVDZ and cc-pVTZ basis sets, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	full Q			Q-(T)		
				MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVDZ	full		503	0.67 0.00132	1.11 0.00221	2.87 0.00570	0.87 0.00172	1.10 0.00219	2.12 0.00422
cc-pVTZ	$1.0{\cdot}10^{-3}$	MP2	364	0.80 0.00219	1.33 0.00366	3.82 0.01049	0.85 0.00233	1.06 0.00291	2.02 0.00556
		CCSD	383	0.74 0.00194	1.23 0.00321	3.41 0.00891	0.89 0.00232	1.12 0.00293	2.28 0.00596
	$5.0{\cdot}10^{-4}$	MP2	422	0.84 0.00199	1.36 0.00322	3.94 0.00934	0.81 0.00192	0.96 0.00227	1.81 0.00429
		CCSD	445	0.70 0.00157	1.19 0.00268	3.68 0.00827	0.80 0.00180	1.00 0.00225	1.79 0.00401
	projected cc-pVDZ	MP2	503	0.72 0.00142	1.14 0.00226	2.93 0.00583	0.60 0.00118	0.71 0.00142	1.33 0.00265
		CCSD	503	0.69 0.00137	1.11 0.00221	2.89 0.00575	0.62 0.00122	0.74 0.00147	1.46 0.00290
	$2.5{\cdot}10^{-4}$	MP2	531	0.67 0.00126	1.01 0.00191	2.84 0.00534	0.59 0.00111	0.72 0.00135	1.34 0.00251
		CCSD	537	0.66 0.00123	1.03 0.00191	3.03 0.00564	0.62 0.00115	0.79 0.00146	1.91 0.00355
	$1.0{\cdot}10^{-4}$	MP2	733	0.43 0.00059	0.57 0.00078	1.36 0.00186	0.42 0.00058	0.57 0.00078	1.34 0.00183
		CCSD	711	0.45 0.00064	0.62 0.00088	1.56 0.00219	0.57 0.00081	0.85 0.00119	2.45 0.00345
	$5.0{\cdot}10^{-5}$	MP2	891	0.33 0.00037	0.44 0.00050	1.06 0.00119	0.30 0.00033	0.40 0.00045	0.99 0.00111
		CCSD	841	0.35 0.00042	0.48 0.00057	1.17 0.00139	0.41 0.00049	0.62 0.00073	1.61 0.00192
	$2.5{\cdot}10^{-5}$	MP2	1001	0.26 0.00026	0.34 0.00034	0.74 0.00074	0.31 0.00031	0.41 0.00041	0.96 0.00096
		CCSD	961	0.26 0.00027	0.36 0.00038	0.79 0.00082	0.33 0.00034	0.48 0.00050	1.42 0.00148
	$1.0{\cdot}10^{-5}$	MP2	1089	0.23 0.00022	0.33 0.00030	0.74 0.00068	0.29 0.00027	0.41 0.00038	0.99 0.00091
		CCSD	1060	0.23 0.00022	0.33 0.00031	0.74 0.00070	0.28 0.00026	0.39 0.00036	0.92 0.00087
	full		1179	0.23 0.00019	0.32 0.00028	0.73 0.00062	0.29 0.00024	0.40 0.00034	0.98 0.00083

TABLE A.12: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the cc-pV(5,6)Z references. Results are reported for the full Q and Q–(T) contributions obtained with the cc-pVQZ basis set, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	full Q			Q-(T)		
				MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVQZ	$1.0{\cdot}10^{-3}$	MP2	384	0.80	1.33	3.80	0.86	1.06	2.04
				0.00210	0.00346	0.00989	0.00224	0.00277	0.00530
		CCSD	401	0.75	1.23	3.45	0.89	1.11	2.26
				0.00188	0.00308	0.00859	0.00221	0.00278	0.00564
	$5.0{\cdot}10^{-4}$	MP2	454	0.76	1.27	3.81	0.77	0.96	1.88
				0.00168	0.00279	0.00840	0.00169	0.00212	0.00413
		CCSD	464	0.70	1.19	3.72	0.78	0.99	1.80
				0.00151	0.00257	0.00802	0.00169	0.00214	0.00387
	projected cc-pVDZ	MP2	503	0.70	1.12	2.92	0.63	0.75	1.45
				0.00138	0.00222	0.00580	0.00126	0.00148	0.00288
		CCSD	503	0.68	1.11	2.93	0.65	0.78	1.56
				0.00136	0.00220	0.00582	0.00130	0.00155	0.00310
	$2.5{\cdot}10^{-4}$	MP2	573	0.60	0.89	2.41	0.54	0.69	1.41
				0.00105	0.00155	0.00420	0.00094	0.00121	0.00246
		CCSD	587	0.60	0.96	3.03	0.60	0.78	1.78
				0.00103	0.00164	0.00516	0.00102	0.00133	0.00304
	$1.0{\cdot}10^{-4}$	MP2	825	0.36	0.50	1.26	0.34	0.47	1.11
				0.00044	0.00060	0.00153	0.00041	0.00057	0.00135
		CCSD	781	0.39	0.56	1.50	0.45	0.64	1.58
				0.00051	0.00072	0.00192	0.00057	0.00082	0.00203
	$5.0{\cdot}10^{-5}$	MP2	1010	0.23	0.31	0.79	0.27	0.37	0.85
				0.00023	0.00031	0.00078	0.00027	0.00036	0.00084
		CCSD	944	0.26	0.37	1.07	0.34	0.53	1.60
				0.00028	0.00039	0.00113	0.00036	0.00056	0.00170
	projected cc-pVTZ	MP2	1179	0.20	0.28	0.64	0.25	0.33	0.77
				0.00017	0.00024	0.00054	0.00021	0.00028	0.00065
		CCSD	1179	0.19	0.27	0.65	0.26	0.36	1.01
				0.00016	0.00023	0.00055	0.00022	0.00031	0.00086
	$2.5{\cdot}10^{-5}$	MP2	1208	0.15	0.20	0.45	0.29	0.38	0.93
				0.00013	0.00017	0.00037	0.00024	0.00031	0.00077
		CCSD	1114	0.17	0.23	0.62	0.27	0.40	1.14
				0.00015	0.00021	0.00056	0.00025	0.00036	0.00102
	$1.0{\cdot}10^{-5}$	MP2	1570	0.11	0.14	0.29	0.23	0.31	0.74
				0.00007	0.00009	0.00018	0.00014	0.00020	0.00047
		CCSD	1419	0.12	0.16	0.36	0.24	0.35	0.94
				0.00009	0.00011	0.00025	0.00017	0.00025	0.00067
	full		2274	0.08	0.12	0.26	0.15	0.23	0.56
				0.00004	0.00005	0.00012	0.00007	0.00010	0.00025

TABLE A.13: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the full Q cc-pV(5,6)Z references. Results are reported for the (Q), $(Q)_\Lambda$, (Q)/A, and (Q)/B contributions obtained with the cc-pVDZ and cc-pVTZ basis sets, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	(Q)			$(Q)_\Lambda$			(Q)/A			(Q)/B		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVDZ	full		503	0.51 0.00101	0.84 0.00166	1.81 0.00361	0.54 0.00107	0.93 0.00185	2.23 0.00444	0.59 0.00117	1.02 0.00203	2.93 0.00582	0.55 0.00109	0.92 0.00183	2.50 0.00496
cc-pVTZ	$1.0{\cdot}10^{-3}$	MP2	364	0.50 0.00138	0.80 0.00220	1.78 0.00490	0.64 0.00175	1.01 0.00277	2.37 0.00650	0.56 0.00153	0.95 0.00260	2.44 0.00671	0.51 0.00141	0.88 0.00242	2.19 0.00603
		CCSD	383	0.53 0.00140	0.79 0.00206	1.67 0.00435	0.61 0.00159	0.96 0.00252	2.34 0.00612	0.53 0.00139	0.90 0.00236	2.29 0.00599	0.52 0.00137	0.85 0.00221	2.05 0.00536
	$5.0{\cdot}10^{-4}$	MP2	422	0.50 0.00118	0.76 0.00179	1.75 0.00416	0.66 0.00157	1.00 0.00237	2.44 0.00579	0.65 0.00153	0.98 0.00231	2.46 0.00583	0.58 0.00138	0.88 0.00209	2.20 0.00521
		CCSD	445	0.40 0.00090	0.62 0.00140	1.62 0.00364	0.53 0.00119	0.85 0.00192	2.41 0.00541	0.50 0.00113	0.82 0.00184	2.00 0.00449	0.44 0.00098	0.72 0.00163	1.84 0.00414
	projected cc-pVDZ	MP2	503	0.49 0.00097	0.75 0.00149	1.59 0.00316	0.53 0.00106	0.87 0.00172	2.11 0.00420	0.55 0.00109	0.93 0.00184	2.44 0.00486	0.52 0.00103	0.83 0.00164	2.08 0.00414
		CCSD	503	0.48 0.00095	0.75 0.00150	1.61 0.00320	0.51 0.00101	0.85 0.00169	2.14 0.00426	0.53 0.00106	0.90 0.00180	2.34 0.00464	0.50 0.00100	0.81 0.00161	1.99 0.00395
	$2.5{\cdot}10^{-4}$	MP2	531	0.45 0.00084	0.67 0.00127	1.58 0.00298	0.48 0.00090	0.77 0.00145	2.37 0.00447	0.51 0.00096	0.82 0.00155	2.01 0.00379	0.48 0.00090	0.73 0.00137	1.82 0.00343
		CCSD	537	0.42 0.00078	0.63 0.00117	1.51 0.00282	0.47 0.00088	0.76 0.00141	2.34 0.00436	0.50 0.00093	0.79 0.00147	1.93 0.00359	0.46 0.00085	0.70 0.00130	1.75 0.00326
	$1.0{\cdot}10^{-4}$	MP2	733	0.40 0.00055	0.81 0.00111	3.03 0.00414	0.37 0.00051	0.57 0.00078	1.77 0.00241	0.49 0.00066	0.72 0.00098	1.58 0.00216	0.45 0.00061	0.69 0.00094	1.99 0.00271
		CCSD	711	0.41 0.00058	0.81 0.00114	2.93 0.00412	0.37 0.00053	0.61 0.00085	2.00 0.00281	0.49 0.00069	0.73 0.00103	1.57 0.00221	0.45 0.00063	0.70 0.00098	1.90 0.00267
	$5.0{\cdot}10^{-5}$	MP2	891	0.37 0.00041	0.87 0.00098	3.39 0.00381	0.33 0.00037	0.54 0.00060	1.56 0.00175	0.44 0.00050	0.71 0.00080	1.87 0.00210	0.41 0.00046	0.70 0.00078	2.31 0.00259
		CCSD	841	0.38 0.00045	0.88 0.00105	3.39 0.00403	0.33 0.00040	0.56 0.00067	1.72 0.00204	0.45 0.00054	0.73 0.00087	1.89 0.00225	0.41 0.00049	0.71 0.00085	2.31 0.00274
	$2.5{\cdot}10^{-5}$	MP2	1001	0.37 0.00037	1.00 0.00100	3.94 0.00394	0.32 0.00032	0.56 0.00056	1.55 0.00155	0.44 0.00044	0.77 0.00077	2.38 0.00237	0.40 0.00040	0.78 0.00078	2.81 0.00281
		CCSD	961	0.36 0.00038	0.96 0.00100	3.80 0.00396	0.31 0.00032	0.55 0.00057	1.50 0.00156	0.43 0.00045	0.75 0.00078	2.25 0.00234	0.39 0.00041	0.76 0.00079	2.68 0.00279
	$1.0{\cdot}10^{-5}$	MP2	1089	0.36 0.00033	0.99 0.00091	3.94 0.00362	0.30 0.00028	0.55 0.00051	1.55 0.00142	0.43 0.00039	0.76 0.00070	2.37 0.00218	0.39 0.00035	0.78 0.00071	2.81 0.00258
		CCSD	1060	0.36 0.00034	0.99 0.00094	3.93 0.00371	0.30 0.00028	0.55 0.00052	1.54 0.00145	0.43 0.00040	0.76 0.00072	2.37 0.00223	0.38 0.00036	0.77 0.00073	2.80 0.00264
	full		1179	0.35 0.00030	1.00 0.00085	3.96 0.00336	0.30 0.00025	0.55 0.00047	1.56 0.00133	0.42 0.00036	0.76 0.00065	2.39 0.00203	0.38 0.00032	0.78 0.00066	2.82 0.00239

TABLE A.14: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the full Q cc-pV(5,6)Z references. Results are reported for the (Q), $(Q)_\Lambda$, (Q)/A, and (Q)/B contributions obtained with the cc-pVQZ basis set, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	(Q)			$(Q)_\Lambda$			(Q)/A			(Q)/B		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVQZ	$1.0\cdot10^{-3}$	MP2	384	0.51	0.80	1.73	0.64	1.00	2.38	0.56	0.94	2.36	0.52	0.88	2.12
				0.00134	0.00209	0.00451	0.00168	0.00262	0.00621	0.00145	0.00244	0.00614	0.00136	0.00229	0.00553
		CCSD	401	0.55	0.79	1.64	0.62	0.96	2.35	0.54	0.90	2.23	0.54	0.84	1.99
				0.00136	0.00197	0.00409	0.00155	0.00240	0.00586	0.00135	0.00224	0.00555	0.00133	0.00210	0.00497
	$5.0\cdot10^{-4}$	MP2	454	0.45	0.69	1.73	0.59	0.91	2.46	0.57	0.87	2.12	0.50	0.79	1.94
				0.00099	0.00151	0.00381	0.00129	0.00200	0.00543	0.00125	0.00193	0.00468	0.00111	0.00173	0.00427
		CCSD	464	0.39	0.61	1.62	0.53	0.84	2.41	0.50	0.80	2.01	0.43	0.71	1.85
				0.00085	0.00131	0.00350	0.00113	0.00182	0.00520	0.00107	0.00172	0.00433	0.00093	0.00152	0.00398
	projected cc-pVDZ	MP2	503	0.48	0.74	1.55	0.52	0.85	2.12	0.53	0.90	2.34	0.50	0.80	2.00
				0.00095	0.00147	0.00309	0.00103	0.00168	0.00421	0.00106	0.00179	0.00466	0.00100	0.00160	0.00398
		CCSD	503	0.47	0.74	1.60	0.50	0.84	2.14	0.52	0.88	2.24	0.49	0.79	1.91
				0.00093	0.00147	0.00318	0.00100	0.00166	0.00426	0.00104	0.00175	0.00445	0.00098	0.00157	0.00379
	$2.5\cdot10^{-4}$	MP2	573	0.43	0.70	1.91	0.41	0.72	2.39	0.49	0.77	2.02	0.46	0.70	1.85
				0.00075	0.00122	0.00333	0.00072	0.00125	0.00418	0.00085	0.00135	0.00353	0.00080	0.00122	0.00323
		CCSD	587	0.36	0.58	1.51	0.41	0.70	2.34	0.43	0.71	1.94	0.40	0.63	1.76
				0.00061	0.00099	0.00258	0.00070	0.00119	0.00399	0.00074	0.00121	0.00330	0.00068	0.00107	0.00299
	$1.0\cdot10^{-4}$	MP2	825	0.37	0.82	3.12	0.32	0.55	1.78	0.43	0.69	1.67	0.40	0.67	2.09
				0.00045	0.00100	0.00379	0.00039	0.00066	0.00215	0.00053	0.00083	0.00203	0.00049	0.00082	0.00254
		CCSD	781	0.38	0.81	3.01	0.33	0.57	1.95	0.44	0.69	1.57	0.40	0.67	1.97
				0.00048	0.00103	0.00385	0.00042	0.00073	0.00250	0.00056	0.00088	0.00201	0.00052	0.00086	0.00253
	$5.0\cdot10^{-5}$	MP2	1010	0.35	0.94	3.70	0.28	0.54	1.55	0.40	0.70	2.17	0.36	0.72	2.61
				0.00035	0.00093	0.00366	0.00028	0.00053	0.00154	0.00040	0.00069	0.00215	0.00036	0.00072	0.00258
		CCSD	944	0.35	0.90	3.51	0.28	0.53	1.67	0.39	0.67	1.99	0.36	0.69	2.41
				0.00037	0.00095	0.00371	0.00029	0.00056	0.00177	0.00042	0.00071	0.00211	0.00038	0.00073	0.00256
	projected cc-pVTZ	MP2	1179	0.35	1.04	4.11	0.29	0.57	1.71	0.42	0.77	2.53	0.38	0.80	2.97
				0.00030	0.00088	0.00349	0.00025	0.00049	0.00145	0.00035	0.00065	0.00214	0.00032	0.00068	0.00252
		CCSD	1179	0.34	1.04	4.13	0.28	0.56	1.69	0.40	0.76	2.51	0.36	0.79	2.96
				0.00029	0.00088	0.00350	0.00024	0.00048	0.00143	0.00034	0.00064	0.00213	0.00031	0.00067	0.00251
	$2.5\cdot10^{-5}$	MP2	1208	0.38	1.06	4.17	0.27	0.57	1.76	0.39	0.75	2.57	0.36	0.80	3.02
				0.00032	0.00087	0.00345	0.00022	0.00047	0.00146	0.00032	0.00062	0.00213	0.00030	0.00066	0.00250
		CCSD	1114	0.38	1.02	4.03	0.26	0.55	1.60	0.38	0.73	2.43	0.36	0.77	2.87
				0.00034	0.00092	0.00361	0.00023	0.00049	0.00144	0.00034	0.00065	0.00218	0.00032	0.00069	0.00258
	$1.0\cdot10^{-5}$	MP2	1570	0.41	1.12	4.40	0.28	0.60	1.96	0.39	0.78	2.77	0.36	0.84	3.22
				0.00026	0.00071	0.00280	0.00018	0.00038	0.00125	0.00025	0.00050	0.00177	0.00023	0.00054	0.00205
		CCSD	1419	0.40	1.10	4.32	0.27	0.59	1.87	0.38	0.77	2.70	0.35	0.82	3.14
				0.00028	0.00077	0.00304	0.00019	0.00041	0.00132	0.00027	0.00054	0.00190	0.00025	0.00058	0.00221
	full		2274	0.40	1.15	4.51	0.27	0.61	2.04	0.37	0.79	2.86	0.35	0.86	3.31
				0.00018	0.00050	0.00198	0.00012	0.00027	0.00090	0.00016	0.00035	0.00126	0.00015	0.00038	0.00146

TABLE A.15: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the Q-(T) cc-pV(5,6)Z references. Results are reported for the (Q)-(T), $(Q)_\Lambda$-(T), (Q)/A-(T), and (Q)/B–(T) contributions obtained with the cc-pVDZ and cc-pVTZ basis sets, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	(Q)-(T)			$(Q)_\Lambda$-(T)			(Q)/A-(T)			(Q)/B-(T)		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVDZ	full		503	1.25 0.00248	1.94 0.00385	5.98 0.01189	1.02 0.00202	1.50 0.00298	3.93 0.00782	1.03 0.00204	1.62 0.00322	4.89 0.00971	1.10 0.00219	1.71 0.00340	5.18 0.01030
cc-pVTZ	$1.0{\cdot}10^{-3}$	MP2	364	1.26 0.00346	1.83 0.00504	5.45 0.01496	1.07 0.00293	1.46 0.00402	3.68 0.01010	1.13 0.00311	1.63 0.00449	4.67 0.01283	1.18 0.00323	1.70 0.00466	4.89 0.01344
		CCSD	383	1.32 0.00343	1.95 0.00509	5.78 0.01509	1.10 0.00288	1.55 0.00403	3.86 0.01007	1.18 0.00307	1.73 0.00451	4.97 0.01299	1.22 0.00319	1.79 0.00467	5.17 0.01350
	$5.0{\cdot}10^{-4}$	MP2	422	1.24 0.00293	1.68 0.00397	4.77 0.01129	1.04 0.00246	1.30 0.00308	3.00 0.00710	1.05 0.00250	1.39 0.00330	3.72 0.00881	1.12 0.00265	1.48 0.00351	4.01 0.00950
		CCSD	445	1.24 0.00278	1.83 0.00411	5.44 0.01222	1.02 0.00229	1.41 0.00317	3.55 0.00799	1.05 0.00235	1.55 0.00348	4.48 0.01007	1.11 0.00250	1.63 0.00367	4.73 0.01063
	projected cc-pVDZ	MP2	503	1.03 0.00205	1.54 0.00305	4.85 0.00965	0.81 0.00160	1.11 0.00220	2.89 0.00574	0.80 0.00160	1.19 0.00237	3.52 0.00700	0.88 0.00175	1.29 0.00257	3.90 0.00776
		CCSD	503	1.06 0.00211	1.62 0.00321	5.19 0.01032	0.83 0.00164	1.16 0.00230	3.13 0.00622	0.83 0.00165	1.26 0.00251	3.83 0.00762	0.91 0.00181	1.36 0.00271	4.22 0.00838
	$2.5{\cdot}10^{-4}$	MP2	531	1.01 0.00190	1.65 0.00310	5.59 0.01053	0.79 0.00148	1.17 0.00221	3.58 0.00675	0.78 0.00146	1.26 0.00237	4.21 0.00793	0.86 0.00161	1.37 0.00259	4.61 0.00869
		CCSD	537	1.04 0.00193	1.67 0.00312	5.63 0.01048	0.81 0.00151	1.18 0.00219	3.58 0.00666	0.83 0.00154	1.35 0.00252	4.56 0.00850	0.90 0.00168	1.44 0.00269	4.84 0.00902
	$1.0{\cdot}10^{-4}$	MP2	733	0.84 0.00114	1.55 0.00211	5.55 0.00757	0.59 0.00081	1.00 0.00136	3.36 0.00459	0.60 0.00082	1.14 0.00155	4.07 0.00555	0.67 0.00091	1.26 0.00171	4.50 0.00614
		CCSD	711	0.99 0.00139	1.75 0.00247	6.06 0.00852	0.75 0.00105	1.18 0.00166	3.83 0.00538	0.74 0.00104	1.34 0.00188	4.61 0.00649	0.82 0.00116	1.46 0.00206	5.02 0.00707
	$5.0{\cdot}10^{-5}$	MP2	891	0.71 0.00080	1.42 0.00160	5.33 0.00598	0.48 0.00054	0.86 0.00097	3.09 0.00347	0.49 0.00055	1.01 0.00113	3.81 0.00428	0.53 0.00060	1.12 0.00126	4.25 0.00477
		CCSD	841	0.83 0.00099	1.63 0.00194	6.00 0.00714	0.57 0.00068	1.03 0.00123	3.71 0.00441	0.58 0.00069	1.19 0.00142	4.50 0.00535	0.65 0.00077	1.32 0.00157	4.92 0.00585
	$2.5{\cdot}10^{-5}$	MP2	1001	0.73 0.00073	1.48 0.00148	5.57 0.00556	0.49 0.00049	0.89 0.00088	3.18 0.00317	0.50 0.00050	1.05 0.00105	4.00 0.00400	0.54 0.00054	1.17 0.00117	4.43 0.00443
		CCSD	961	0.75 0.00079	1.48 0.00154	5.47 0.00570	0.49 0.00051	0.86 0.00090	3.10 0.00322	0.53 0.00055	1.05 0.00109	3.92 0.00408	0.57 0.00059	1.16 0.00121	4.35 0.00453
	$1.0{\cdot}10^{-5}$	MP2	1089	0.71 0.00066	1.48 0.00136	5.56 0.00511	0.47 0.00044	0.88 0.00081	3.17 0.00291	0.48 0.00044	1.05 0.00096	4.00 0.00367	0.52 0.00048	1.16 0.00107	4.43 0.00407
		CCSD	1060	0.69 0.00066	1.47 0.00139	5.53 0.00522	0.47 0.00044	0.88 0.00083	3.14 0.00297	0.46 0.00043	1.04 0.00098	3.97 0.00374	0.50 0.00047	1.15 0.00109	4.40 0.00415
	full		1179	0.70 0.00059	1.48 0.00125	5.54 0.00470	0.46 0.00039	0.87 0.00074	3.15 0.00267	0.46 0.00039	1.04 0.00088	3.97 0.00337	0.51 0.00043	1.16 0.00098	4.40 0.00374

TABLE A.16: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the Q-(T) cc-pV(5,6)Z references. Results are reported for the (Q)-(T), $(Q)_\Lambda$-(T), (Q)/A-(T), and (Q)/B–(T) contributions obtained with the cc-pVQZ basis set, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	(Q)-(T)			$(Q)_\Lambda$-(T)			(Q)/A-(T)			(Q)/B-(T)		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVQZ	$1.0\cdot10^{-3}$	MP2	384	1.28	1.85	5.47	1.09	1.48	3.69	1.16	1.65	4.69	1.20	1.71	4.91
				0.00334	0.00482	0.01424	0.00284	0.00387	0.00962	0.00301	0.00431	0.01222	0.00313	0.00446	0.01280
		CCSD	401	1.32	1.94	5.73	1.11	1.55	3.80	1.18	1.72	4.91	1.23	1.79	5.11
				0.00330	0.00485	0.01428	0.00276	0.00386	0.00949	0.00295	0.00430	0.01224	0.00306	0.00445	0.01274
	$5.0\cdot10^{-4}$	MP2	454	1.20	1.68	4.65	1.00	1.30	2.82	1.02	1.39	3.52	1.08	1.48	3.84
				0.00265	0.00369	0.01025	0.00220	0.00287	0.00621	0.00225	0.00307	0.00775	0.00238	0.00326	0.00845
		CCSD	464	1.22	1.82	5.36	1.01	1.41	3.48	1.04	1.54	4.39	1.10	1.63	4.65
				0.00264	0.00392	0.01156	0.00217	0.00304	0.00751	0.00223	0.00333	0.00946	0.00237	0.00350	0.01002
	projected cc-pVDZ	MP2	503	1.07	1.59	5.00	0.85	1.17	3.04	0.85	1.25	3.67	0.93	1.35	4.05
				0.00213	0.00316	0.00995	0.00169	0.00232	0.00604	0.00169	0.00248	0.00729	0.00184	0.00268	0.00806
		CCSD	503	1.10	1.68	5.37	0.88	1.23	3.31	0.88	1.33	4.00	0.96	1.43	4.39
				0.00219	0.00335	0.01068	0.00174	0.00245	0.00658	0.00174	0.00264	0.00795	0.00190	0.00285	0.00873
	$2.5\cdot10^{-4}$	MP2	573	0.95	1.65	5.72	0.73	1.17	3.68	0.73	1.29	4.41	0.81	1.39	4.78
				0.00166	0.00288	0.00999	0.00128	0.00204	0.00643	0.00127	0.00224	0.00769	0.00141	0.00243	0.00835
		CCSD	587	1.02	1.66	5.54	0.79	1.18	3.49	0.81	1.35	4.47	0.88	1.44	4.75
				0.00174	0.00284	0.00943	0.00135	0.00201	0.00594	0.00138	0.00230	0.00762	0.00150	0.00245	0.00810
	$1.0\cdot10^{-4}$	MP2	825	0.74	1.47	5.49	0.50	0.93	3.31	0.52	1.07	4.04	0.57	1.19	4.46
				0.00090	0.00179	0.00666	0.00061	0.00112	0.00401	0.00063	0.00130	0.00490	0.00069	0.00144	0.00541
		CCSD	781	0.87	1.65	6.00	0.62	1.07	3.75	0.61	1.23	4.56	0.70	1.35	4.96
				0.00111	0.00211	0.00768	0.00080	0.00137	0.00480	0.00078	0.00157	0.00583	0.00090	0.00172	0.00635
	$5.0\cdot10^{-5}$	MP2	1010	0.69	1.41	5.33	0.48	0.85	3.06	0.47	1.00	3.80	0.51	1.11	4.24
				0.00068	0.00140	0.00528	0.00047	0.00084	0.00303	0.00046	0.00099	0.00377	0.00050	0.00110	0.00420
		CCSD	944	0.76	1.52	5.61	0.50	0.91	3.28	0.53	1.09	4.09	0.58	1.21	4.51
				0.00081	0.00161	0.00594	0.00053	0.00096	0.00348	0.00056	0.00115	0.00434	0.00061	0.00128	0.00478
	projected cc-pVTZ	MP2	1179	0.66	1.41	5.38	0.45	0.82	2.98	0.45	1.00	3.79	0.49	1.11	4.24
				0.00056	0.00120	0.00456	0.00038	0.00070	0.00253	0.00038	0.00085	0.00322	0.00042	0.00094	0.00360
		CCSD	1179	0.66	1.43	5.42	0.42	0.81	2.99	0.44	1.00	3.81	0.48	1.11	4.26
				0.00056	0.00121	0.00460	0.00035	0.00069	0.00253	0.00037	0.00085	0.00323	0.00041	0.00094	0.00361
	$2.5\cdot10^{-5}$	MP2	1208	0.72	1.47	5.55	0.50	0.88	3.14	0.48	1.04	3.95	0.53	1.15	4.40
				0.00059	0.00122	0.00459	0.00041	0.00072	0.00260	0.00040	0.00086	0.00327	0.00044	0.00095	0.00364
		CCSD	1114	0.70	1.46	5.53	0.43	0.84	3.10	0.46	1.03	3.93	0.51	1.14	4.37
				0.00063	0.00131	0.00496	0.00039	0.00076	0.00278	0.00042	0.00092	0.00353	0.00046	0.00103	0.00393
	$1.0\cdot10^{-5}$	MP2	0.67	1.42	5.39	0.45	0.82	2.95	0.44	0.99	3.76	0.48	1.10	4.21	
				0.00043	0.00091	0.00343	0.00029	0.00052	0.00188	0.00028	0.00063	0.00240	0.00030	0.00070	0.00268
		CCSD	1419	0.68	1.45	5.51	0.44	0.83	3.06	0.44	1.01	3.89	0.49	1.13	4.33
				0.00048	0.00102	0.00388	0.00031	0.00059	0.00216	0.00031	0.00071	0.00274	0.00034	0.00079	0.00305
	full		2274	0.60	1.39	5.28	0.40	0.78	2.81	0.39	0.95	3.64	0.41	1.06	4.08
				0.00026	0.00061	0.00232	0.00018	0.00034	0.00124	0.00017	0.00042	0.00160	0.00018	0.00047	0.00180

TABLE A.17: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the full basis references. Results are reported for the (T), full T, (Q), and (Q)–(T) contributions obtained with the cc-pVTZ basis set, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	(T)			full T			(Q)			(Q)-(T)		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVTZ	$1.0{\cdot}10^{-3}$	MP2	364	8.66	10.55	20.93	0.98	1.42	3.45	0.66	1.14	3.55	0.58	0.76	1.59
				0.02379	0.02898	0.05749	0.00269	0.00389	0.00948	0.00181	0.00313	0.00975	0.00158	0.00210	0.00436
		CCSD	383	7.06	8.22	16.13	0.96	1.38	3.34	0.62	1.03	3.10	0.62	0.84	1.92
				0.01842	0.02147	0.04211	0.00252	0.00361	0.00872	0.00163	0.00269	0.00810	0.00162	0.00220	0.00500
	$5.0{\cdot}10^{-4}$	MP2	422	4.75	5.57	10.30	1.01	1.37	2.90	0.68	1.14	3.67	0.64	0.75	1.36
				0.01126	0.01319	0.02440	0.00239	0.00324	0.00686	0.00160	0.00271	0.00871	0.00152	0.00177	0.00322
		CCSD	445	4.45	5.33	11.47	0.93	1.35	3.15	0.53	0.96	3.26	0.57	0.74	1.67
				0.01000	0.01198	0.02578	0.00210	0.00303	0.00708	0.00119	0.00217	0.00732	0.00127	0.00167	0.00375
	projected cc-pVDZ	MP2	503	2.36	2.81	5.22	0.79	1.05	2.42	0.51	0.86	2.52	0.43	0.52	1.00
				0.00469	0.00559	0.01038	0.00157	0.00209	0.00480	0.00102	0.00171	0.00501	0.00085	0.00104	0.00199
		CCSD	503	2.46	2.99	5.35	0.81	1.06	2.42	0.48	0.81	2.35	0.42	0.52	1.14
				0.00488	0.00595	0.01064	0.00161	0.00212	0.00481	0.00096	0.00162	0.00467	0.00084	0.00103	0.00227
	$2.5{\cdot}10^{-4}$	MP2	531	1.76	2.18	4.40	0.73	1.00	2.45	0.44	0.76	2.40	0.31	0.40	1.01
				0.00331	0.00410	0.00829	0.00137	0.00188	0.00462	0.00083	0.00144	0.00453	0.00059	0.00075	0.00189
		CCSD	537	1.89	2.53	7.15	0.77	1.06	2.64	0.43	0.78	2.56	0.35	0.43	0.95
				0.00352	0.00470	0.01332	0.00143	0.00197	0.00492	0.00080	0.00145	0.00476	0.00066	0.00081	0.00176
	$1.0{\cdot}10^{-4}$	MP2	733	0.87	1.18	3.01	0.32	0.44	0.93	0.21	0.31	0.92	0.15	0.23	0.51
				0.00119	0.00160	0.00410	0.00044	0.00060	0.00127	0.00029	0.00042	0.00126	0.00021	0.00032	0.00069
		CCSD	711	0.85	1.09	2.80	0.50	0.77	2.14	0.23	0.36	1.03	0.30	0.46	1.40
				0.00119	0.00154	0.00394	0.00071	0.00108	0.00301	0.00032	0.00050	0.00144	0.00042	0.00064	0.00196
	$5.0{\cdot}10^{-5}$	MP2	891	0.60	0.68	1.26	0.10	0.15	0.37	0.11	0.17	0.57	0.08	0.14	0.49
				0.00068	0.00076	0.00141	0.00011	0.00017	0.00041	0.00012	0.00019	0.00064	0.00009	0.00016	0.00055
		CCSD	841	0.70	0.84	1.35	0.25	0.40	1.03	0.12	0.19	0.57	0.16	0.23	0.62
				0.00084	0.00100	0.00160	0.00030	0.00048	0.00122	0.00015	0.00022	0.00067	0.00019	0.00028	0.00074
	$2.5{\cdot}10^{-5}$	MP2	1001	0.28	0.36	1.04	0.05	0.10	0.39	0.03	0.04	0.10	0.06	0.13	0.49
				0.00028	0.00036	0.00104	0.00005	0.00010	0.00039	0.00003	0.00004	0.00010	0.00006	0.00013	0.00049
		CCSD	961	0.37	0.44	1.05	0.09	0.16	0.56	0.04	0.05	0.15	0.08	0.14	0.44
				0.00039	0.00046	0.00109	0.00009	0.00017	0.00058	0.00004	0.00006	0.00016	0.00009	0.00014	0.00046
	$1.0{\cdot}10^{-5}$	MP2	1089	0.08	0.10	0.17	0.02	0.04	0.12	0.01	0.01	0.03	0.02	0.04	0.14
				0.00007	0.00009	0.00015	0.00002	0.00003	0.00011	0.00001	0.00001	0.00002	0.00002	0.00004	0.00013
		CCSD	1060	0.10	0.13	0.37	0.02	0.04	0.12	0.01	0.01	0.03	0.02	0.04	0.13
				0.00009	0.00012	0.00035	0.00002	0.00003	0.00011	0.00001	0.00001	0.00002	0.00002	0.00004	0.00013

TABLE A.18: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the full basis references. Results are reported for the (T), full T, (Q), and (Q)–(T) contributions obtained with the cc-pVQZ basis set, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	(T)			full T			(Q)			(Q)-(T)		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVQZ	$1.0\cdot10^{-3}$	MP2	384	9.23	11.06	22.21	1.24	1.75	4.25	0.79	1.35	4.07	0.69	0.89	1.84
				0.02403	0.02880	0.05785	0.00322	0.00456	0.01107	0.00206	0.00352	0.01059	0.00178	0.00231	0.00478
		CCSD	401	7.88	9.26	18.12	1.21	1.71	4.11	0.74	1.24	3.67	0.72	0.95	2.13
				0.01964	0.02309	0.04518	0.00302	0.00427	0.01026	0.00185	0.00310	0.00915	0.00181	0.00237	0.00531
	$5.0\cdot10^{-4}$	MP2	454	5.70	6.32	10.60	1.17	1.60	3.34	0.74	1.26	3.97	0.68	0.85	1.68
				0.01256	0.01392	0.02336	0.00257	0.00353	0.00735	0.00163	0.00277	0.00874	0.00151	0.00188	0.00370
		CCSD	464	5.31	6.31	13.40	1.14	1.66	3.92	0.67	1.18	3.84	0.63	0.84	1.62
				0.01144	0.01360	0.02888	0.00246	0.00358	0.00845	0.00144	0.00253	0.00827	0.00136	0.00180	0.00350
	projected cc-pVDZ	MP2	503	3.10	3.75	7.04	1.06	1.41	2.89	0.65	1.08	3.04	0.51	0.61	1.28
				0.00616	0.00746	0.01400	0.00210	0.00280	0.00574	0.00128	0.00215	0.00604	0.00101	0.00122	0.00254
		CCSD	503	3.23	3.94	6.99	1.08	1.44	3.00	0.62	1.04	2.91	0.51	0.64	1.40
				0.00643	0.00783	0.01391	0.00215	0.00287	0.00597	0.00124	0.00207	0.00579	0.00102	0.00126	0.00279
	$2.5\cdot10^{-4}$	MP2	573	2.14	2.60	5.89	0.87	1.21	3.04	0.55	0.89	2.60	0.36	0.47	1.18
				0.00373	0.00454	0.01027	0.00151	0.00212	0.00531	0.00095	0.00155	0.00454	0.00063	0.00082	0.00207
		CCSD	587	2.11	2.99	9.10	0.96	1.33	3.36	0.54	0.95	3.11	0.43	0.57	1.29
				0.00359	0.00510	0.01550	0.00163	0.00226	0.00573	0.00092	0.00162	0.00529	0.00073	0.00097	0.00220
	$1.0\cdot10^{-4}$	MP2	825	1.07	1.34	2.59	0.43	0.60	1.59	0.32	0.48	1.38	0.19	0.24	0.48
				0.00130	0.00162	0.00313	0.00052	0.00073	0.00193	0.00038	0.00058	0.00168	0.00023	0.00029	0.00058
		CCSD	781	0.87	1.09	2.21	0.59	0.86	2.22	0.34	0.52	1.50	0.28	0.38	0.94
				0.00112	0.00139	0.00284	0.00076	0.00110	0.00284	0.00043	0.00067	0.00192	0.00036	0.00049	0.00120
	$5.0\cdot10^{-5}$	MP2	1010	1.09	1.27	2.67	0.25	0.33	0.86	0.18	0.27	0.81	0.12	0.16	0.50
				0.00108	0.00126	0.00264	0.00025	0.00033	0.00085	0.00018	0.00027	0.00081	0.00012	0.00016	0.00049
		CCSD	944	0.90	1.12	2.52	0.35	0.56	1.55	0.20	0.32	1.00	0.19	0.30	0.99
				0.00096	0.00119	0.00267	0.00037	0.00059	0.00164	0.00022	0.00034	0.00106	0.00021	0.00032	0.00105
	projected cc-pVTZ	MP2	1179	0.69	0.91	1.63	0.19	0.24	0.50	0.14	0.21	0.49	0.12	0.16	0.41
				0.00058	0.00077	0.00138	0.00016	0.00020	0.00042	0.00012	0.00018	0.00042	0.00010	0.00014	0.00035
		CCSD	1179	0.68	0.90	1.71	0.18	0.24	0.58	0.13	0.20	0.50	0.13	0.18	0.46
				0.00057	0.00077	0.00145	0.00015	0.00021	0.00049	0.00011	0.00017	0.00043	0.00011	0.00015	0.00039
	$2.5\cdot10^{-5}$	MP2	1208	0.33	0.42	0.92	0.19	0.24	0.61	0.09	0.13	0.34	0.13	0.16	0.44
				0.00028	0.00035	0.00077	0.00016	0.00020	0.00050	0.00007	0.00011	0.00028	0.00011	0.00013	0.00036
		CCSD	1114	0.53	0.60	1.17	0.21	0.31	0.83	0.11	0.16	0.48	0.14	0.19	0.54
				0.00047	0.00054	0.00105	0.00019	0.00028	0.00074	0.00009	0.00014	0.00043	0.00012	0.00017	
	$1.0\cdot10^{-5}$	MP2	1570	0.23	0.27	0.46	0.10	0.13	0.31	0.03	0.05	0.14	0.07	0.11	0.35
				0.00014	0.00017	0.00029	0.00006	0.00008	0.00020	0.00002	0.00003	0.00009	0.00005	0.00007	0.00022
		CCSD	1419	0.22	0.28	0.58	0.12	0.18	0.50	0.05	0.07	0.19	0.09	0.12	0.36
				0.00016	0.00020	0.00041	0.00009	0.00013	0.00035	0.00003	0.00005	0.00014	0.00006	0.00009	0.00026

TABLE A.19: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the full basis references. Results are reported for the (T), full T, (Q), and (Q)–(T) contributions obtained with the cc-pV5Z basis set, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	(T)			full T			(Q)			(Q)-(T)		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pV5Z	$1.0{\cdot}10^{-3}$	MP2	409	9.59	11.45	22.96	1.35	1.92	4.62	0.83	1.42	4.23	0.76	0.96	1.98
				0.02346	0.02800	0.05614	0.00331	0.00469	0.01131	0.00204	0.00347	0.01034	0.00185	0.00236	0.00484
		CCSD	424	8.50	10.12	19.34	1.31	1.86	4.49	0.77	1.31	3.83	0.78	1.00	2.13
				0.02005	0.02388	0.04560	0.00310	0.00439	0.01059	0.00182	0.00309	0.00904	0.00184	0.00236	0.00503
	$5.0{\cdot}10^{-4}$	MP2	479	5.96	6.66	11.10	1.28	1.76	3.71	0.78	1.33	4.14	0.72	0.91	1.71
				0.01244	0.01390	0.02317	0.00267	0.00367	0.00775	0.00164	0.00278	0.00863	0.00151	0.00189	0.00357
		CCSD	493	5.41	6.43	14.14	1.26	1.82	4.30	0.67	1.23	4.00	0.68	0.89	1.73
				0.01098	0.01305	0.02869	0.00256	0.00369	0.00873	0.00137	0.00250	0.00812	0.00138	0.00181	0.00352
	projected cc-pVDZ	MP2	503	3.34	4.11	7.83	1.18	1.57	3.15	0.70	1.16	3.21	0.55	0.67	1.41
				0.00665	0.00818	0.01556	0.00235	0.00312	0.00625	0.00139	0.00231	0.00638	0.00109	0.00133	0.00281
		CCSD	503	3.58	4.36	7.84	1.20	1.61	3.39	0.67	1.12	3.09	0.58	0.71	1.56
				0.00712	0.00867	0.01558	0.00238	0.00320	0.00675	0.00134	0.00222	0.00615	0.00116	0.00141	0.00310
	$2.5{\cdot}10^{-4}$	MP2	604	2.21	2.81	6.46	0.98	1.37	3.39	0.61	0.96	2.76	0.41	0.55	1.32
				0.00367	0.00464	0.01069	0.00163	0.00227	0.00562	0.00101	0.00159	0.00458	0.00068	0.00090	0.00218
		CCSD	616	2.36	3.20	9.05	1.06	1.46	3.61	0.60	1.01	3.17	0.49	0.64	1.44
				0.00383	0.00520	0.01470	0.00172	0.00237	0.00586	0.00097	0.00164	0.00515	0.00079	0.00104	0.00234
	$1.0{\cdot}10^{-4}$	MP2	862	1.41	1.67	3.32	0.52	0.73	1.92	0.35	0.52	1.52	0.22	0.28	0.69
				0.00164	0.00194	0.00385	0.00060	0.00085	0.00223	0.00041	0.00061	0.00176	0.00025	0.00033	0.00080
		CCSD	812	1.24	1.52	3.17	0.68	0.98	2.54	0.38	0.59	1.65	0.32	0.44	1.12
				0.00152	0.00187	0.00390	0.00084	0.00121	0.00313	0.00047	0.00072	0.00203	0.00039	0.00055	0.00138
	$5.0{\cdot}10^{-5}$	MP2	1066	1.11	1.29	2.45	0.32	0.43	1.03	0.19	0.27	0.64	0.18	0.22	0.47
				0.00104	0.00121	0.00230	0.00030	0.00040	0.00097	0.00018	0.00025	0.00060	0.00017	0.00021	0.00044
		CCSD	991	1.02	1.27	3.15	0.44	0.68	1.79	0.24	0.36	1.06	0.24	0.37	1.17
				0.00103	0.00128	0.00317	0.00044	0.00068	0.00180	0.00024	0.00036	0.00107	0.00025	0.00038	0.00118
	projected cc-pVTZ	MP2	1179	0.80	1.06	2.30	0.30	0.38	0.87	0.18	0.28	0.68	0.18	0.21	0.39
				0.00067	0.00090	0.00195	0.00026	0.00032	0.00074	0.00015	0.00023	0.00058	0.00015	0.00018	0.00033
		CCSD	1179	0.97	1.33	2.61	0.30	0.41	0.95	0.17	0.26	0.68	0.20	0.27	0.74
				0.00082	0.00112	0.00221	0.00026	0.00035	0.00080	0.00014	0.00022	0.00058	0.00017	0.00023	0.00063
	$2.5{\cdot}10^{-5}$	MP2	1320	0.42	0.49	1.08	0.25	0.32	0.77	0.11	0.15	0.36	0.16	0.20	0.42
				0.00031	0.00037	0.00082	0.00019	0.00024	0.00059	0.00008	0.00012	0.00027	0.00012	0.00015	0.00031
		CCSD	1199	0.69	0.79	1.38	0.29	0.42	1.10	0.14	0.20	0.52	0.18	0.25	0.76
				0.00057	0.00066	0.00115	0.00024	0.00035	0.00092	0.00012	0.00016	0.00043	0.00015	0.00021	0.00063
	$1.0{\cdot}10^{-5}$	MP2	1800	0.28	0.32	0.73	0.16	0.20	0.42	0.06	0.08	0.17	0.13	0.15	0.32
				0.00015	0.00018	0.00040	0.00009	0.00011	0.00023	0.00003	0.00004	0.00010	0.00007	0.00008	0.00018
		CCSD	1541	0.26	0.32	0.71	0.19	0.27	0.68	0.07	0.10	0.28	0.13	0.18	0.50
				0.00017	0.00021	0.00046	0.00012	0.00018	0.00044	0.00005	0.00007	0.00018	0.00008	0.00012	0.00033

TABLE A.20: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the full basis references. Results are reported for the full Q, Q-(T), and (P) contributions obtained with the cc-pVTZ basis set, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

				full Q			Q-(T)			(P)		
			n_{CO}	MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVTZ	$1.0{\cdot}10^{-3}$	MP2	364	0.63	1.05	3.17	0.58	0.74	1.52	0.08	0.15	0.44
				0.00172	0.00289	0.00871	0.00159	0.00204	0.00418	0.00021	0.00041	0.00120
		CCSD	383	0.60	0.95	2.76	0.62	0.80	1.78	0.09	0.18	0.48
				0.00157	0.00249	0.00721	0.00161	0.00209	0.00464	0.00025	0.00047	0.00126
	$5.0{\cdot}10^{-4}$	MP2	422	0.67	1.07	3.29	0.59	0.69	1.28	0.08	0.15	0.41
				0.00159	0.00255	0.00780	0.00139	0.00163	0.00303	0.00019	0.00035	0.00097
		CCSD	445	0.53	0.92	3.03	0.54	0.70	1.59	0.08	0.15	0.51
				0.00119	0.00206	0.00680	0.00121	0.00157	0.00357	0.00018	0.00034	0.00114
	projected cc-pVDZ	MP2	503	0.51	0.83	2.28	0.39	0.47	0.94	0.05	0.09	0.23
				0.00101	0.00164	0.00454	0.00077	0.00094	0.00186	0.00011	0.00017	0.00046
		CCSD	503	0.48	0.80	2.24	0.38	0.47	1.03	0.06	0.10	0.28
				0.00096	0.00160	0.00446	0.00077	0.00094	0.00204	0.00012	0.00020	0.00056
	$2.5{\cdot}10^{-4}$	MP2	531	0.45	0.72	2.18	0.32	0.38	0.90	0.04	0.07	0.16
				0.00084	0.00136	0.00411	0.00060	0.00072	0.00170	0.00008	0.00013	0.00031
		CCSD	537	0.44	0.74	2.38	0.35	0.43	0.92	0.07	0.14	0.49
				0.00081	0.00138	0.00443	0.00066	0.00080	0.00172	0.00014	0.00027	0.00091
	$1.0{\cdot}10^{-4}$	MP2	733	0.20	0.27	0.71	0.17	0.24	0.51	0.01	0.02	0.06
				0.00028	0.00037	0.00097	0.00023	0.00033	0.00070	0.00002	0.00002	0.00008
		CCSD	711	0.23	0.33	0.91	0.31	0.48	1.47	0.02	0.03	0.09
				0.00032	0.00047	0.00128	0.00043	0.00067	0.00207	0.00003	0.00005	0.00013
	$5.0{\cdot}10^{-5}$	MP2	891	0.10	0.14	0.41	0.07	0.14	0.51	0.01	0.01	0.03
				0.00011	0.00015	0.00046	0.00008	0.00016	0.00057	0.00001	0.00001	0.00003
		CCSD	841	0.13	0.18	0.52	0.16	0.24	0.63	0.01	0.01	0.03
				0.00015	0.00022	0.00061	0.00019	0.00028	0.00075	0.00001	0.00002	0.00004
	$2.5{\cdot}10^{-5}$	MP2	1001	0.03	0.04	0.12	0.06	0.13	0.51	0.00	0.00	0.01
				0.00003	0.00004	0.00012	0.00006	0.00013	0.00051	0.00000	0.00000	0.00001
		CCSD	961	0.04	0.05	0.14	0.08	0.14	0.44	0.00	0.01	0.01
				0.00004	0.00005	0.00014	0.00008	0.00015	0.00046	0.00000	0.00001	0.00001
	$1.0{\cdot}10^{-5}$	MP2	1089	0.01	0.01	0.03	0.02	0.04	0.15	0.00	0.00	0.00
				0.00001	0.00001	0.00003	0.00002	0.00004	0.00014	0.00000	0.00000	0.00000
		CCSD	1060	0.01	0.01	0.03	0.02	0.04	0.13	0.00	0.00	0.00
				0.00001	0.00001	0.00002	0.00002	0.00004	0.00013	0.00000	0.00000	0.00000

TABLE A.21: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the full basis references. Results are reported for the full Q, Q-(T), and (P) contributions obtained with the cc-pVQZ basis set, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	full Q			Q-(T)			(P)		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVQZ	$1.0\cdot10^{-3}$	MP2	384	0.75	1.23	3.59	0.72	0.89	1.77	0.08	0.15	0.45
				0.00194	0.00321	0.00936	0.00187	0.00231	0.00462	0.00021	0.00038	0.00116
		CCSD	401	0.70	1.14	3.24	0.74	0.94	2.00	0.09	0.17	0.46
				0.00173	0.00284	0.00807	0.00185	0.00233	0.00498	0.00023	0.00043	0.00115
	$5.0\cdot10^{-4}$	MP2	454	0.70	1.17	3.61	0.66	0.82	1.71	0.08	0.14	0.41
				0.00155	0.00257	0.00794	0.00145	0.00181	0.00378	0.00017	0.00031	0.00090
		CCSD	464	0.64	1.09	3.51	0.64	0.82	1.63	0.08	0.15	0.46
				0.00138	0.00236	0.00757	0.00138	0.00176	0.00352	0.00017	0.00031	0.00100
	projected cc-pVDZ	MP2	503	0.62	1.01	2.71	0.49	0.59	1.22	0.06	0.09	0.25
				0.00123	0.00200	0.00538	0.00097	0.00117	0.00242	0.00012	0.00018	0.00050
		CCSD	503	0.60	1.00	2.72	0.51	0.62	1.33	0.07	0.10	0.29
				0.00120	0.00198	0.00541	0.00101	0.00122	0.00264	0.00013	0.00020	0.00058
	$2.5\cdot10^{-4}$	MP2	573	0.52	0.79	2.20	0.39	0.50	1.14	0.05	0.07	0.18
				0.00090	0.00138	0.00384	0.00069	0.00086	0.00198	0.00009	0.00013	0.00031
		CCSD	587	0.52	0.87	2.82	0.45	0.59	1.25	0.07	0.14	0.47
				0.00088	0.00148	0.00481	0.00077	0.00100	0.00212	0.00013	0.00024	0.00080
	$1.0\cdot10^{-4}$	MP2	825	0.28	0.39	1.06	0.22	0.28	0.58	0.02	0.03	0.07
				0.00034	0.00047	0.00128	0.00026	0.00034	0.00070	0.00002	0.00003	0.00008
		CCSD	781	0.31	0.46	1.29	0.30	0.42	1.05	0.03	0.04	0.09
				0.00040	0.00059	0.00166	0.00038	0.00054	0.00134	0.00004	0.00005	0.00011
	$5.0\cdot10^{-5}$	MP2	1010	0.15	0.21	0.58	0.14	0.19	0.51	0.01	0.02	0.04
				0.00015	0.00021	0.00058	0.00014	0.00018	0.00051	0.00001	0.00002	0.00004
		CCSD	944	0.18	0.28	0.86	0.21	0.33	1.06	0.01	0.02	0.04
				0.00019	0.00029	0.00091	0.00022	0.00035	0.00113	0.00001	0.00002	0.00004
	projected cc-pVTZ	MP2	1179	0.12	0.17	0.37	0.13	0.16	0.42	0.01	0.02	0.04
				0.00010	0.00014	0.00032	0.00011	0.00014	0.00036	0.00001	0.00002	0.00004
		CCSD	1179	0.10	0.16	0.39	0.13	0.18	0.48	0.01	0.02	0.03
				0.00009	0.00013	0.00033	0.00011	0.00015	0.00040	0.00001	0.00001	0.00003
	$2.5\cdot10^{-5}$	MP2	1208	0.07	0.10	0.24	0.14	0.18	0.45	0.01	0.01	0.03
				0.00006	0.00008	0.00020	0.00012	0.00015	0.00038	0.00000	0.00001	0.00003
		CCSD	1114	0.09	0.13	0.41	0.15	0.20	0.60	0.01	0.01	0.04
				0.00008	0.00012	0.00037	0.00013	0.00018	0.00054	0.00001	0.00001	0.00003
	$1.0\cdot10^{-5}$	MP2	1570	0.03	0.04	0.11	0.08	0.12	0.36	0.00	0.00	0.01
				0.00002	0.00003	0.00007	0.00005	0.00008	0.00023	0.00000	0.00000	0.00001
		CCSD	1419	0.04	0.05	0.15	0.10	0.14	0.41	0.00	0.01	0.02
				0.00003	0.00004	0.00011	0.00007	0.00010	0.00029	0.00000	0.00000	0.00001

TABLE A.22: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the full Q full basis set references. Results are reported for the (Q), $(Q)_\Lambda$, (Q)/A and (Q)/B contributions obtained with the cc-pVDZ and cc-pVTZ basis sets, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	(Q)			$(Q)_\Lambda$			(Q)/A			(Q)/B		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVDZ	full		503	0.44 0.00087	1.15 0.00228	4.46 0.00887	0.32 0.00063	0.69 0.00138	2.41 0.00480	0.39 0.00077	0.90 0.00179	3.37 0.00669	0.36 0.00073	0.95 0.00188	3.66 0.00728
cc-pVTZ	$1.0{\cdot}10^{-3}$	MP2	364	0.49 0.00135	0.68 0.00186	1.48 0.00407	0.53 0.00144	0.80 0.00219	2.17 0.00597	0.49 0.00135	0.75 0.00207	1.82 0.00500	0.49 0.00134	0.71 0.00195	1.59 0.00436
		CCSD	383	0.53 0.00138	0.72 0.00188	1.51 0.00393	0.52 0.00136	0.77 0.00202	2.15 0.00562	0.52 0.00137	0.75 0.00196	1.67 0.00437	0.52 0.00135	0.71 0.00187	1.52 0.00396
	$5.0{\cdot}10^{-4}$	MP2	422	0.42 0.00099	0.61 0.00145	1.51 0.00357	0.48 0.00115	0.77 0.00182	2.25 0.00534	0.47 0.00111	0.75 0.00178	1.87 0.00443	0.43 0.00101	0.67 0.00159	1.72 0.00407
		CCSD	445	0.37 0.00083	0.56 0.00126	1.43 0.00321	0.38 0.00086	0.65 0.00147	2.21 0.00497	0.40 0.00090	0.63 0.00142	1.80 0.00405	0.37 0.00083	0.56 0.00127	1.65 0.00370
	projected cc-pVDZ	MP2	503	0.40 0.00079	0.67 0.00133	2.09 0.00415	0.37 0.00073	0.64 0.00126	1.92 0.00382	0.45 0.00089	0.71 0.00142	1.82 0.00362	0.42 0.00083	0.64 0.00128	1.46 0.00290
		CCSD	503	0.40 0.00080	0.70 0.00139	2.26 0.00449	0.36 0.00072	0.63 0.00126	1.95 0.00388	0.45 0.00089	0.71 0.00140	1.71 0.00340	0.42 0.00084	0.65 0.00129	1.36 0.00271
	$2.5{\cdot}10^{-4}$	MP2	531	0.32 0.00060	0.67 0.00126	2.21 0.00415	0.28 0.00053	0.59 0.00111	2.18 0.00411	0.37 0.00069	0.65 0.00122	1.80 0.00339	0.34 0.00064	0.60 0.00113	1.63 0.00306
		CCSD	537	0.30 0.00055	0.62 0.00116	2.05 0.00382	0.25 0.00047	0.57 0.00106	2.15 0.00400	0.36 0.00067	0.63 0.00117	1.73 0.00323	0.33 0.00061	0.58 0.00107	1.56 0.00291
	$1.0{\cdot}10^{-4}$	MP2	733	0.37 0.00051	0.95 0.00129	3.68 0.00503	0.27 0.00036	0.55 0.00075	1.57 0.00215	0.38 0.00052	0.69 0.00094	2.21 0.00301	0.36 0.00049	0.73 0.00099	2.64 0.00360
		CCSD	711	0.38 0.00053	0.93 0.00131	3.58 0.00504	0.26 0.00037	0.57 0.00080	1.81 0.00254	0.38 0.00054	0.70 0.00098	2.14 0.00301	0.36 0.00051	0.73 0.00102	2.55 0.00358
	$5.0{\cdot}10^{-5}$	MP2	891	0.41 0.00046	1.04 0.00117	4.04 0.00454	0.28 0.00032	0.58 0.00065	1.80 0.00202	0.37 0.00042	0.73 0.00082	2.52 0.00283	0.35 0.00040	0.78 0.00088	2.96 0.00332
		CCSD	841	0.41 0.00049	1.04 0.00124	4.04 0.00481	0.28 0.00034	0.59 0.00070	1.74 0.00207	0.38 0.00045	0.74 0.00088	2.54 0.00302	0.35 0.00042	0.79 0.00094	2.96 0.00352
	$2.5{\cdot}10^{-5}$	MP2	1001	0.45 0.00045	1.18 0.00118	4.59 0.00459	0.32 0.00032	0.66 0.00066	2.20 0.00220	0.39 0.00039	0.83 0.00083	3.03 0.00302	0.37 0.00036	0.90 0.00090	3.46 0.00346
		CCSD	961	0.44 0.00046	1.14 0.00119	4.45 0.00464	0.31 0.00032	0.63 0.00066	2.08 0.00216	0.38 0.00039	0.80 0.00083	2.90 0.00302	0.36 0.00037	0.87 0.00090	3.33 0.00347
	$1.0{\cdot}10^{-5}$	MP2	1089	0.45 0.00041	1.18 0.00108	4.59 0.00421	0.31 0.00029	0.65 0.00060	2.20 0.00202	0.38 0.00035	0.82 0.00075	3.02 0.00278	0.35 0.00033	0.89 0.00082	3.46 0.00317
		CCSD	1060	0.45 0.00042	1.18 0.00111	4.58 0.00432	0.31 0.00029	0.65 0.00061	2.19 0.00207	0.38 0.00035	0.82 0.00077	3.02 0.00285	0.35 0.00033	0.89 0.00084	3.45 0.00325
	full		1179	0.45 0.00038	1.18 0.00100	4.61 0.00391	0.31 0.00027	0.65 0.00055	2.21 0.00188	0.37 0.00032	0.82 0.00070	3.04 0.00258	0.35 0.00030	0.90 0.00076	3.47 0.00294

TABLE A.23: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the full Q full basis set references. Results are reported for the (Q), $(Q)_\Lambda$, (Q)/A and (Q)/B contributions obtained with the cc-pVQZ basis set, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	(Q)			$(Q)_\Lambda$			(Q)/A			(Q)/B		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVQZ	$1.0\cdot10^{-3}$	MP2	384	0.51	0.75	1.66	0.59	0.93	2.34	0.52	0.86	2.11	0.50	0.81	1.87
				0.00132	0.00195	0.00432	0.00155	0.00241	0.00609	0.00136	0.00225	0.00550	0.00131	0.00211	0.00488
		CCSD	401	0.54	0.75	1.53	0.59	0.89	2.30	0.54	0.83	1.98	0.53	0.79	1.75
				0.00136	0.00187	0.00381	0.00148	0.00223	0.00574	0.00134	0.00208	0.00494	0.00133	0.00196	0.00436
	$5.0\cdot10^{-4}$	MP2	454	0.41	0.63	1.68	0.52	0.83	2.42	0.50	0.79	2.05	0.44	0.71	1.89
				0.00091	0.00140	0.00370	0.00115	0.00183	0.00532	0.00111	0.00175	0.00451	0.00097	0.00157	0.00417
		CCSD	464	0.38	0.57	1.58	0.46	0.77	2.36	0.44	0.72	1.96	0.40	0.64	1.80
				0.00083	0.00123	0.00340	0.00100	0.00166	0.00509	0.00096	0.00156	0.00423	0.00086	0.00138	0.00388
	projected cc-pVDZ	MP2	503	0.44	0.69	1.68	0.45	0.76	2.07	0.49	0.81	2.10	0.46	0.73	1.75
				0.00087	0.00138	0.00334	0.00090	0.00151	0.00411	0.00098	0.00162	0.00417	0.00092	0.00145	0.00349
		CCSD	503	0.44	0.70	1.81	0.44	0.75	2.09	0.48	0.80	1.99	0.46	0.72	1.66
				0.00088	0.00140	0.00359	0.00087	0.00150	0.00416	0.00096	0.00159	0.00396	0.00091	0.00143	0.00330
	$2.5\cdot10^{-4}$	MP2	573	0.37	0.69	2.11	0.34	0.66	2.34	0.43	0.71	1.98	0.40	0.66	1.80
				0.00065	0.00121	0.00369	0.00059	0.00114	0.00409	0.00075	0.00124	0.00345	0.00070	0.00114	0.00314
		CCSD	587	0.31	0.57	1.61	0.33	0.63	2.29	0.38	0.65	1.89	0.34	0.58	1.71
				0.00052	0.00097	0.00274	0.00056	0.00108	0.00391	0.00064	0.00111	0.00321	0.00058	0.00098	0.00291
	$1.0\cdot10^{-4}$	MP2	825	0.34	0.86	3.33	0.29	0.54	1.73	0.40	0.67	1.88	0.36	0.68	2.30
				0.00042	0.00105	0.00404	0.00035	0.00065	0.00210	0.00048	0.00081	0.00228	0.00044	0.00082	0.00279
		CCSD	781	0.34	0.85	3.22	0.29	0.55	1.90	0.40	0.67	1.77	0.37	0.67	2.18
				0.00044	0.00108	0.00412	0.00037	0.00071	0.00243	0.00051	0.00086	0.00227	0.00047	0.00086	0.00279
	$5.0\cdot10^{-5}$	MP2	1010	0.39	1.00	3.90	0.26	0.57	1.63	0.38	0.71	2.38	0.36	0.76	2.81
				0.00039	0.00099	0.00386	0.00026	0.00056	0.00161	0.00038	0.00070	0.00235	0.00035	0.00075	0.00279
		CCSD	944	0.38	0.96	3.71	0.25	0.54	1.62	0.37	0.68	2.20	0.35	0.72	2.62
				0.00041	0.00101	0.00393	0.00027	0.00058	0.00172	0.00039	0.00072	0.00233	0.00037	0.00076	0.00277
	projected cc-pVTZ	MP2	1179	0.37	1.09	4.32	0.27	0.60	1.92	0.38	0.78	2.73	0.35	0.83	3.18
				0.00031	0.00093	0.00366	0.00023	0.00050	0.00163	0.00032	0.00066	0.00232	0.00029	0.00071	0.00270
		CCSD	1179	0.37	1.10	4.33	0.27	0.59	1.90	0.37	0.77	2.72	0.34	0.83	3.17
				0.00032	0.00093	0.00367	0.00023	0.00050	0.00161	0.00031	0.00065	0.00231	0.00028	0.00070	0.00269
	$2.5\cdot10^{-5}$	MP2	1208	0.43	1.12	4.37	0.30	0.61	1.97	0.38	0.78	2.78	0.36	0.84	3.22
				0.00036	0.00093	0.00362	0.00025	0.00051	0.00163	0.00032	0.00064	0.00230	0.00030	0.00070	0.00267
		CCSD	1114	0.42	1.09	4.23	0.29	0.59	1.81	0.38	0.75	2.64	0.35	0.81	3.08
				0.00038	0.00098	0.00380	0.00026	0.00053	0.00162	0.00034	0.00067	0.00237	0.00032	0.00073	0.00277
	$1.0\cdot10^{-5}$	MP2	1570	0.46	1.19	4.61	0.33	0.66	2.17	0.38	0.81	2.98	0.36	0.89	3.43
				0.00029	0.00076	0.00294	0.00021	0.00042	0.00138	0.00024	0.00052	0.00190	0.00023	0.00057	0.00219
		CCSD	1419	0.45	1.16	4.52	0.31	0.64	2.08	0.38	0.80	2.90	0.35	0.87	3.35
				0.00031	0.00082	0.00319	0.00022	0.00045	0.00146	0.00027	0.00056	0.00205	0.00025	0.00061	0.00236
	full		2274	0.46	1.22	4.72	0.33	0.66	2.24	0.37	0.83	3.07	0.35	0.91	3.52
				0.00020	0.00053	0.00207	0.00014	0.00029	0.00099	0.00016	0.00036	0.00135	0.00016	0.00040	0.00155

TABLE A.24: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the Q-(T) full basis set references. Results are reported for the (Q)-(T), $(Q)_\Lambda$-(T), (Q)/A-(T) and (Q)/B-(T) contributions obtained with the cc-pVDZ and cc-pVTZ basis sets, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	(Q)-(T)			$(Q)_\Lambda$-(T)			(Q)/A-(T)			(Q)/B-(T)		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVDZ	full		503	0.44 0.00087	1.15 0.00228	4.46 0.00887	0.32 0.00063	0.69 0.00138	2.42 0.00480	0.39 0.00077	0.90 0.00179	3.37 0.00669	0.36 0.00072	0.95 0.00188	3.66 0.00728
cc-pVTZ	$1.0{\cdot}10^{-3}$	MP2	364	0.99 0.00272	1.49 0.00410	4.52 0.01241	0.83 0.00228	1.15 0.00315	2.75 0.00755	0.86 0.00237	1.30 0.00358	3.74 0.01028	0.91 0.00249	1.36 0.00373	3.96 0.01089
		CCSD	383	1.05 0.00273	1.61 0.00421	4.85 0.01266	0.90 0.00235	1.24 0.00324	2.93 0.00764	0.91 0.00236	1.40 0.00366	4.05 0.01057	0.95 0.00249	1.46 0.00381	4.24 0.01107
	$5.0{\cdot}10^{-4}$	MP2	422	0.97 0.00229	1.33 0.00316	3.84 0.00909	0.79 0.00187	0.99 0.00236	2.07 0.00491	0.79 0.00186	1.07 0.00255	2.79 0.00661	0.85 0.00201	1.15 0.00272	3.08 0.00730
		CCSD	445	0.97 0.00219	1.49 0.00335	4.51 0.01013	0.80 0.00179	1.11 0.00248	2.63 0.00590	0.79 0.00177	1.23 0.00277	3.55 0.00799	0.85 0.00192	1.30 0.00293	3.80 0.00855
	projected cc-pVDZ	MP2	503	0.78 0.00154	1.22 0.00242	3.92 0.00780	0.66 0.00132	0.84 0.00167	1.96 0.00390	0.65 0.00129	0.91 0.00182	2.59 0.00515	0.66 0.00132	0.99 0.00197	2.98 0.00592
		CCSD	503	0.80 0.00159	1.30 0.00258	4.26 0.00848	0.68 0.00135	0.89 0.00177	2.20 0.00438	0.66 0.00131	0.98 0.00195	2.90 0.00577	0.69 0.00136	1.06 0.00211	3.29 0.00654
	$2.5{\cdot}10^{-4}$	MP2	531	0.74 0.00139	1.32 0.00249	4.66 0.00878	0.61 0.00115	0.88 0.00166	2.66 0.00500	0.55 0.00103	0.96 0.00180	3.28 0.00618	0.61 0.00115	1.06 0.00200	3.69 0.00694
		CCSD	537	0.78 0.00144	1.33 0.00248	4.70 0.00875	0.57 0.00106	0.85 0.00158	2.65 0.00494	0.57 0.00107	1.03 0.00192	3.64 0.00677	0.64 0.00119	1.11 0.00207	3.92 0.00729
	$1.0{\cdot}10^{-4}$	MP2	733	0.57 0.00078	1.24 0.00169	4.62 0.00630	0.43 0.00058	0.74 0.00100	2.44 0.00332	0.45 0.00062	0.88 0.00120	3.14 0.00429	0.44 0.00060	0.97 0.00132	3.57 0.00488
		CCSD	711	0.72 0.00102	1.41 0.00198	5.13 0.00722	0.48 0.00067	0.84 0.00118	2.90 0.00408	0.51 0.00072	1.02 0.00143	3.69 0.00519	0.55 0.00078	1.13 0.00158	4.10 0.00576
	$5.0{\cdot}10^{-5}$	MP2	891	0.48 0.00053	1.14 0.00128	4.40 0.00494	0.35 0.00039	0.65 0.00073	2.16 0.00243	0.41 0.00046	0.80 0.00090	2.88 0.00323	0.39 0.00044	0.87 0.00098	3.32 0.00373
		CCSD	841	0.56 0.00067	1.31 0.00156	5.08 0.00604	0.37 0.00044	0.74 0.00088	2.78 0.00330	0.37 0.00044	0.92 0.00109	3.57 0.00425	0.39 0.00047	1.02 0.00121	3.99 0.00475
	$2.5{\cdot}10^{-5}$	MP2	1001	0.49 0.00049	1.20 0.00120	4.64 0.00464	0.35 0.00035	0.68 0.00067	2.25 0.00225	0.42 0.00042	0.84 0.00084	3.07 0.00307	0.40 0.00039	0.92 0.00092	3.51 0.00350
		CCSD	961	0.48 0.00050	1.17 0.00122	4.55 0.00473	0.31 0.00033	0.60 0.00063	2.17 0.00226	0.38 0.00040	0.80 0.00084	2.99 0.00311	0.36 0.00038	0.88 0.00092	3.42 0.00356
	$1.0{\cdot}10^{-5}$	MP2	1089	0.47 0.00043	1.19 0.00109	4.64 0.00426	0.33 0.00030	0.66 0.00061	2.25 0.00206	0.39 0.00036	0.83 0.00076	3.07 0.00282	0.37 0.00034	0.91 0.00083	3.50 0.00322
		CCSD	1060	0.46 0.00044	1.18 0.00112	4.60 0.00434	0.33 0.00031	0.67 0.00063	2.22 0.00209	0.39 0.00037	0.83 0.00079	3.04 0.00287	0.37 0.00035	0.90 0.00085	3.47 0.00327
	full		1179	0.45 0.00038	1.18 0.00100	4.61 0.00391	0.31 0.00027	0.65 0.00056	2.22 0.00188	0.37 0.00032	0.82 0.00070	3.05 0.00258	0.35 0.00030	0.90 0.00076	3.48 0.00295

TABLE A.25: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the Q-(T) full basis set references. Results are reported for the (Q)-(T), $(Q)_\Lambda$-(T), (Q)/A-(T) and (Q)/B-(T) contributions obtained with the cc-pVQZ basis set, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	(Q)-(T)			$(Q)_\Lambda$-(T)			(Q)/A-(T)			(Q)/B-(T)		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVQZ	$1.0{\cdot}10^{-3}$	MP2	384	1.14	1.66	4.90	0.94	1.30	3.13	1.01	1.46	4.13	1.06	1.52	4.35
				0.00296	0.00431	0.01277	0.00246	0.00340	0.00816	0.00264	0.00381	0.01075	0.00275	0.00396	0.01133
		CCSD	401	1.18	1.75	5.17	0.99	1.37	3.24	1.04	1.53	4.35	1.08	1.59	4.55
				0.00294	0.00436	0.01288	0.00246	0.00342	0.00808	0.00259	0.00383	0.01084	0.00270	0.00398	0.01134
	$5.0{\cdot}10^{-4}$	MP2	454	1.06	1.48	4.09	0.86	1.13	2.26	0.88	1.21	2.96	0.94	1.29	3.27
				0.00233	0.00326	0.00901	0.00188	0.00249	0.00497	0.00195	0.00267	0.00651	0.00207	0.00285	0.00721
		CCSD	464	1.08	1.62	4.80	0.87	1.23	2.92	0.89	1.35	3.83	0.96	1.43	4.09
				0.00233	0.00350	0.01035	0.00187	0.00265	0.00629	0.00193	0.00292	0.00825	0.00207	0.00309	0.00881
	projected cc-pVDZ	MP2	503	0.93	1.40	4.44	0.77	1.00	2.48	0.73	1.08	3.10	0.78	1.17	3.49
				0.00184	0.00279	0.00883	0.00154	0.00200	0.00492	0.00145	0.00214	0.00617	0.00155	0.00232	0.00694
		CCSD	503	0.96	1.49	4.81	0.80	1.07	2.74	0.75	1.15	3.44	0.81	1.25	3.83
				0.00191	0.00297	0.00957	0.00159	0.00212	0.00546	0.00150	0.00229	0.00683	0.00161	0.00248	0.00761
	$2.5{\cdot}10^{-4}$	MP2	573	0.81	1.46	5.16	0.64	1.00	3.12	0.61	1.11	3.84	0.66	1.21	4.22
				0.00141	0.00255	0.00900	0.00112	0.00174	0.00544	0.00106	0.00194	0.00671	0.00116	0.00211	0.00737
		CCSD	587	0.88	1.46	4.97	0.65	0.99	2.92	0.68	1.16	3.91	0.74	1.24	4.19
				0.00149	0.00249	0.00847	0.00111	0.00168	0.00498	0.00116	0.00198	0.00666	0.00126	0.00212	0.00714
	$1.0{\cdot}10^{-4}$	MP2	825	0.60	1.29	4.93	0.42	0.77	2.74	0.41	0.92	3.48	0.44	1.02	3.90
				0.00072	0.00157	0.00597	0.00051	0.00093	0.00333	0.00050	0.00111	0.00421	0.00053	0.00123	0.00472
		CCSD	781	0.72	1.45	5.44	0.48	0.89	3.18	0.49	1.05	3.99	0.55	1.16	4.40
				0.00092	0.00186	0.00696	0.00062	0.00114	0.00408	0.00062	0.00134	0.00511	0.00071	0.00148	0.00563
	$5.0{\cdot}10^{-5}$	MP2	1010	0.55	1.24	4.77	0.40	0.71	2.49	0.42	0.86	3.24	0.41	0.96	3.68
				0.00054	0.00123	0.00472	0.00040	0.00071	0.00247	0.00041	0.00086	0.00321	0.00040	0.00095	0.00364
		CCSD	944	0.62	1.33	5.04	0.37	0.73	2.72	0.41	0.92	3.53	0.45	1.03	3.95
				0.00065	0.00141	0.00534	0.00039	0.00077	0.00288	0.00044	0.00098	0.00374	0.00047	0.00109	0.00419
	projected cc-pVTZ	MP2	1179	0.52	1.24	4.82	0.38	0.70	2.41	0.42	0.87	3.23	0.40	0.95	3.68
				0.00044	0.00105	0.00409	0.00032	0.00059	0.00205	0.00035	0.00073	0.00274	0.00034	0.00081	0.00312
		CCSD	1179	0.52	1.25	4.86	0.35	0.67	2.42	0.38	0.85	3.25	0.36	0.95	3.70
				0.00044	0.00106	0.00412	0.00030	0.00057	0.00206	0.00032	0.00072	0.00275	0.00031	0.00080	0.00314
	$2.5{\cdot}10^{-5}$	MP2	1208	0.57	1.30	4.98	0.42	0.74	2.58	0.44	0.90	3.39	0.43	0.99	3.83
				0.00047	0.00107	0.00412	0.00035	0.00061	0.00213	0.00036	0.00074	0.00280	0.00035	0.00082	0.00317
		CCSD	1114	0.55	1.28	4.96	0.36	0.69	2.54	0.38	0.88	3.36	0.39	0.97	3.81
				0.00050	0.00115	0.00445	0.00032	0.00062	0.00228	0.00034	0.00079	0.00302	0.00035	0.00087	0.00342
	$1.0{\cdot}10^{-5}$	MP2	1570	0.53	1.25	4.83	0.38	0.69	2.39	0.41	0.85	3.20	0.39	0.94	3.65
				0.00034	0.00080	0.00307	0.00024	0.00044	0.00152	0.00026	0.00054	0.00204	0.00025	0.00060	0.00232
		CCSD	1419	0.53	1.28	4.94	0.36	0.69	2.50	0.39	0.87	3.32	0.37	0.97	3.77
				0.00038	0.00090	0.00348	0.00025	0.00049	0.00176	0.00027	0.00061	0.00234	0.00026	0.00068	0.00266
	full		2274	0.46	1.22	4.72	0.33	0.66	2.25	0.37	0.83	3.07	0.35	0.91	3.52
				0.00020	0.00053	0.00208	0.00014	0.00029	0.00099	0.00016	0.00036	0.00135	0.00016	0.00040	0.00155

TABLE A.26: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the full basis set references. Results are reported for the $(Q)_\Lambda$, (Q)/A, and (Q)/B contributions obtained with the cc-pVTZ basis set, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	$(Q)_\Lambda$			Q/A			Q/B		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVTZ	$1.0\cdot10^{-3}$	MP2	364	0.59	0.98	2.92	0.57	0.91	2.75	0.60	0.98	2.96
				0.00162	0.00268	0.00803	0.00157	0.00251	0.00756	0.00165	0.00269	0.00814
		CCSD	383	0.57	0.91	2.63	0.54	0.82	2.34	0.57	0.89	2.57
				0.00148	0.00237	0.00687	0.00141	0.00215	0.00610	0.00148	0.00232	0.00672
	$5.0\cdot10^{-4}$	MP2	422	0.60	0.98	3.05	0.58	0.97	3.15	0.61	1.02	3.29
				0.00143	0.00233	0.00722	0.00137	0.00230	0.00747	0.00145	0.00243	0.00780
		CCSD	445	0.47	0.83	2.75	0.45	0.79	2.64	0.47	0.85	2.83
				0.00106	0.00186	0.00617	0.00101	0.00179	0.00594	0.00106	0.00190	0.00635
	projected cc-pVDZ	MP2	503	0.47	0.76	2.09	0.46	0.76	2.29	0.47	0.79	2.33
				0.00093	0.00151	0.00416	0.00091	0.00152	0.00455	0.00094	0.00157	0.00464
		CCSD	503	0.44	0.73	2.02	0.42	0.72	2.14	0.44	0.75	2.19
				0.00088	0.00146	0.00401	0.00084	0.00143	0.00426	0.00088	0.00149	0.00435
	$2.5\cdot10^{-4}$	MP2	531	0.39	0.65	2.01	0.38	0.68	2.21	0.40	0.70	2.24
				0.00074	0.00123	0.00379	0.00071	0.00127	0.00417	0.00075	0.00132	0.00422
		CCSD	537	0.39	0.68	2.21	0.35	0.62	2.05	0.38	0.67	2.20
				0.00073	0.00126	0.00412	0.00065	0.00116	0.00382	0.00070	0.00125	0.00410
	$1.0\cdot10^{-4}$	MP2	733	0.18	0.25	0.71	0.18	0.27	0.83	0.19	0.28	0.83
				0.00025	0.00035	0.00097	0.00025	0.00037	0.00113	0.00026	0.00038	0.00114
		CCSD	711	0.20	0.30	0.86	0.19	0.30	0.90	0.20	0.32	0.92
				0.00029	0.00042	0.00121	0.00027	0.00042	0.00127	0.00029	0.00045	0.00130
	$5.0\cdot10^{-5}$	MP2	891	0.09	0.13	0.41	0.10	0.15	0.52	0.10	0.15	0.51
				0.00010	0.00015	0.00046	0.00011	0.00017	0.00058	0.00011	0.00017	0.00058
		CCSD	841	0.11	0.16	0.47	0.11	0.16	0.50	0.11	0.17	0.51
				0.00013	0.00019	0.00056	0.00013	0.00020	0.00060	0.00013	0.00020	0.00061
	$2.5\cdot10^{-5}$	MP2	1001	0.03	0.03	0.10	0.02	0.03	0.10	0.02	0.04	0.10
				0.00003	0.00003	0.00010	0.00002	0.00003	0.00010	0.00002	0.00004	0.00010
		CCSD	961	0.03	0.05	0.14	0.03	0.05	0.14	0.03	0.05	0.14
				0.00004	0.00005	0.00014	0.00003	0.00005	0.00014	0.00003	0.00005	0.00014
	$1.0\cdot10^{-5}$	MP2	1089	0.01	0.01	0.02	0.01	0.01	0.03	0.01	0.01	0.03
				0.00001	0.00001	0.00002	0.00001	0.00001	0.00002	0.00001	0.00001	0.00002
		CCSD	1060	0.01	0.01	0.02	0.01	0.01	0.02	0.01	0.01	0.02
				0.00001	0.00001	0.00002	0.00001	0.00001	0.00002	0.00001	0.00001	0.00002

TABLE A.27: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the full basis set references. Results are reported for the $(Q)_\Lambda$, (Q)/A, and (Q)/B contributions obtained with the cc-pVQZ basis set, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	$(Q)_\Lambda$			Q/A			Q/B		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVQZ	$1.0\cdot10^{-3}$	MP2	384	0.70	1.17	3.37	0.69	1.10	3.19	0.72	1.17	3.42
				0.00183	0.00305	0.00877	0.00179	0.00286	0.00832	0.00186	0.00305	0.00891
		CCSD	401	0.67	1.10	3.12	0.65	1.01	2.84	0.68	1.09	3.09
				0.00168	0.00275	0.00778	0.00162	0.00253	0.00708	0.00169	0.00271	0.00770
	$5.0\cdot10^{-4}$	MP2	454	0.66	1.09	3.33	0.64	1.08	3.45	0.67	1.13	3.58
				0.00146	0.00240	0.00733	0.00140	0.00238	0.00760	0.00147	0.00249	0.00790
		CCSD	464	0.60	1.02	3.25	0.57	0.99	3.16	0.60	1.04	3.35
				0.00130	0.00220	0.00699	0.00122	0.00213	0.00682	0.00129	0.00225	0.00723
	projected cc-pVDZ	MP2	503	0.59	0.96	2.53	0.58	0.96	2.73	0.60	1.00	2.79
				0.00117	0.00192	0.00503	0.00115	0.00192	0.00543	0.00119	0.00198	0.00555
		CCSD	503	0.58	0.95	2.51	0.56	0.93	2.64	0.58	0.96	2.70
				0.00115	0.00188	0.00498	0.00111	0.00185	0.00524	0.00115	0.00192	0.00536
	$2.5\cdot10^{-4}$	MP2	573	0.49	0.76	2.17	0.47	0.76	2.27	0.49	0.80	2.34
				0.00086	0.00133	0.00379	0.00081	0.00133	0.00396	0.00086	0.00139	0.00409
		CCSD	587	0.49	0.83	2.69	0.44	0.78	2.53	0.47	0.83	2.69
				0.00084	0.00141	0.00458	0.00075	0.00132	0.00430	0.00081	0.00141	0.00459
	$1.0\cdot10^{-4}$	MP2	825	0.28	0.40	1.10	0.27	0.41	1.19	0.28	0.42	1.22
				0.00034	0.00049	0.00133	0.00033	0.00050	0.00144	0.00034	0.00051	0.00148
		CCSD	781	0.31	0.46	1.28	0.29	0.45	1.30	0.30	0.47	1.34
				0.00039	0.00058	0.00164	0.00037	0.00057	0.00166	0.00039	0.00060	0.00172
	$5.0\cdot10^{-5}$	MP2	1010	0.16	0.22	0.62	0.16	0.23	0.69	0.16	0.24	0.71
				0.00016	0.00022	0.00061	0.00016	0.00023	0.00069	0.00016	0.00024	0.00070
		CCSD	944	0.19	0.28	0.85	0.18	0.28	0.87	0.18	0.29	0.90
				0.00020	0.00030	0.00090	0.00019	0.00029	0.00092	0.00019	0.00031	0.00095
	projected cc-pVTZ	MP2	1179	0.13	0.20	0.46	0.12	0.19	0.47	0.13	0.19	0.48
				0.00011	0.00017	0.00039	0.00011	0.00016	0.00040	0.00011	0.00016	0.00040
		CCSD	1179	0.12	0.19	0.48	0.11	0.18	0.48	0.11	0.18	0.49
				0.00010	0.00016	0.00041	0.00009	0.00015	0.00041	0.00010	0.00015	0.00042
	$2.5\cdot10^{-5}$	MP2	1208	0.08	0.11	0.28	0.08	0.11	0.29	0.08	0.11	0.30
				0.00007	0.00009	0.00023	0.00006	0.00009	0.00024	0.00007	0.00009	0.00024
		CCSD	1114	0.10	0.14	0.43	0.09	0.14	0.43	0.10	0.14	0.44
				0.00009	0.00013	0.00039	0.00008	0.00013	0.00039	0.00009	0.00013	0.00039
	$1.0\cdot10^{-5}$	MP2	1570	0.03	0.04	0.10	0.03	0.04	0.12	0.03	0.04	0.12
				0.00002	0.00003	0.00007	0.00002	0.00003	0.00007	0.00002	0.00003	0.00008
		CCSD	1419	0.04	0.06	0.17	0.04	0.06	0.17	0.04	0.06	0.17
				0.00003	0.00004	0.00012	0.00003	0.00004	0.00012	0.00003	0.00004	0.00012

TABLE A.28: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the full basis set references. Results are reported for the $(Q)_\Lambda$-(T), (Q)/A-(T), and (Q)/B-(T) contributions obtained with the cc-pVTZ basis set, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	$(Q)_\Lambda$-(T)			(Q)/A-(T)			(Q)/B-(T)		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVTZ	$1.0\cdot10^{-3}$	MP2	364	0.64	0.81	1.65	0.73	0.94	2.12	0.67	0.87	1.94
				0.00175	0.00223	0.00453	0.00201	0.00258	0.00582	0.00185	0.00239	0.00532
		CCSD	383	0.67	0.87	1.93	0.77	1.01	2.34	0.72	0.95	2.20
				0.00176	0.00228	0.00504	0.00202	0.00264	0.00610	0.00188	0.00248	0.00574
	$5.0\cdot10^{-4}$	MP2	422	0.63	0.75	1.43	0.68	0.81	1.46	0.66	0.78	1.43
				0.00149	0.00179	0.00339	0.00162	0.00191	0.00346	0.00158	0.00184	0.00339
		CCSD	445	0.61	0.77	1.72	0.66	0.84	1.74	0.62	0.80	1.72
				0.00136	0.00174	0.00385	0.00148	0.00188	0.00392	0.00140	0.00180	0.00387
	projected cc-pVDZ	MP2	503	0.42	0.51	1.02	0.46	0.56	1.06	0.45	0.54	1.04
				0.00083	0.00101	0.00202	0.00092	0.00111	0.00210	0.00089	0.00108	0.00207
		CCSD	503	0.41	0.52	1.11	0.46	0.57	1.21	0.45	0.55	1.19
				0.00082	0.00103	0.00222	0.00092	0.00113	0.00240	0.00089	0.00109	0.00236
	$2.5\cdot10^{-4}$	MP2	531	0.36	0.43	0.98	0.38	0.47	1.06	0.36	0.44	1.04
				0.00068	0.00081	0.00185	0.00071	0.00088	0.00200	0.00067	0.00083	0.00197
		CCSD	537	0.40	0.48	1.06	0.44	0.53	1.00	0.41	0.49	0.98
				0.00074	0.00090	0.00198	0.00082	0.00098	0.00186	0.00076	0.00092	0.00183
	$1.0\cdot10^{-4}$	MP2	733	0.17	0.25	0.50	0.18	0.25	0.53	0.17	0.24	0.52
				0.00024	0.00034	0.00069	0.00024	0.00034	0.00072	0.00023	0.00033	0.00071
		CCSD	711	0.32	0.50	1.55	0.33	0.50	1.51	0.32	0.49	1.46
				0.00046	0.00071	0.00218	0.00047	0.00071	0.00212	0.00046	0.00069	0.00206
	$5.0\cdot10^{-5}$	MP2	891	0.07	0.13	0.49	0.08	0.14	0.49	0.08	0.14	0.49
				0.00008	0.00015	0.00055	0.00008	0.00015	0.00055	0.00008	0.00015	0.00055
		CCSD	841	0.17	0.26	0.69	0.17	0.26	0.67	0.17	0.25	0.66
				0.00020	0.00031	0.00082	0.00020	0.00030	0.00080	0.00020	0.00030	0.00078
	$2.5\cdot10^{-5}$	MP2	1001	0.05	0.13	0.49	0.05	0.13	0.49	0.05	0.13	0.49
				0.00005	0.00013	0.00049	0.00005	0.00013	0.00049	0.00005	0.00013	0.00049
		CCSD	961	0.08	0.14	0.46	0.08	0.14	0.46	0.08	0.14	0.45
				0.00009	0.00015	0.00048	0.00008	0.00015	0.00047	0.00008	0.00015	0.00047
	$1.0\cdot10^{-5}$	MP2	1089	0.02	0.04	0.14	0.02	0.04	0.14	0.02	0.04	0.14
				0.00002	0.00004	0.00013	0.00002	0.00004	0.00013	0.00002	0.00004	0.00013
		CCSD	1060	0.02	0.04	0.13	0.02	0.04	0.13	0.02	0.04	0.13
				0.00002	0.00004	0.00012	0.00002	0.00004	0.00012	0.00002	0.00004	0.00012

TABLE A.29: MUE, RMSD, and MAX values, together with their counterparts normalized to the total number of COs (n_{CO}), for the atomization energies of the Karton test set relative to the full basis set references. Results are reported for the $(Q)_\Lambda$-(T), (Q)/A-(T), and (Q)/B-(T) contributions obtained with the cc-pVQZ basis set, at different FNO thresholds and using both MP2- and CCSD-derived densities. All values are in kJ mol^{-1}.

			n_{CO}	$(Q)_\Lambda$-(T)			(Q)/A-(T)			(Q)/B-(T)		
				MUE	RMSD	MAX	MUE	RMSD	MAX	MUE	RMSD	MAX
cc-pVQZ	$1.0\cdot10^{-3}$	MP2	384	0.77	0.95	1.88	0.86	1.08	2.37	0.80	1.01	2.19
				0.00199	0.00248	0.00490	0.00225	0.00282	0.00616	0.00210	0.00263	0.00570
		CCSD	401	0.78	0.99	2.13	0.89	1.14	2.55	0.83	1.07	2.41
				0.00196	0.00248	0.00531	0.00221	0.00284	0.00635	0.00208	0.00267	0.00601
	$5.0\cdot10^{-4}$	MP2	454	0.68	0.87	1.76	0.74	0.93	1.76	0.72	0.89	1.74
				0.00149	0.00193	0.00387	0.00163	0.00204	0.00389	0.00158	0.00197	0.00383
		CCSD	464	0.69	0.88	1.68	0.75	0.95	1.73	0.71	0.91	1.67
				0.00149	0.00190	0.00362	0.00162	0.00205	0.00374	0.00153	0.00196	0.00361
	projected cc-pVDZ	MP2	503	0.52	0.63	1.25	0.55	0.68	1.34	0.53	0.65	1.33
				0.00104	0.00125	0.00248	0.00110	0.00135	0.00267	0.00106	0.00130	0.00264
		CCSD	503	0.55	0.66	1.37	0.59	0.71	1.49	0.57	0.69	1.46
				0.00110	0.00131	0.00273	0.00117	0.00142	0.00295	0.00113	0.00137	0.00291
	$2.5\cdot10^{-4}$	MP2	573	0.41	0.53	1.16	0.44	0.56	1.24	0.41	0.53	1.22
				0.00072	0.00092	0.00203	0.00076	0.00097	0.00216	0.00072	0.00092	0.00214
		CCSD	587	0.48	0.63	1.37	0.52	0.68	1.36	0.49	0.64	1.34
				0.00081	0.00107	0.00233	0.00089	0.00115	0.00231	0.00084	0.00109	0.00228
	$1.0\cdot10^{-4}$	MP2	825	0.22	0.28	0.66	0.23	0.28	0.59	0.22	0.27	0.55
				0.00026	0.00034	0.00079	0.00028	0.00034	0.00071	0.00026	0.00032	0.00066
		CCSD	781	0.31	0.44	1.13	0.33	0.45	1.09	0.31	0.43	1.03
				0.00039	0.00056	0.00145	0.00042	0.00058	0.00140	0.00040	0.00055	0.00132
	$5.0\cdot10^{-5}$	MP2	1010	0.14	0.18	0.49	0.14	0.18	0.49	0.14	0.17	0.49
				0.00014	0.00018	0.00049	0.00014	0.00018	0.00049	0.00014	0.00017	0.00049
		CCSD	944	0.21	0.33	1.09	0.22	0.33	1.07	0.21	0.32	1.04
				0.00022	0.00035	0.00116	0.00023	0.00035	0.00113	0.00023	0.00034	0.00110
	projected cc-pVTZ	MP2	1179	0.13	0.16	0.41	0.13	0.17	0.41	0.13	0.16	0.41
				0.00011	0.00014	0.00035	0.00011	0.00014	0.00034	0.00011	0.00014	0.00035
		CCSD	1179	0.13	0.18	0.45	0.14	0.19	0.45	0.13	0.18	0.45
				0.00011	0.00015	0.00038	0.00011	0.00016	0.00038	0.00011	0.00016	0.00038
	$2.5\cdot10^{-5}$	MP2	1208	0.14	0.17	0.44	0.14	0.17	0.44	0.14	0.17	0.44
				0.00011	0.00014	0.00036	0.00011	0.00014	0.00036	0.00011	0.00014	0.00036
		CCSD	1114	0.14	0.20	0.59	0.15	0.20	0.58	0.15	0.20	0.57
				0.00013	0.00018	0.00053	0.00013	0.00018	0.00052	0.00013	0.00018	0.00051
	$1.0\cdot10^{-5}$	MP2	1570	0.08	0.12	0.35	0.08	0.12	0.35	0.08	0.11	0.35
				0.00005	0.00007	0.00022	0.00005	0.00007	0.00022	0.00005	0.00007	0.00022
		CCSD	1419	0.09	0.13	0.38	0.09	0.13	0.38	0.09	0.13	0.37
				0.00006	0.00009	0.00027	0.00007	0.00009	0.00027	0.00007	0.00009	0.00026

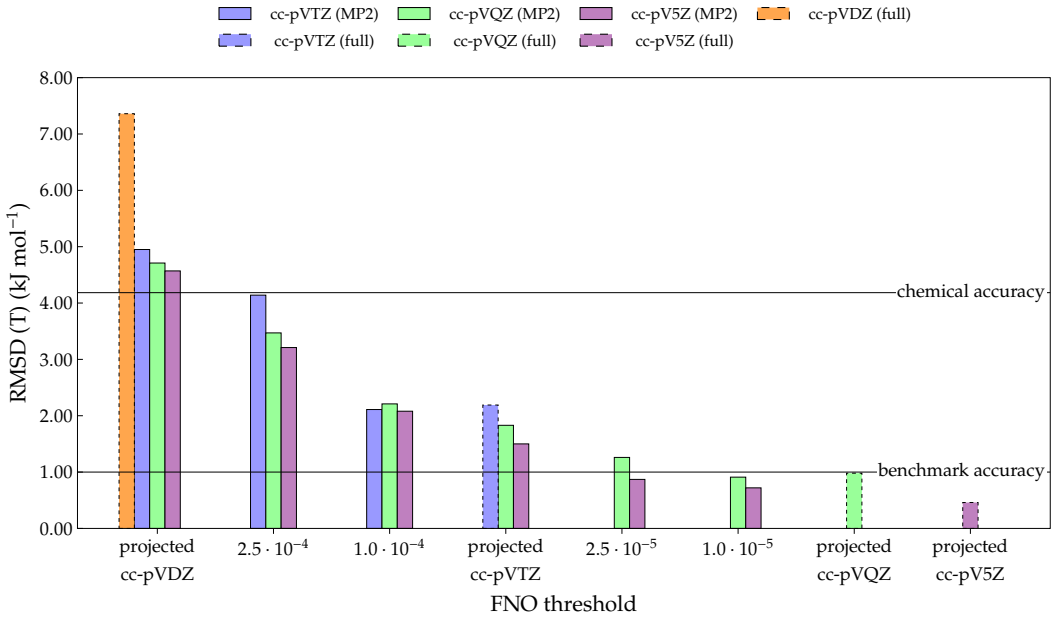


FIGURE A.5: RMSD for the (T) contribution to the atomization energies of the Karton test set, relative to the cc-pV(5,6)Z reference, as a function of the FNO truncation threshold. Results based on MP2-derived orbitals are labeled as (MP2) and calculations with the full virtual space as (full). Projected cc-pVnZ thresholds retain the same number of COs as the corresponding cc-pVnZ basis sets.

Bibliography

(1) Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Al-Laham, M. A.; Shirley, W. A.; Mantzaris, J. *The Journal of Chemical Physics* **1988**, *89*, 2193–2218.

(2) Pople, J. A.; Head-Gordon, M.; Fox, D. J.; Raghavachari, K.; Curtiss, L. A. *The Journal of Chemical Physics* **1989**, *90*, 5622–5629.

(3) DeYonker, N. J.; Cundari, T. R.; Wilson, A. K. *The Journal of Chemical Physics* **2006**, *124*, 114104.

(4) Bell, A. T.; Head-Gordon, M. *Annual Review of Chemical and Biomolecular Engineering* **2011**, *2*, 453–477.

(5) Vereecken, L.; Francisco, J. S. *Chemical Society Reviews* **2012**, *41*, 6259.

(6) Ruscic, B. *International Journal of Quantum Chemistry* **2014**, *114*, 1097–1101.

(7) Peterson, K. A.; Feller, D.; Dixon, D. A. *Theoretical Chemistry Accounts* **2012**, *131*, 1079.

(8) Mardirossian, N.; Head-Gordon, M. *Molecular Physics* **2017**, *115*, 2315–2372.

(9) Mehta, N.; Fellowes, T.; White, J. M.; Goerigk, L. *Journal of Chemical Theory and Computation* **2021**, *17*, 2783–2806.

(10) Karton, A.; Gruzman, D.; Martin, J. M. L. *The Journal of Physical Chemistry A* **2009**, *113*, 8434–8447.

(11) Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. *The Journal of Chemical Physics* **2006**, *125*, 144108.

(12) Harding, M. E.; Vázquez, J.; Ruscic, B.; Wilson, A. K.; Gauss, J.; Stanton, J. F. *The Journal of Chemical Physics* **2008**, *128*, 114111.

(13) Feller, D.; Peterson, K.; Dixon, D. The Impact of Larger Basis Sets and Explicitly Correlated Coupled Cluster Theory on the Feller–Peterson–Dixon Composite Method, In *Annual Reports in Computational Chemistry*, Dixon, D. A., Ed.; Elsevier: 2016; Vol. 12, pp 47–78.

(14) Lam Nguyen, T.; Ravishankara, A.; Stanton, J. F. *Chemical Physics Letters* **2018**, *708*, 100–105.

(15) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chemical Physics Letters* **1989**, *157*, 479–483.

(16) Werner, H.-J.; Hansen, A. *Journal of Chemical Theory and Computation* **2023**, *19*, 7007–7030.

(17) Kendall, R. A.; Früchtl, H. A. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling* **1997**, *97*, 158–163.

(18) Almlöf, J. *Chemical Physics Letters* **1991**, *181*, 319–320.

(19) Häser, M.; Almlöf, J. *The Journal of Chemical Physics* **1992**, *96*, 489–494.

(20) Wilson, A. K.; Almlöf, J. *Theoretica Chimica Acta* **1997**, *95*, 49–62.

(21) Lambrecht, D. S.; Doser, B.; Ochsenfeld, C. *The Journal of Chemical Physics* **2005**, *123*, 184102.

(22) Schütz, M.; Hetzer, G.; Werner, H.-J. *The Journal of Chemical Physics* **1999**, *111*, 5691–5705.

(23) Saebø, S.; Pulay, P. *Annual Review of Physical Chemistry* **1993**, *44*, 213–236.

(24) Handy, N. C.; Carter, S. *Molecular Physics* **2004**, *102*, 2201–2205.

(25) Scribano, Y.; Benoit, D. M. *Chemical Physics Letters* **2008**, *458*, 384–387.

(26) Curtiss, L. A.; Jones, C.; Trucks, G. W.; Raghavachari, K.; Pople, J. A. *The Journal of Chemical Physics* **1990**, *93*, 2537–2545.

(27) Helgaker, T.; Klopper, W.; Tew, D. P. *Molecular Physics* **2008**, *106*, 2107–2143.

(28) Wan, W.; Karton, A. *Chemical Physics Letters* **2016**, *643*, 34–38.

(29) Karton, A. Quantum Mechanical Thermochemical Predictions 100 Years After the Schrödinger Equation, In *Annual Reports in Computational Chemistry*, Dixon, D. A., Ed.; Elsevier: 2022; Vol. 18, pp 123–166.

(30) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *The Journal of Chemical Physics* **1991**, *94*, 7221–7230.

(31) Curtiss, L. A.; Carpenter, J. E.; Raghavachari, K.; Pople, J. A. *The Journal of Chemical Physics* **1992**, *96*, 9030–9034.

(32) Curtiss, L. A.; Raghavachari, K.; Pople, J. A. *The Journal of Chemical Physics* **1993**, *98*, 1293–1298.

(33) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *The Journal of Chemical Physics* **1998**, *109*, 7764–7776.

(34) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *The Journal of Chemical Physics* **2007**, *126*, 084108.

(35) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *The Journal of Chemical Physics* **2007**, *127*, 124105.

(36) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *WIREs Computational Molecular Science* **2011**, *1*, 810–825.

(37) Petersson, G. A.; Tensfeldt, T. G.; Montgomery, J. A. *The Journal of Chemical Physics* **1991**, *94*, 6091–6101.

(38) Petersson, G. A.; Al-Laham, M. A. *The Journal of Chemical Physics* **1991**, *94*, 6081–6090.

(39) Montgomery, J. A.; Ochterski, J. W.; Petersson, G. A. *The Journal of Chemical Physics* **1994**, *101*, 5900–5909.

(40) Ochterski, J. W.; Petersson, G. A.; Montgomery, J. A. *The Journal of Chemical Physics* **1996**, *104*, 2598–2619.

(41) Montgomery, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *The Journal of Chemical Physics* **1999**, *110*, 2822–2827.

(42) Montgomery, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *The Journal of Chemical Physics* **2000**, *112*, 6532–6542.

(43) Wood, G. P. F.; Radom, L.; Petersson, G. A.; Barnes, E. C.; Frisch, M. J.; Montgomery, J. A. *The Journal of Chemical Physics* **2006**, *125*, 094106.

(44) Martin, J. M. L.; De Oliveira, G. *The Journal of Chemical Physics* **1999**, *111*, 1843–1856.

(45) Boese, A. D.; Oren, M.; Atasoylu, O.; Martin, J. M. L.; Kállay, M.; Gauss, J. *The Journal of Chemical Physics* **2004**, *120*, 4129–4141.

(46) Karton, A.; Taylor, P. R.; Martin, J. M. L. *The Journal of Chemical Physics* **2007**, *127*, 064104.

(47) Karton, A.; Martin, J. M. L. *The Journal of Chemical Physics* **2012**, *136*, 124114.

(48) Sylvetsky, N.; Peterson, K. A.; Karton, A.; Martin, J. M. L. *The Journal of Chemical Physics* **2016**, *144*, 214101.

(49) Jiang, W.; DeYonker, N. J.; Determan, J. J.; Wilson, A. K. *The Journal of Physical Chemistry A* **2012**, *116*, 870–885.

(50) DeYonker, N. J.; Wilson, B. R.; Pierpont, A. W.; Cundari, T. R.; Wilson, A. K. *Molecular Physics* **2009**, *107*, 1107–1121.

(51) Oyedepo, G. A.; Wilson, A. K. *The Journal of Physical Chemistry A* **2010**, *114*, 8806–8816.

(52) Laury, M. L.; DeYonker, N. J.; Jiang, W.; Wilson, A. K. *The Journal of Chemical Physics* **2011**, *135*, 214103.

(53) Oyedepo, G. A.; Wilson, A. K. *ChemPhysChem* **2011**, *12*, 3320–3330.

(54) Nedd, S. A.; DeYonker, N. J.; Wilson, A. K.; Piecuch, P.; Gordon, M. S. *The Journal of Chemical Physics* **2012**, *136*, 144109.

(55) Mahler, A.; Wilson, A. K. *Journal of Chemical Theory and Computation* **2013**, *9*, 1402–1407.

(56) Riojas, A. G.; Wilson, A. K. *Journal of Chemical Theory and Computation* **2014**, *10*, 1500–1510.

(57) Feller, D.; Peterson, K. A.; De Jong, W. A.; Dixon, D. A. *The Journal of Chemical Physics* **2003**, *118*, 3510–3522.

(58) Feller, D.; Dixon, D. A. *The Journal of Physical Chemistry A* **2003**, *107*, 9641–9651.

(59) East, A. L. L.; Allen, W. D. *The Journal of Chemical Physics* **1993**, *99*, 4638–4650.

(60) Császár, A. G.; Allen, W. D.; Schaefer, H. F. *The Journal of Chemical Physics* **1998**, *108*, 9751–9764.

(61) Császár, A. G.; Tarczay, G.; Leininger, M. L.; Polyansky, O. L.; Tennyson, J.; Allen, W. D. Dream or Reality: Complete Basis Set Full Configuration Interaction Potential Energy Hypersurfaces, In *Spectroscopy from Space*, Demaison, J., Sarka, K., Cohen, E. A., Eds.; Springer Netherlands: Dordrecht, Netherlands, 2001; Vol. 20, pp 317–339.

(62) Kenny, J. P.; Allen, W. D.; Schaefer, H. F. *The Journal of Chemical Physics* **2003**, *118*, 7353–7365.

(63) Gonzales, J. M.; Pak, C.; Cox, R. S.; Allen, W. D.; Schaefer Iii, H. F.; Császár, A. G.; Tarczay, G. *Chemistry – A European Journal* **2003**, *9*, 2173–2192.

(64) Schuurman, M. S.; Muir, S. R.; Allen, W. D.; Schaefer, H. F. *The Journal of Chemical Physics* **2004**, *120*, 11586–11599.

(65) Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vázquez, J.; Stanton, J. F. *The Journal of Chemical Physics* **2004**, *121*, 11599–11613.

(66) Bomble, Y. J.; Vázquez, J.; Kállay, M.; Michauk, C.; Szalay, P. G.; Császár, A. G.; Gauss, J.; Stanton, J. F. *The Journal of Chemical Physics* **2006**, *125*, 064108.

(67) Puzzarini, C.; Bloino, J.; Tasinato, N.; Barone, V. *Chemical Reviews* **2019**, *119*, 8131–8191.

(68) Ventura, O. N.; Kieninger, M.; Katz, A.; Vega-Teijido, M.; Segovia, M.; Irving, K. *International Journal of Quantum Chemistry* **2021**, *121*, e26745.

(69) Bakowies, D. *The Journal of Chemical Physics* **2009**, *130*, 144113.

(70) Bakowies, D. *Journal of Chemical Theory and Computation* **2019**, *15*, 5230–5251.

(71) Bakowies, D. *Journal of Chemical Theory and Computation* **2020**, *16*, 399–426.

(72) Patel, P.; Melin, T. R.; North, S. C.; Wilson, A. K. Ab initio composite methodologies: Their significance for the chemistry community, In *Annual Reports in Computational Chemistry*, Dixon, D. A., Ed.; Elsevier: 2021; Vol. 17, pp 113–161.

(73) Ruscic, B.; Pinzon, R. E.; Morton, M. L.; Von Laszevski, G.; Bittner, S. J.; Nijsure, S. G.; Amin, K. A.; Minkoff, M.; Wagner, A. F. *The Journal of Physical Chemistry A* **2004**, *108*, 9979–9997.

(74) Ruscic, B.; Pinzon, R. E.; Laszewski, G. V.; Kodeboyina, D.; Burcat, A.; Leahy, D.; Montoy, D.; Wagner, A. F. *Journal of Physics: Conference Series* **2005**, *16*, 561–570.

(75) Afeefy, H. Y.; Liebman, J. F.; Stein, S. E. Neutral Thermochemical Data, In *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*, Linstrom, P. J., Mallard, W. G., Eds., http:// webbook.nist.gov; National Institute of Standards and Technology: Gaithersburg, Maryland.

(76) Johnson III, R. D. *Computational Chemistry Comparison and Benchmark DataBase, Release 21 Standard Reference Database 101*, http://cccbdb.nist.gov.

(77) Bartlett, R. J. Coupled-Cluster Theory: an Overview of Recent Developments, In *Modern Electronic Structure Therory, Part II*, Yarkony, D. R., Ed.; Advanced Series in Physical Chemistry, Vol. 2; World Scientific: Singapore, 1995, pp 1047–1131.

(78) Karton, A. *The Journal of Chemical Physics* **2020**, *153*, 024102.

(79) Karton, A. *WIREs Computational Molecular Science* **2016**, *6*, 292–310.

(80) Karton, A. *The Journal of Chemical Physics* **2018**, *149*, 034102.

(81) Karton, A. *Chemical Physics Letters* **2019**, *737*, 136810.

(82) Semidalas, E.; Karton, A.; Martin, J. M. L. *The Journal of Physical Chemistry A* **2024**, *128*, 1715–1724.

(83) Karton, A.; Daon, S.; Martin, J. M. *Chemical Physics Letters* **2011**, *510*, 165–178.

(84) Karton, A.; Sylvetsky, N.; Martin, J. M. L. *Journal of Computational Chemistry* **2017**, *38*, 2063–2075.

(85) Karton, A.; Martin, J. M. L. *The Journal of Chemical Physics* **2012**, *136*, 197101.

(86) Karton, A. *Chemical Physics Letters* **2016**, *645*, 118–122.

(87) Bomble, Y. J.; Stanton, J. F.; Kállay, M.; Gauss, J. *The Journal of Chemical Physics* **2005**, *123*, 054101.

(88) Semidalas, E.; Boese, A. D.; Martin, J. M. *Chemical Physics Letters* **2025**, *863*, 141874.

(89) Alessandrini, S.; Barone, V.; Puzzarini, C. *Journal of Chemical Theory and Computation* **2020**, *16*, 988–1006.

(90) Lupi, J.; Alessandrini, S.; Puzzarini, C.; Barone, V. *Journal of Chemical Theory and Computation* **2021**, *17*, 6974–6992.

(91) Nagy, P. R. *Chemical Science* **2024**, *15*, 14556–14584.

(92) Gyevi-Nagy, L.; Kállay, M.; Nagy, P. R. *Journal of Chemical Theory and Computation* **2020**, *16*, 366–384.

(93) Gyevi-Nagy, L.; Kállay, M.; Nagy, P. R. *Journal of Chemical Theory and Computation* **2021**, *17*, 860–878.

(94) Nagy, P. R.; Gyevi-Nagy, L.; Kállay, M. *Molecular Physics* **2021**, *119*, e1963495.

(95) Kállay, M.; Horváth, R. A.; Gyevi-Nagy, L.; Nagy, P. R. *Journal of Chemical Theory and Computation* **2023**, *19*, 174–189.

(96) Karton, A.; Gregory, K. P.; Chan, B. *Chemical Physics* **2025**, *595*, 112712.

(97) Alfonso-Ramos, J. E.; Adamo, C.; Brémond, É.; Stuyver, T. *Journal of Chemical Theory and Computation* **2025**, *21*, 1752–1761.

(98) Karton, A. *Physical Chemistry Chemical Physics* **2024**, *26*, 14594–14606.

(99) Karton, A.; Tarnopolsky, A.; Lamère, J.-F.; Schatz, G. C.; Martin, J. M. L. *The Journal of Physical Chemistry A* **2008**, *112*, 12868–12886.

(100) Ehlert, S.; Hermann, J.; Vogels, T.; Satorras, V. G.; Lanius, S.; Segler, M.; Kooi, D. P.; Takeda, K.; Huang, C.-W.; Luise, G.; van den Berg, R.; Gori-Giorgi, P.; Karton, A., arXiv:2506.14492, 2025.

(101) Luise, G.; Huang, C.-W.; Vogels, T.; Kooi, D. P.; Ehlert, S.; Lanius, S.; Giesbertz, K. J. H.; Karton, A.; Gunceler, D.; Stanley, M.; Bruinsma, W. P.; Huang, L.; Wei, X.; Torres, J. G.; Katbashev, A.; Zavaleta, R. C.; Máté, B.; Kaba, S.-O.; Sordillo, R.; Chen, Y.; Williams-Young, D. B.; Bishop, C. M.; Hermann, J.; van den Berg, R.; Gori-Giorgi, P., arXiv:2506.14665, 2025.

(102) Redfern, P. C.; Zapol, P.; Curtiss, L. A.; Raghavachari, K. *The Journal of Physical Chemistry A* **2000**, *104*, 5850–5854.

(103) Salta, Z.; Tasinato, N.; Liebman, J. F.; Ventura, O. N. *Chemical Thermodynamics and Thermal Analysis* **2023**, *12*, 100121.

(104) Cundari, T. R.; Arturo Ruiz Leza, H.; Grimes, T.; Steyl, G.; Waters, A.; Wilson, A. K. *Chemical Physics Letters* **2005**, *401*, 58–61.

(105) Neese, F.; Wennmohs, F.; Hansen, A. *The Journal of Chemical Physics* **2009**, *130*, 114108.

(106) Neese, F.; Hansen, A.; Liakos, D. G. *The Journal of Chemical Physics* **2009**, *131*, 064103.

(107) Guo, Y.; Riplinger, C.; Liakos, D. G.; Becker, U.; Saitow, M.; Neese, F. *The Journal of Chemical Physics* **2020**, *152*, 024116.

(108) Schmitz, G.; Hättig, C. *The Journal of Chemical Physics* **2016**, *145*, 234107.

(109) Ma, Q.; Werner, H.-J. *Journal of Chemical Theory and Computation* **2021**, *17*, 902–926.

(110) Bensberg, M.; Neugebauer, J. *Journal of Chemical Theory and Computation* **2020**, *16*, 3607–3619.

(111) D'Cunha, R.; Crawford, T. D. *Journal of Chemical Theory and Computation* **2021**, *17*, 290–301.

(112) Nagy, P. R.; Kállay, M. *Journal of Chemical Theory and Computation* **2019**, *15*, 5275–5298.

(113) Al-Hamdani, Y. S.; Nagy, P. R.; Zen, A.; Barton, D.; Kállay, M.; Brandenburg, J. G.; Tkatchenko, A. *Nature Communications* **2021**, *12*, 3927.

(114) Pogrebetsky, J.; Siklitskaya, A.; Kubas, A. *Journal of Chemical Theory and Computation* **2023**, *19*, 4023–4032.

(115) Barone, V. *WIREs Computational Molecular Science* **2016**, *6*, 86–110.

(116) Sugiki, T.; Kobayashi, N.; Fujiwara, T. *Computational and Structural Biotechnology Journal* **2017**, *15*, 328–339.

(117) Puzzarini, C.; Barone, V. *Accounts of Chemical Research* **2018**, *51*, 548–556.

(118) He, Y.; Tang, L.; Wu, X.; Hou, X.; Lee, Y.-I. *Applied Spectroscopy Reviews* **2007**, *42*, 119–138.

(119) Martin-Drumel, M.-A.; McCarthy, M. C.; Patterson, D.; McGuire, B. A.; Crabtree, K. N. *The Journal of Chemical Physics* **2016**, *144*, 124202.

(120) Cleeton, C. E.; Williams, N. H. *Physical Review* **1934**, *45*, 234–237.

(121) Lesarri, A.; Mata, S.; López, J. C.; Alonso, J. L. *Review of Scientific Instruments* **2003**, *74*, 4799–4804.

(122) Brown, G. G.; Dian, B. C.; Douglass, K. O.; Geyer, S. M.; Shipman, S. T.; Pate, B. H. *Review of Scientific Instruments* **2008**, *79*, 053103.

(123) Steber, A. L.; Neill, J. L.; Zaleski, D. P.; Pate, B. H.; Lesarri, A.; Bird, R. G.; Vaquero-Vara, V.; Pratt, D. W. *Faraday Discussions* **2011**, *150*, 227.

(124) Grabow, J.-U.; Caminati, W. Microwave Spectroscopy, In *Frontiers of Molecular Spectroscopy*, Laane, J., Ed., 1st ed.; Elsevier: Amsterdam, Netherlands, 2009, pp 383–454.

(125) León, I.; Alonso, E. R.; Mata, S.; Cabezas, C.; Alonso, J. L. *Angewandte Chemie International Edition* **2019**, *58*, 16002–16007.

(126) León, I.; Alonso, E. R.; Mata, S.; Alonso, J. L. *The Journal of Physical Chemistry Letters* **2021**, *12*, 6983–6987.

(127) Insausti, A.; Alonso, E. R.; Municio, S.; León, I.; Kolesniková, L.; Mata, S. *The Journal of Physical Chemistry Letters* **2025**, *16*, 2425–2432.

(128) Pinacho, P.; Caliebe, S. V. M.; Quesada-Moreno, M. M.; Zinn, S.; Schnell, M. *Physical Chemistry Chemical Physics* **2022**, *24*, 5539–5545.

(129) Peña, I.; Cabezas, C.; Alonso, J. L. *Angewandte Chemie International Edition* **2015**, *54*, 2991–2994.

(130) Domingos, S. R.; Cnossen, A.; Buma, W. J.; Browne, W. R.; Feringa, B. L.; Schnell, M. *Angewandte Chemie International Edition* **2017**, *56*, 11209–11212.

(131) Pérez, C.; Muckle, M. T.; Zaleski, D. P.; Seifert, N. A.; Temelso, B.; Shields, G. C.; Kisiel, Z.; Pate, B. H. *Science* **2012**, *336*, 897–901.

(132) Chrayteh, M.; Burevschi, E.; Loru, D.; Huet, T. R.; Dréan, P.; Sanz, M. E. *Physical Chemistry Chemical Physics* **2021**, *23*, 20686–20694.

(133) Medel, R.; Camiruaga, A.; Saragi, R. T.; Pinacho, P.; Pérez, C.; Schnell, M.; Lesarri, A.; Suhm, M. A.; Fernández, J. A. *Physical Chemistry Chemical Physics* **2021**, *23*, 23610–23624.

(134) León, I.; Fusè, M.; Alonso, E. R.; Mata, S.; Mancini, G.; Puzzarini, C.; Alonso, J. L.; Barone, V. *The Journal of Chemical Physics* **2022**, *157*, 074107.

(135) Calabrese, C.; Uriarte, I.; Insausti, A.; Vallejo-López, M.; Basterretxea, F. J.; Cochrane, S. A.; Davis, B. G.; Corzana, F.; Cocinero, E. J. *ACS Central Science* **2020**, *6*, 293–303.

(136) Melosso, M.; Bizzocchi, L.; Gazzeh, H.; Tonolo, F.; Guillemin, J.-C.; Alessandrini, S.; Rivilla, V. M.; Dore, L.; Barone, V.; Puzzarini, C. *Chemical Communications* **2022**, *58*, 2750–2753.

(137) Melli, A.; Melosso, M.; Bizzocchi, L.; Alessandrini, S.; Jiang, N.; Tonolo, F.; Boi, S.; Castellan, G.; Sapienza, C.; Guillemin, J.-C.; Dore, L.; Puzzarini, C. *The Journal of Physical Chemistry A* **2022**, *126*, 6210–6220.

(138) McCarthy, M. C.; Lee, K. L. K.; Carroll, P. B.; Porterfield, J. P.; Changala, P. B.; Thorpe, J. H.; Stanton, J. F. *The Journal of Physical Chemistry A* **2020**, *124*, 5170–5181.

(139) Puzzarini, C.; Stanton, J. F.; Gauss, J. *International Reviews in Physical Chemistry* **2010**, *29*, 273–367.

(140) Barone, V.; Alessandrini, S.; Biczysko, M.; Cheeseman, J. R.; Clary, D. C.; McCoy, A. B.; DiRisio, R. J.; Neese, F.; Melosso, M.; Puzzarini, C. *Nature Reviews Methods Primers* **2021**, *1*, 38.

(141) Pulay, P.; Meyer, W.; Boggs, J. E. *The Journal of Chemical Physics* **1978**, *68*, 5077–5085.

(142) Pawłowski, F.; Jørgensen, P.; Olsen, J.; Hegelund, F.; Helgaker, T.; Gauss, J.; Bak, K. L.; Stanton, J. F. *The Journal of Chemical Physics* **2002**, *116*, 6482–6496.

(143) Demaison, J. *Molecular Physics* **2007**, *105*, 3109–3138.

(144) Demaison, J.; Boggs, J. E.; Császár, A. G.; CRC press, Taylor & Francis Group: Boca Raton, Florida, 2011.

(145) Penocchio, E.; Piccardo, M.; Barone, V. *Journal of Chemical Theory and Computation* **2015**, *11*, 4689–4707.

(146) Mendolicchio, M.; Penocchio, E.; Licari, D.; Tasinato, N.; Barone, V. *Journal of Chemical Theory and Computation* **2017**, *13*, 3060–3075.

(147) Vogt, N.; Demaison, J.; Lecture Notes in Chemistry, Vol. 111; Springer International Publishing: Cham, Switzerland, 2023.

(148) Ceselin, G.; Barone, V.; Tasinato, N. *Journal of Chemical Theory and Computation* **2021**, *17*, 7290–7311.

(149) Barone, V.; Ceselin, G.; Lazzari, F.; Tasinato, N. *The Journal of Physical Chemistry A* **2023**, *127*, 5183–5192.

(150) Demaison, J.; Jahn, M. K.; Cocinero, E. J.; Lesarri, A.; Grabow, J.-U.; Guillemin, J.-C.; Rudolph, H. D. *The Journal of Physical Chemistry A* **2013**, *117*, 2278–2284.

(151) Vogt, N.; Demaison, J.; Cocinero, E. J.; Écija, P.; Lesarri, A.; Rudolph, H. D.; Vogt, J. *Physical Chemistry Chemical Physics* **2016**, *18*, 15555–15563.

(152) Puzzarini, C.; Stanton, J. F. *Physical Chemistry Chemical Physics* **2023**, *25*, 1421–1429.

(153) Thorpe, J. H.; Lopez, C. A.; Nguyen, T. L.; Baraban, J. H.; Bross, D. H.; Ruscic, B.; Stanton, J. F. *The Journal of Chemical Physics* **2019**, *150*, 224102.

(154) Franke, P. R.; Stanton, J. F. *The Journal of Physical Chemistry A* **2023**, *127*, 924–937.

(155) Sahoo, N. P.; Franke, P. R.; Stanton, J. F. *Journal of Computational Chemistry* **2024**, *45*, 1419–1427.

(156) Lupi, J.; Puzzarini, C.; Cavallotti, C.; Barone, V. *Journal of Chemical Theory and Computation* **2020**, *16*, 5090–5104.

(157) Puzzarini, C.; Barone, V. *Physical Chemistry Chemical Physics* **2020**, *22*, 6507–6523.

(158) Heim, Z. N.; Amberger, B. K.; Esselman, B. J.; Stanton, J. F.; Woods, R. C.; McMahon, R. J. *The Journal of Chemical Physics* **2020**, *152*, 104303.

(159) Watrous, A. G.; Westbrook, B. R.; Fortenberry, R. C. *The Journal of Physical Chemistry A* **2021**, *125*, 10532–10540.

(160) Puzzarini, C.; Barone, V. *Physical Chemistry Chemical Physics* **2011**, *13*, 7189.

(161) Spackman, P. R.; Jayatilaka, D.; Karton, A. *The Journal of Chemical Physics* **2016**, *145*, 104101.

(162) Warden, C. E.; Smith, D. G. A.; Burns, L. A.; Bozkaya, U.; Sherrill, C. D. *The Journal of Chemical Physics* **2020**, *152*, 124109.

(163) Tasinato, N.; Puzzarini, C.; Barone, V. *Angewandte Chemie International Edition* **2017**, *56*, 13838–13841.

(164) Mancini, G.; Fusè, M.; Lazzari, F.; Barone, V. *Digital Discovery* **2022**, *1*, 790–805.

(165) Barone, V.; Fusè, M.; Lazzari, F.; Mancini, G. *Journal of Chemical Theory and Computation* **2023**, *19*, 1243–1260.

(166) Barone, V.; Puzzarini, C. *Annual Review of Physical Chemistry* **2023**, *74*, 29–52.

(167) Barone, V. *Physical Chemistry Chemical Physics* **2024**, *26*, 5802–5821.

(168) Møller, C.; Plesset, M. S. *Physical Review* **1934**, *46*, 618–622.

(169) Alonso, J. L.; Vaquero, V.; Peña, I.; López, J. C.; Mata, S.; Caminati, W. *Angewandte Chemie International Edition* **2013**, *52*, 2331–2334.

(170) Caliebe, S. V. M.; Pinacho, P.; Schnell, M. *The Journal of Physical Chemistry Letters* **2022**, *13*, 11913–11917.

(171) Vats, A.; Pathak, A. *Monthly Notices of the Royal Astronomical Society* **2022**, *517*, 5780–5790.

(172) Vats, A.; Srivastav, S.; Pandey, A.; Pathak, A. *Physical Chemistry Chemical Physics* **2023**, *25*, 19066–19072.

(173) Alonso, E. R.; Fusè, M.; León, I.; Puzzarini, C.; Alonso, J. L.; Barone, V. *The Journal of Physical Chemistry A* **2021**, *125*, 2121–2129.

(174) Piccardo, M.; Penocchio, E.; Puzzarini, C.; Biczysko, M.; Barone, V. *The Journal of Physical Chemistry A* **2015**, *119*, 2058–2082.

(175) Melli, A.; Tonolo, F.; Barone, V.; Puzzarini, C. *The Journal of Physical Chemistry A* **2021**, *125*, 9904–9916.

(176) Barone, V.; Crisci, L.; Di Grande, S. *Journal of Chemical Theory and Computation* **2023**, *19*, 7273–7286.

(177) Barone, V.; Uribe Grajales, L. M.; Di Grande, S.; Lazzari, F.; Mendolicchio, M. *The Journal of Physical Chemistry A* **2023**, *127*, 7534–7543.

(178) Alessandrini, S.; Melosso, M.; Bizzocchi, L.; Barone, V.; Puzzarini, C. *The Journal of Physical Chemistry A* **2024**, *128*, 5833–5855.

(179) Ye, H.; Alessandrini, S.; Melosso, M.; Puzzarini, C. *Physical Chemistry Chemical Physics* **2022**, *24*, 23254–23264.

(180) Melli, A.; Barone, V.; Puzzarini, C. *The Journal of Physical Chemistry A* **2021**, *125*, 2989–2998.

(181) Lazzari, F.; Mendolicchio, M.; Barone, V. *The Journal of Physical Chemistry A* **2024**, *128*, 1385–1395.

(182) Di Grande, S.; Barone, V. *The Journal of Physical Chemistry A* **2024**, *128*, 4886–4900.

(183) Di Grande, S.; Lazzari, F.; Barone, V. *Journal of Chemical Theory and Computation* **2024**, *20*, 9243–9258.

(184) Di Grande, S.; Kállay, M.; Barone, V. *Journal of Computational Chemistry* **2023**, *44*, 2149–2157.

(185) Kállay, M.; Horváth, R. A.; Gyevi-Nagy, L.; Nagy, P. R. *The Journal of Chemical Physics* **2021**, *155*, 034107.

(186) Barone, V.; Di Grande, S.; Lazzari, F.; Mendolicchio, M. *The Journal of Physical Chemistry A* **2023**, *127*, 6771–6778.

(187) Barone, V. *The Journal of Chemical Physics* **2023**, *159*, 081102.

(188) Barone, V.; Lazzari, F. *The Journal of Physical Chemistry A* **2023**, *127*, 10517–10527.

(189) Uribe, L.; Lazzari, F.; Di Grande, S.; Crisci, L.; Mendolicchio, M.; Barone, V. *The Journal of Chemical Physics* **2024**, *161*, 014307.

(190) Uribe, L.; Di Grande, S.; Mendolicchio, M.; Tasinato, N.; Barone, V. *The Journal of Physical Chemistry A* **2024**, *128*, 10474–10488.

(191) Barone, V.; Lazzari, F.; Di Grande, S. *The Journal of Physical Chemistry A* **2025**, *129*, 2876–2886.

(192) Barone, V.; Lazzari, F.; Grande, S. D.; Crisci, L. *Chemical Physics Letters* **2025**, *866*, 141956.

(193) Lazzari, F.; Di Grande, S.; Crisci, L.; Mendolicchio, M.; Barone, V. *The Journal of Chemical Physics* **2025**, *162*, 114310.

(194) Helgaker, T.; Jørgensen, P.; Olsen, J.; John Wiley & Sons: Chichester, England, 2000.

(195) Jensen, F., 2nd ed.; John Wiley & Sons: Chichester, England, 2007.

(196) Åstrand, P.-O., 1st ed.; BookBon: ebook, 2016.

(197) Szabo, A.; Ostlund, N. S.; Dover Publications: Mineola, New York, 1996.

(198) Shavitt, I.; Bartlett, R. J.; Cambridge University Press: Cambridge, New York, 2009.

(199) Kong, L.; Bischoff, F. A.; Valeev, E. F. *Chemical Reviews* **2012**, *112*, 75–107.

(200) Köhn, A. Complete Basis Set Results in Electron Correlation Methods Using F12 Theory, In *Chemical Modelling: Applications and Theory*, Springborg, M., Joswig, J.-O., Eds.; Royal Society of Chemistry: Cambridge, England, 2013; Vol. 10, pp 32–63.

(201) Hättig, C.; Klopper, W.; Köhn, A.; Tew, D. P. *Chemical Reviews* **2012**, *112*, 4–74.

(202) Butler, L. J. *Annual Review of Physical Chemistry* **1998**, *49*, 125–171.

(203) Roothaan, C. C. J. *Reviews of Modern Physics* **1951**, *23*, 69–89.

(204) Olsen, J.; Christiansen, O.; Koch, H.; Jørgensen, P. *The Journal of Chemical Physics* **1996**, *105*, 5082–5090.

(205) Cremer, D.; He, Z. *The Journal of Physical Chemistry* **1996**, *100*, 6173–6188.

(206) Pople, J. A.; Binkley, J. S.; Seeger, R. *International Journal of Quantum Chemistry* **1976**, *10*, 1–19.

(207) Bartlett, R. J.; Purvis, G. D. *International Journal of Quantum Chemistry* **1978**, *14*, 561–581.

(208) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *International Journal of Quantum Chemistry* **1978**, *14*, 545–560.

(209) Purvis, G. D.; Bartlett, R. J. *The Journal of Chemical Physics* **1982**, *76*, 1910–1918.

(210) Noga, J.; Bartlett, R. J. *The Journal of Chemical Physics* **1987**, *86*, 7041–7050.

(211) Scuseria, G. E.; Schaefer, H. F. *Chemical Physics Letters* **1988**, *152*, 382–386.

(212) Oliphant, N.; Adamowicz, L. *The Journal of Chemical Physics* **1991**, *95*, 6645–6651.

(213) Kucharski, S. A.; Bartlett, R. J. *The Journal of Chemical Physics* **1992**, *97*, 4282–4288.

(214) Kállay, M.; Surján, P. R. *The Journal of Chemical Physics* **2001**, *115*, 2945–2954.

(215) Kállay, M.; Gauss, J. *The Journal of Chemical Physics* **2005**, *123*, 214105.

(216) Stanton, J. F.; Gauss, J. *Theoretica Chimica Acta* **1996**, *93*, 303–313.

(217) Crawford, T. D.; Stanton, J. F. *International Journal of Quantum Chemistry* **1998**, *70*, 601–611.

(218) Kucharski, S. A.; Bartlett, R. J. *The Journal of Chemical Physics* **1998**, *108*, 5243–5254.

(219) Kucharski, S. A.; Bartlett, R. J. *The Journal of Chemical Physics* **1998**, *108*, 5255–5264.

(220) Kállay, M.; Gauss, J. *The Journal of Chemical Physics* **2008**, *129*, 144101.

(221) Kato, T. *Communications on Pure and Applied Mathematics* **1957**, *10*, 151–177.

(222) Hylleraas, E. A. *Zeitschrift für Physik* **1928**, *48*, 469–494.

(223) Hylleraas, E. A. *Zeitschrift für Physik* **1929**, *54*, 347–366.

(224) Kutzelnigg, W. *Theoretica chimica acta* **1985**, *68*, 445–469.

(225) Klopper, W.; Kutzelnigg, W. *The Journal of Physical Chemistry* **1990**, *94*, 5625–5630.

(226) Kutzelnigg, W.; Klopper, W. *The Journal of Chemical Physics* **1991**, *94*, 1985–2001.

(227) Termath, V.; Klopper, W.; Kutzelnigg, W. *The Journal of Chemical Physics* **1991**, *94*, 2002–2019.

(228) Klopper, W.; Kutzelnigg, W. *The Journal of Chemical Physics* **1991**, *94*, 2020–2030.

(229) Klopper, W. *Chemical Physics Letters* **1991**, *186*, 583–585.

(230) Klopper, W.; Samson, C. C. M. *The Journal of Chemical Physics* **2002**, *116*, 6397–6410.

(231) Ten-no, S. *The Journal of Chemical Physics* **2004**, *121*, 117–129.

(232) Ten-no, S. *Chemical Physics Letters* **2004**, *398*, 56–61.

(233) May, A. J.; Manby, F. R. *The Journal of Chemical Physics* **2004**, *121*, 4479–4485.

(234) Werner, H.-J.; Adler, T. B.; Manby, F. R. *The Journal of Chemical Physics* **2007**, *126*, 164102.

(235) Adler, T. B.; Knizia, G.; Werner, H.-J. *The Journal of Chemical Physics* **2007**, *127*, 221106.

(236) Bokhan, D.; Ten-no, S.; Noga, J. *Physical Chemistry Chemical Physics* **2008**, *10*, 3320.

(237) Noga, J.; Kedžuch, S.; Šimunek, J.; Ten-no, S. *The Journal of Chemical Physics* **2008**, *128*, 174103.

(238) Köhn, A.; Richings, G. W.; Tew, D. P. *The Journal of Chemical Physics* **2008**, *129*, 201103.

(239) Knizia, G.; Adler, T. B.; Werner, H.-J. *The Journal of Chemical Physics* **2009**, *130*, 054104.

(240) Bokhan, D.; Bernadotte, S.; Ten-no, S. *Chemical Physics Letters* **2009**, *469*, 214–218.

(241) Hättig, C.; Tew, D. P.; Köhn, A. *The Journal of Chemical Physics* **2010**, *132*, 231102.

(242) Werner, H.-J.; Knizia, G.; Manby, F. R. *Molecular Physics* **2011**, *109*, 407–417.

(243) Marchetti, O.; Werner, H.-J. *Physical Chemistry Chemical Physics* **2008**, *10*, 3400.

(244) Kállay, M. *The Journal of Chemical Physics* **2014**, *141*, 244113.

(245) Fliegl, H.; Klopper, W.; Hättig, C. *The Journal of Chemical Physics* **2005**, *122*, 084107.

(246) Tew, D. P.; Klopper, W.; Neiss, C.; Hättig, C. *Physical Chemistry Chemical Physics* **2007**, *9*, 1921–1930.

(247) Torheyden, M.; Valeev, E. F. *Physical Chemistry Chemical Physics* **2008**, *10*, 3410.

(248) Valeev, E. F.; Daniel Crawford, T. *The Journal of Chemical Physics* **2008**, *128*, 244113.

(249) Köhn, A. *The Journal of Chemical Physics* **2009**, *130*, 104104.

(250) Köhn, A. *The Journal of Chemical Physics* **2009**, *130*, 131101.

(251) Hohenberg, P.; Kohn, W. *Physical Review* **1964**, *136*, B864–B871.

(252) Kohn, W.; Sham, L. J. *Physical Review* **1965**, *140*, A1133–A1138.

(253) Perdew, J. P.; Schmidt, K. Jacob's Ladder of Density Functional Approximations for the Exchange-Correlation Energy, In *AIP Conference Proceedings*, AIP: Antwerp, Belgium, 2001; Vol. 577, pp 1–20.

(254) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *The Journal of Chemical Physics* **2010**, *132*, 154104.

(255) Smith, D. G. A.; Burns, L. A.; Patkowski, K.; Sherrill, C. D. *The Journal of Physical Chemistry Letters* **2016**, *7*, 2197–2203.

(256) Grimme, S.; Ehrlich, S.; Goerigk, L. *Journal of Computational Chemistry* **2011**, *32*, 1456–1465.

(257) Caldeweyher, E.; Bannwarth, C.; Grimme, S. *The Journal of Chemical Physics* **2017**, *147*, 034112.

(258) Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. *The Journal of Chemical Physics* **2019**, *150*, 154122.

(259) Hehre, W. J.; Stewart, R. F.; Pople, J. A. *The Journal of Chemical Physics* **1969**, *51*, 2657–2664.

(260) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *Journal of the American Chemical Society* **1980**, *102*, 939–947.

(261) Gordon, M. S.; Binkley, J. S.; Pople, J. A.; Pietro, W. J.; Hehre, W. J. *Journal of the American Chemical Society* **1982**, *104*, 2797–2803.

(262) Hariharan, P. C.; Pople, J. A. *Theoretica Chimica Acta* **1973**, *28*, 213–222.

(263) Dunning, T. H. *The Journal of Chemical Physics* **1989**, *90*, 1007–1023.

(264) Dunning, T. H.; Peterson, K. A.; Wilson, A. K. *The Journal of Chemical Physics* **2001**, *114*, 9244–9253.

(265) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *The Journal of Chemical Physics* **1992**, *96*, 6796–6806.

(266) Woon, D. E.; Dunning, T. H. *The Journal of Chemical Physics* **1995**, *103*, 4572–4585.

(267) Peterson, K. A.; Dunning, T. H. *The Journal of Chemical Physics* **2002**, *117*, 10548–10560.

(268) Papajak, E.; Zheng, J.; Xu, X.; Leverentz, H. R.; Truhlar, D. G. *Journal of Chemical Theory and Computation* **2011**, *7*, 3027–3034.

(269) Peterson, K. A.; Adler, T. B.; Werner, H.-J. *The Journal of Chemical Physics* **2008**, *128*, 084102.

(270) Yousaf, K. E.; Peterson, K. A. *The Journal of Chemical Physics* **2008**, *129*, 184108.

(271) Hill, J. G.; Mazumder, S.; Peterson, K. A. *The Journal of Chemical Physics* **2010**, *132*, 054108.

(272) Hill, J. G.; Peterson, K. A.; Knizia, G.; Werner, H.-J. *The Journal of Chemical Physics* **2009**, *131*, 194105.

(273) DePrince, A. E.; Sherrill, C. D. *Journal of Chemical Theory and Computation* **2013**, *9*, 2687–2696.

(274) Janowski, T.; Pulay, P. *Journal of Chemical Theory and Computation* **2008**, *4*, 1585–1592.

(275) Anisimov, V. M.; Bauer, G. H.; Chadalavada, K.; Olson, R. M.; Glenski, J. W.; Kramer, W. T. C.; Aprà, E.; Kowalski, K. *Journal of Chemical Theory and Computation* **2014**, *10*, 4307–4316.

(276) Peng, C.; Calvin, J. A.; Valeev, E. F. *International Journal of Quantum Chemistry* **2019**, *119*, e25894.

(277) Parrish, R. M.; Sherrill, C. D.; Hohenstein, E. G.; Kokkila, S. I. L.; Martínez, T. J. *The Journal of Chemical Physics* **2014**, *140*, 181102.

(278) Ma, Q.; Werner, H.-J. *WIREs Computational Molecular Science* **2018**, *8*, e1371.

(279) Guo, Y.; Riplinger, C.; Becker, U.; Liakos, D. G.; Minenkov, Y.; Cavallo, L.; Neese, F. *The Journal of Chemical Physics* **2018**, *148*, 011101.

(280) Schmitz, G.; Hättig, C.; Tew, D. P. *Physical Chemistry Chemical Physics* **2014**, *16*, 22167–22178.

(281) Li, W.; Piecuch, P.; Gour, J. R.; Li, S. *The Journal of Chemical Physics* **2009**, *131*, 114109.

(282) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *The Journal of Chemical Physics* **1997**, *106*, 9639–9646.

(283) Marshall, M. S.; Burns, L. A.; Sherrill, C. D. *The Journal of Chemical Physics* **2011**, *135*, 194102.

(284) Varandas, A.; Piecuch, P. *Chemical Physics Letters* **2006**, *430*, 448–453.

(285) Adamowicz, L.; Bartlett, R. J. *The Journal of Chemical Physics* **1987**, *86*, 6314–6324.

(286) Adamowicz, L.; Bartlett, R. J.; Sadlej, A. J. *The Journal of Chemical Physics* **1988**, *88*, 5749–5758.

(287) Neogrády, P.; Pitoňák, M.; Urban, M. *Molecular Physics* **2005**, *103*, 2141–2157.

(288) Sosa, C.; Geertsen, J.; Trucks, G. W.; Bartlett, R. J.; Franz, J. A. *Chemical Physics Letters* **1989**, *159*, 148–154.

(289) Pitoňák, M.; Aquilante, F.; Hobza, P.; Neogrády, P.; Noga, J.; Urban, M. *Collection of Czechoslovak Chemical Communications* **2011**, *76*, 713–742.

(290) Clement, M. C.; Zhang, J.; Lewis, C. A.; Yang, C.; Valeev, E. F. *Journal of Chemical Theory and Computation* **2018**, *14*, 4581–4589.

(291) Löwdin, P.-O. *Physical Review* **1955**, *97*, 1474–1489.

(292) Edmiston, C.; Krauss, M. *The Journal of Chemical Physics* **1966**, *45*, 1833–1839.

(293) Barr, T. L.; Davidson, E. R. *Physical Review A* **1970**, *1*, 644–658.

(294) Meyer, W. *The Journal of Chemical Physics* **1973**, *58*, 1017–1035.

(295) Hay, P. J. *The Journal of Chemical Physics* **1973**, *59*, 2468–2476.

(296) Ahlrichs, R.; Lischka, H.; Staemmler, V.; Kutzelnigg, W. *The Journal of Chemical Physics* **1975**, *62*, 1225–1234.

(297) Taube, A. G.; Bartlett, R. J. *Collection of Czechoslovak Chemical Communications* **2005**, *70*, 837–850.

(298) Taube, A. G.; Bartlett, R. J. *The Journal of Chemical Physics* **2008**, *128*, 164101.

(299) Landau, A.; Khistyaev, K.; Dolgikh, S.; Krylov, A. I. *The Journal of Chemical Physics* **2010**, *132*, 014109.

(300) Rolik, Z.; Kállay, M. *The Journal of Chemical Physics* **2011**, *134*, 124111.

(301) Grüneis, A.; Booth, G. H.; Marsman, M.; Spencer, J.; Alavi, A.; Kresse, G. *Journal of Chemical Theory and Computation* **2011**, *7*, 2780–2785.

(302) Mester, D.; Nagy, P. R.; Kállay, M. *The Journal of Chemical Physics* **2018**, *148*, 094111.

(303) Ramberger, B.; Sukurma, Z.; Schäfer, T.; Kresse, G. *The Journal of Chemical Physics* **2019**, *151*, 214106.

(304) Segarra-Martí, J.; Garavelli, M.; Aquilante, F. *Journal of Chemical Theory and Computation* **2015**, *11*, 3772–3784.

(305) DePrince, A. E.; Sherrill, C. D. *Journal of Chemical Theory and Computation* **2013**, *9*, 293–299.

(306) Mester, D.; Nagy, P. R.; Kállay, M. *The Journal of Chemical Physics* **2017**, *146*, 194102.

(307) Pokhilko, P.; Izmodenov, D.; Krylov, A. I. *The Journal of Chemical Physics* **2020**, *152*, 034105.

(308) Folkestad, S. D.; Koch, H. *Journal of Chemical Theory and Computation* **2020**, *16*, 6869–6879.

(309) Kumar, A.; Crawford, T. D. *The Journal of Physical Chemistry A* **2017**, *121*, 708–716.

(310) Segarra-Martí, J.; Garavelli, M.; Aquilante, F. *The Journal of Chemical Physics* **2018**, *148*, 034107.

(311) Pitoňák, M.; Holka, F.; Neogrády, P.; Urban, M. *Journal of Molecular Structure: Theochem* **2006**, *768*, 79–89.

(312) Irmler, A.; Gallo, A.; Hummel, F.; Grüneis, A. *Physical Review Letters* **2019**, *123*, 156401.

(313) Irmler, A.; Grüneis, A. *The Journal of Chemical Physics* **2019**, *151*, 104107.

(314) Nyden, M. R.; Petersson, G. A. *The Journal of Chemical Physics* **1981**, *75*, 1843–1862.

(315) Petersson, G. A.; Nyden, M. R. *The Journal of Chemical Physics* **1981**, *75*, 3423–3425.

(316) Ranasinghe, D. S.; Petersson, G. A. *The Journal of Chemical Physics* **2013**, *138*, 144104.

(317) Valeev, E. F. *Physical Chemistry Chemical Physics* **2008**, *10*, 106–113.

(318) Bartlett, R. J.; Musiał, M. *Reviews of Modern Physics* **2007**, *79*, 291–352.

(319) Peterson, K. A. Gaussian Basis Sets Exhibiting Systematic Convergence to the Complete Basis Set Limit, In *Annual Reports in Computational Chemistry*, Wheeler, R. A., Spellmeyer, D. C., Eds.; Elsevier: 2007; Vol. 3, pp 195–206.

(320) Dixon, D. A.; Feller, D.; Peterson, K. A. A Practical Guide to Reliable First Principles Computational Thermochemistry Predictions Across the Periodic Table, In *Annual Reports in Computational Chemistry*, Wheeler, R. A., Ed.; Elsevier: 2012; Vol. 8, pp 1–28.

(321) Cowan, R. D.; Griffin, D. C. *Journal of the Optical Society of America* **1976**, *66*, 1010.

(322) Martin, R. L. *The Journal of Physical Chemistry* **1983**, *87*, 750–754.

(323) Sellers, H.; Pulay, P. *Chemical Physics Letters* **1984**, *103*, 463–465.

(324) Handy, N. C.; Yamaguchi, Y.; Schaefer, H. F. *The Journal of Chemical Physics* **1986**, *84*, 4481–4484.

(325) Kutzelnigg, W. *Molecular Physics* **1997**, *90*, 909–916.

(326) Handy, N. C.; Lee, A. M. *Chemical Physics Letters* **1996**, *252*, 425–430.

(327) Karton, A.; Martin, J. M. L. *Theoretical Chemistry Accounts* **2006**, *115*, 330–333.

(328) Jensen, F. *Theoretical Chemistry Accounts* **2005**, *113*, 267–273.

(329) Klopper, W. *Molecular Physics* **2001**, *99*, 481–507.

(330) Becke, A. D. *The Journal of Chemical Physics* **1993**, *98*, 5648–5652.

(331) Lee, C.; Yang, W.; Parr, R. G. *Physical Review B* **1988**, *37*, 785–789.

(332) Becke, A. D. *Physical Review A* **1988**, *38*, 3098–3100.

(333) Feller, D. *The Journal of Chemical Physics* **1993**, *98*, 7059–7071.

(334) Grimme, S. *The Journal of Chemical Physics* **2006**, *124*, 034108.

(335) Schwabe, T.; Grimme, S. *Physical Chemistry Chemical Physics* **2006**, *8*, 4398.

(336) Santra, G.; Sylvetsky, N.; Martin, J. M. L. *The Journal of Physical Chemistry A* **2019**, *123*, 5129–5143.

(337) Zhao, Y.; Truhlar, D. G. *Theoretical Chemistry Accounts* **2008**, *120*, 215–241.

(338) Martin, J. M. *Chemical Physics Letters* **1996**, *259*, 669–678.

(339) Heckert, M.; Kállay, M.; Gauss, J. *Molecular Physics* **2005**, *103*, 2109–2115.

(340) Heckert, M.; Kállay, M.; Tew, D. P.; Klopper, W.; Gauss, J. *The Journal of Chemical Physics* **2006**, *125*, 044108.

(341) Mills, I. M. Vibration-Rotation Structure in Asymmetric- and Symmetric-Top Molecules, In *Molecular Spectroscopy: Modern Research*, Rao, K. N., Mathews, C. W., Eds.; Academic Press: New York, New York, 1972, pp 115–140.

(342) Puzzarini, C.; Heckert, M.; Gauss, J. *The Journal of Chemical Physics* **2008**, *128*, 194108.

(343) Puzzarini, C.; Cazzoli, G.; Baldacci, A.; Baldan, A.; Michauk, C.; Gauss, J. *The Journal of Chemical Physics* **2007**, *127*, 164302.

(344) Puzzarini, C. *The Journal of Physical Chemistry A* **2009**, *113*, 14530–14535.

(345) Tew, D. P.; Klopper, W.; Heckert, M.; Gauss, J. *The Journal of Physical Chemistry A* **2007**, *111*, 11242–11248.

(346) Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini, C. *Journal of Chemical Theory and Computation* **2013**, *9*, 1533–1547.

(347) Carnimeo, I.; Puzzarini, C.; Tasinato, N.; Stoppa, P.; Charmet, A. P.; Biczysko, M.; Cappelli, C.; Barone, V. *The Journal of Chemical Physics* **2013**, *139*, 074310.

(348) Martin, J. M. L.; Kesharwani, M. K. *Journal of Chemical Theory and Computation* **2014**, *10*, 2085–2090.

(349) Kállay, M.; Nagy, P. R.; Rolik, Z.; Mester, D.; Samu, G.; Csontos, J.; Csónka, J.; Szabó, B. P.; Gyevi-Nagy, L.; Ladjánszki, I.; Szegedy, L.; Ladóczki, b.; Petrov, K.; Farkas, M.; Mezei, P. D.; Hégely, B., https://www.mrcc.hu/.

(350) Kállay, M.; Nagy, P. R.; Mester, D.; Rolik, Z.; Samu, G.; Csontos, J.; Csóka, J.; Szabó, P. B.; Gyevi-Nagy, L.; Hégely, B.; Ladjánszki, I.; Szegedy, L.; Ladóczki, B.; Petrov, K.; Farkas, M.; Mezei, P. D.; Ganyecz, Á. *The Journal of Chemical Physics* **2020**, *152*, 074107.

(351) Frisch, M. J.; Trucks G., W.; Schlegel H., B.; Scuseria G., E.; Robb M., A.; Cheeseman J., R.; Scalmani, G.; Barone, V.; Petersson G., A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J., 2016.

(352) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. *WIREs Computational Molecular Science* **2012**, *2*, 242–253.

(353) 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.; Ma, Q.; Miller, T. F.; Mitrushchenkov, A.; Peterson, K. A.; Polyak, I.; Rauhut, G.; Sibaev, M. *The Journal of Chemical Physics* **2020**, *152*, 144107.

(354) Weigend, F. *Physical Chemistry Chemical Physics* **2002**, *4*, 4285–4291.

(355) Rauhut, G.; Knizia, G.; Werner, H.-J. *The Journal of Chemical Physics* **2009**, *130*, 054105.

(356) Řezáč, J.; Hobza, P. *Journal of Chemical Theory and Computation* **2013**, *9*, 2151–2155.

(357) Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini, C. *Physical Chemistry Chemical Physics* **2013**, *15*, 10094.

(358) Karton, A. *Australian Journal of Chemistry* **2019**, *72*, 607.

(359) Demaison, J.; Császár, A. G.; Dehayem-Kamadjeu, A. *The Journal of Physical Chemistry A* **2006**, *110*, 13609–13617.

(360) Baraban, J. H.; Changala, P. B.; Stanton, J. F. *Journal of Molecular Spectroscopy* **2018**, *343*, 92–95.

(361) Cook, R. L.; De Lucia, F. C.; Helminger, P. *Journal of Molecular Structure* **1975**, *28*, 237–246.

(362) Małyszek, P.; Koput, J. *Journal of Computational Chemistry* **2013**, *34*, 337–345.

(363) Huber, K. P.; Herzberg, G.; Springer United States of America: Boston, Massachusetts, 1979.

(364) Bürger, H.; Schulz, P.; Jacob, E.; Fähnle, M. *Zeitschrift für Naturforschung A* **1986**, *41*, 1015–1020.

(365) Irikura, K. K. *Journal of Physical and Chemical Reference Data* **2007**, *36*, 389–397.

(366) Temsamani, M. A.; Herman, M. *The Journal of Chemical Physics* **1995**, *102*, 6371–6384.

(367) Martin, J. M. L.; Lee, T. J.; Taylor, P. R. *The Journal of Chemical Physics* **1998**, *108*, 676–691.

(368) Strey, G.; Mills, I. *Journal of Molecular Spectroscopy* **1976**, *59*, 103–115.

(369) Saouli, A.; Bredohl, H.; Dubois, I.; Fayt, A. *Journal of Molecular Spectroscopy* **1995**, *174*, 20–50.

(370) Gershikov, A.; Spiridonov, V. *Journal of Molecular Structure* **1983**, *96*, 141–149.

(371) Strey, G. *Journal of Molecular Spectroscopy* **1967**, *24*, 87–99.

(372) Strey, G.; Mills, I. *Molecular Physics* **1973**, *26*, 129–138.

(373) Teffo, J.-L.; Chédin, A. *Journal of Molecular Spectroscopy* **1989**, *135*, 389–409.

(374) Huang, X.; Schwenke, D. W.; Lee, T. J. *The Journal of Chemical Physics* **2011**, *134*, 044320.

(375) Martin, J. M.; Taylor, P. R. *Chemical Physics Letters* **1996**, *248*, 336–344.

(376) Tarczay, G.; Miller, T. A.; Czakó, G.; Császár, A. G. *Physical Chemistry Chemical Physics* **2005**, *7*, 2881.

(377) Demaison, J.; Margulès, L.; Martin, J. M. L.; Boggs, J. E. *Physical Chemistry Chemical Physics* **2002**, *4*, 3282–3288.

(378) Wang, X.-G.; Carrington Jr, T. An accurate potential energy surface for methane, In *68th international symposium on molecular spectroscopy*, 2013, EWG10.

(379) Burleigh, D. C.; McCoy, A. B.; Sibert, E. L. *The Journal of Chemical Physics* **1996**, *104*, 480–487.

(380) Yachmenev, A.; Yurchenko, S. N.; Ribeyre, T.; Thiel, W. *The Journal of Chemical Physics* **2011**, *135*, 074302.

(381) Wang, D.; Shi, Q.; Zhu, Q.-S. *The Journal of Chemical Physics* **2000**, *112*, 9624–9631.

(382) East, A. L. L.; Allen, W. D.; Klippenstein, S. J. *The Journal of Chemical Physics* **1995**, *102*, 8506–8532.

(383) Dateo, C. E.; Lee, T. J.; Schwenke, D. W. *The Journal of Chemical Physics* **1994**, *101*, 5853–5859.

(384) Skokov, S.; Peterson, K. A.; Bowman, J. M. *The Journal of Chemical Physics* **1998**, *109*, 2662–2671.

(385) Schuurman, M. S.; Allen, W. D.; Schaefer, H. F. *Journal of Computational Chemistry* **2005**, *26*, 1106–1112.

(386) Brauer, B.; Kesharwani, M. K.; Kozuch, S.; Martin, J. M. L. *Physical Chemistry Chemical Physics* **2016**, *18*, 20905–20925.

(387) Weigend, F. *Journal of Computational Chemistry* **2008**, *29*, 167–175.

(388) Hättig, C. *Phys. Chem. Chem. Phys.* **2005**, *7*, 59–66.

(389) Tew, D. P.; Klopper, W. *The Journal of Chemical Physics* **2005**, *123*, 074101.

(390) Martin, J. M. *Chemical Physics Letters* **1995**, *242*, 343–350.

(391) Margulès, L.; Demaison, J.; Rudolph, H. D. *Journal of Molecular Structure* **2001**, *599*, 23±30.

(392) Barone, V. *The Journal of Physical Chemistry Letters* **2023**, *14*, 5883–5890.

(393) Yang, Q.; Mendolicchio, M.; Barone, V.; Bloino, J. *Frontiers in Astronomy and Space Sciences* **2021**, *8*, 665232.

(394) Zheng, J.; Zhao, Y.; Truhlar, D. G. *Journal of Chemical Theory and Computation* **2009**, *5*, 808–821.

(395) Barone, V.; Lupi, J.; Salta, Z.; Tasinato, N. *Journal of Chemical Theory and Computation* **2021**, *17*, 4913–4928.

(396) Chan, B. *Pure and Applied Chemistry* **2017**, *89*, 699–713.

(397) Nacsa, A. B.; Czakó, G. *Physical Chemistry Chemical Physics* **2021**, *23*, 9663–9671.

(398) Bazsó, G.; Tarczay, G.; Fogarasi, G.; Szalay, P. G. *Physical Chemistry Chemical Physics* **2011**, *13*, 6799.

(399) Barone, V. *Journal of Chemical Theory and Computation* **2023**, *19*, 4970–4981.

(400) Alonso, J. L.; Peña, I.; López, J. C.; Vaquero, V. *Angewandte Chemie International Edition* **2009**, *48*, 6141–6143.

(401) Ganyecz, Á.; Kállay, M.; Csontos, J. *The Journal of Physical Chemistry A* **2019**, *123*, 4057–4067.

(402) Dorofeeva, O. V.; Vogt, N. *Journal of Chemical & Engineering Data* **2009**, *54*, 1348–1352.

(403) Paulechka, E.; Kazakov, A. *Journal of Chemical Theory and Computation* **2018**, *14*, 5920–5932.

(404) Paulechka, E.; Kazakov, A. *The Journal of Physical Chemistry A* **2021**, *125*, 8116–8131.

(405) Karton, A.; Chan, B. *Journal of Chemical & Engineering Data* **2021**, *66*, 3453–3462.

(406) Roux, M. V.; Temprado, M.; Chickos, J. S.; Nagano, Y. *Journal of Physical and Chemical Reference Data* **2008**, *37*, 1855–1996.

(407) Ruscic, B.; Bross, D. H., 2022.

(408) Demaison, J.; Craig, N. C.; Groner, P.; Écija, P.; Cocinero, E. J.; Lesarri, A.; Rudolph, H. D. *The Journal of Physical Chemistry A* **2015**, *119*, 1486–1493.

(409) Groner, P.; Kukolich, S. G. *Journal of Molecular Structure* **2006**, *780-781*, 178–181.

(410) Choplin, A. *Chemical Physics Letters* **1980**, *71*, 503–506.

(411) Chiang, J. F.; Chiang, R.; Lu, K. C.; Sung, E.-M.; Harmony, M. D. *Journal of Molecular Structure* **1977**, *41*, 67–77.

(412) Knuchel, G.; Grassi, G.; Vogelsanger, B.; Bauder, A. *Journal of the American Chemical Society* **1993**, *115*, 10845–10848.

(413) Uribe, L.; Mendolicchio, M.; Municio, S.; Mato, S.; Alonso, E. R.; Alonso, J. L.; León, I.; Barone, V. *The Journal of Physical Chemistry Letters* **2025**, *16*, 6523–6532.

(414) Godfrey, P. D.; Robertson, E. G. *The Journal of Chemical Physics* **2012**, *137*, 064306.

(415) Huber, S.; Grassi, G.; Bauder *, A. *Molecular Physics* **2005**, *103*, 1395–1409.

(416) Kabir, M. H.; Kasahara, S.; Demtröder, W.; Tatamitani, Y.; Doi, A.; Katô, H.; Baba, M. *The Journal of Chemical Physics* **2003**, *119*, 3691–3698.

(417) Kisiel, Z.; Desyatnyk, O.; Pszczołkowski, L.; Charnley, S. B.; Ehrenfreund, P. *Journal of Molecular Spectroscopy* **2003**, *217*, 115–122.

(418) Rudolph, H. D.; Demaison, J.; Császár, A. G. *The Journal of Physical Chemistry A* **2013**, *117*, 12969–12982.

(419) Zdanovskaia, M. A.; Esselman, B. J.; Woods, R. C.; McMahon, R. J. *The Journal of Chemical Physics* **2019**, *151*, 024301.

(420) Thorwirth, S.; Theule, P.; Gottlieb, C. A.; McCarthy, M. C.; Thaddeus, P. *The Astrophysical Journal* **2007**, *662*, 1309–1314.

(421) Dorofeeva, O. V.; Vishnevskiy, Y. V.; Vogt, N.; Vogt, J.; Khristenko, L. V.; Krasnoshchekov, S. V.; Shishkov, I. F.; Hargittai, I.; Vilkov, L. V. *Structural Chemistry* **2007**, *18*, 739–753.

(422) Leavell, S.; Curl, R. *Journal of Molecular Spectroscopy* **1973**, *45*, 428–442.

(423) Roucou, A.; Goubet, M.; Kleiner, I.; Bteich, S.; Cuisset, A. *ChemPhysChem* **2020**, *21*, 2523–2538.

(424) Barone, V.; Lupi, J.; Salta, Z.; Tasinato, N. *Journal of Chemical Theory and Computation* **2023**, *19*, 3526–3537.

(425) Irikura, K. K.; Johnson, R. D.; Kacker, R. N.; Kessel, R. *The Journal of Chemical Physics* **2009**, *130*, 114102.

(426) Kesharwani, M. K.; Brauer, B.; Martin, J. M. L. *The Journal of Physical Chemistry A* **2015**, *119*, 1701–1714.

(427) Chan, B.; Radom, L. *Journal of Chemical Theory and Computation* **2016**, *12*, 3774–3780.

(428) Zapata Trujillo, J. C.; McKemmish, L. K. *The Journal of Physical Chemistry A* **2023**, *127*, 1715–1735.

(429) Schuurman, M. S.; Allen, W. D.; Von Ragué Schleyer, P.; Schaefer, H. F. *The Journal of Chemical Physics* **2005**, *122*, 104302.

(430) Piccardo, M.; Bloino, J.; Barone, V. *International Journal of Quantum Chemistry* **2015**, *115*, 948–982.

(431) Bloino, J.; Biczysko, M.; Barone, V. *Journal of Chemical Theory and Computation* **2012**, *8*, 1015–1036.

(432) Ayala, P. Y.; Schlegel, H. B. *The Journal of Chemical Physics* **1998**, *108*, 2314–2325.

(433) Truhlar, D. G.; Isaacson, A. D. *The Journal of Chemical Physics* **1991**, *94*, 357–359.

(434) Oliphant, N.; Adamowicz, L. *The Journal of Chemical Physics* **1991**, *94*, 1229–1235.

(435) Piecuch, P.; Oliphant, N.; Adamowicz, L. *The Journal of Chemical Physics* **1993**, *99*, 1875–1900.

(436) Piecuch, P.; Kucharski, S. A.; Bartlett, R. J. *The Journal of Chemical Physics* **1999**, *110*, 6103–6122.

(437) Kállay, M.; Szalay, P. G.; Surján, P. R. *The Journal of Chemical Physics* **2002**, *117*, 980–990.

(438) Köhn, A.; Olsen, J. *The Journal of Chemical Physics* **2006**, *125*, 174110.

(439) Klopper, W.; Noga, J.; Koch, H.; Helgaker, T. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)* **1997**, *97*, 164–176.

(440) Koch, H.; Sánchez De Merás, A.; Pedersen, T. B. *The Journal of Chemical Physics* **2003**, *118*, 9481–9484.

(441) Koch, H.; Sánchez De Merás, A. *The Journal of Chemical Physics* **2000**, *113*, 508–513.

(442) Kinoshita, T.; Hino, O.; Bartlett, R. J. *The Journal of Chemical Physics* **2003**, *119*, 7756–7762.

(443) Hino, O.; Kinoshita, T.; Bartlett, R. J. *The Journal of Chemical Physics* **2004**, *121*, 1206–1213.

(444) Lesiuk, M. *Journal of Chemical Theory and Computation* **2020**, *16*, 453–467.

(445) Mester, D.; Nagy, P. R.; Csóka, J.; Gyevi-Nagy, L.; Szabó, P. B.; Horváth, R. A.; Petrov, K.; Hégely, B.; Ladóczki, B.; Samu, G.; Lőrincz, B. D.; Kállay, M. *The Journal of Physical Chemistry A* **2025**, *129*, 2086–2107.

(446) Varandas, A. J. *Annual Review of Physical Chemistry* **2018**, *69*, 177–203.

(447) Varandas, A. J. C. *International Journal of Quantum Chemistry* **2021**, *121*, e26598.