Doctoral theses at NTNU, 2022:343

Anders Hutcheson

Coupled cluster investigations of spectroscopic properties in molecular and crystalline systems

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Natural Sciences
Department of Chemistry

NTNU
Norwegian University of
Science and Technology

Anders Hutcheson

Coupled cluster investigations of spectroscopic properties in molecular and crystalline systems

Thesis for the Degree of Philosophiae Doctor

Trondheim, November 2022

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Chemistry

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Natural Sciences
Department of Chemistry

ISBN 978-82-326-5466-6 (printed ver.)
ISBN 978-82-326-6388-0 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2022:343

Printed by NTNU Grafisk senter

Abstract

Theoretical studies of spectroscopic properties are of great value, both by themselves and in combination with experiments. The highly accurate coupled cluster methods are attractive tools for such studies. However, due to their high polynomial scaling, coupled cluster methods are limited by the system size which can be treated and also run into issues when problematic degeneracies occur. The thesis consists of three papers and is concerned with how coupled cluster can be used to obtain spectroscopic properties of systems exhibiting such problematic attributes.

In the first paper, the spectroscopic properties of photoswitches are studied, where CC3 generates ground and excited state potential energy surfaces in good agreement with high-level multireference methods. Therefore, CC3 could constitute a valuable tool for investigating photoswitches, so long as the limitations associated with degeneracies are kept in mind.

In the second paper, a cluster approach for obtaining size-intensive properties of molecular crystals is presented. The approach relies on the near-sightedness of electrons, allowing us to compute size-intensive properties of a target region even though the ground state energy is not converged with cluster size. This highlights a currently underexploited feature, since most approaches first target the energy. Local excited state properties can be obtained in this manner, which is demonstrated using coupled cluster in a reduced orbital space to compute local valence excitations in NH_3 clusters.

In the third and final paper, the cluster approach is used to generate core excitation spectra for CH_4 and CO_2 clathrates. The cluster approach appears to be well suited to obtain properties of core excited states. Further, the spectra indicate that XAS has potential for the study of such systems.

Acknowledgments

First and foremost, I thank my supervisor Ida-Marie Høyvik for her unwavering support and can-do attitude. I can confidently say that you would not be reading these words without her encouragement. I am also grateful to her for giving me the opportunity to explore a new field of research and for her guidance.

Thanks should also go to everyone who has developed the $e^{\mathcal{T}}$ program, which I use in all papers in the thesis, with special thanks going to the maintainers of the code, Eirik F. Kjønstad, Sarai D. Folkestad, Rolf H. Myhre and Alexander C. Paul.

I'm also grateful for all my colleagues who have made the last four years a great experience. Thanks for making the work environment fun, organizing events, and always having time to help.

For proofreading, I would like to thank Regina Matveeva, Alexander C. Paul, and Sarai D. Folkestad. Additionally, I would like to thank Sarai D. Folkestad and Alexander C. Paul for the LaTeX template used for this thesis.

Finally, I would like to thank my family for their support and Tina Danielsen for her love and for putting up with me through these four years.

Contents

1 Introduction

The field of spectroscopy is concerned with how matter interacts with or generates electromagnetic radiation. This gives rise to nondestructive methods with the capability to investigate various properties of matter. Spectroscopic techniques that probe electronic excitations from core orbitals require X-ray radiation. X-ray absorption spectroscopy (XAS) is one such technique that can be used to investigate the local chemical environment surrounding a given element. Excitations from valence orbitals can be probed with ultraviolet or visible (UV-VIS) radiation, with UV-VIS spectroscopy being widely used to study transition metal compounds and conjugated organic compounds. However, interpreting spectroscopic data is not always straightforward, and theoretical studies can be of great value in this respect. The need for theoretical studies increases as more complex spectroscopic methods are developed. The appropriate choice of theoretical method to investigate spectroscopic properties is, however, dependent on the system and type of spectroscopy.

In cases where a single-reference description is reasonable, coupled cluster methods are among the most accurate options.[1,2] Coupled cluster methods have several attractive properties, such as accounting for dynamical correlation in a compact fashion, offering systematic improvement by changing truncation level, and being black box methods.[3] Further, coupled cluster methods are well suited for the study of excited state properties, where both core[4–8] and valence[9–13] excitations can be accurately represented. However, the steep polynomial scaling of even low-level coupled cluster methods limits the system size which can be

treated, which has motivated the development of low-scaling coupled cluster methods.[14–31] Coupled cluster methods also run into problems when treating degenerate or near-degenerate states. In particular, this is a problem for ground state degeneracies since a single-reference description becomes unsuitable.

Photoswitches pose a challenge for coupled cluster methods, with respect to their size and the frequent occurrence, and importance, of degenerate states and conical intersections.[32,33] Since multireference methods are apt at describing degenerate states and conical intersections, the majority of photoswitch studies are conducted using such methods.[34–51] The question is then why would we use coupled cluster for the study of photoswitches? In paper A we try to answer this question by investigating the ability of CC2,[52] CCSD[53] and CC3[54] to generate ground and excited potential energy surfaces of two photoswitches, a retinal model system (PSB3) and azobenzene. CC3 was found to produce accurate potential energy surfaces in agreement with high-level multireference methods, even for regions close to a conical intersection. Regions dominated by states of doubly excited character were well described already at the CC3 level of theory. Since coupled cluster methods are black box, CC3 has an advantage over multireference methods in that active space selection is avoided. Therefore, there is merit to using coupled cluster for the study of photoswitches, despite the frequent occurrence of degeneracies for such systems. Particularly CC3 is shown to be a robust and versatile choice.

A central concept for reducing the computational cost of electronic structure methods is the "nearsightedness" of electrons in many-atom systems.[55] The concept explains that local electronic properties depend significantly on the effective external potential only at nearby points. This is valid at the Hartree-Fock (HF) level of theory and allows size-intensive properties to be obtained for a target region. However, to compute size-intensive properties, the electronic density of the target region must be converged, which can be achieved if sufficiently large parts of the surroundings are included. An important caveat is that, unlike the electronic density, the energy of the target region converges slowly as larger parts of the surroundings are included. Therefore, the traditional requirement of converging the ground state energy prior to computing properties must be abandoned for such an approach. The ability to compute size-intensive

properties of a target region without converging the energy is currently an underexploited feature, as most approaches first target the energy. Conversely, the locality of electron correlation is exploited to a much larger extent, as it is fundamental in several low-scaling electronic structure models.[14–30,56–60]

The ability to obtain size-intensive properties of a target region without converging the energy is particularly useful for molecular crystals. A cluster approach can then be used, thereby avoiding a periodic approach. A cluster approach allows for different basis sets to be used for different regions of the cluster. This can alleviate problems of orthogonalization tails and near-linear dependencies associated with periodic wave function models when high-level basis sets are used. Further, any of the aforementioned low-scaling electronic structure models can be used to compute size-intensive properties of the target region, so long as one does not aim to converge the energy with cluster size. In addition to these methods, coupled cluster performed in a reduced orbital space is an attractive option when properties of local excitations are of interest.

In paper B we use a cluster model of crystalline NH_3 to illustrate that the HF electronic density of a target region converges with cluster size, and as a result, size-intensive properties of the target region are obtained. The properties computed are HF dipole moments and CC2-in-HF excitation energies of a local excitation. The approach was extended to more chemically interesting systems in paper C, where core excitation spectra of CH_4 and CO_2 clathrates were generated. The evaluation of the core excitation spectra indicates that XAS may constitute a valuable tool for investigating gas clathrates and thereby expand the range of spectroscopic methods for the characterization of gas clathrates.

The thesis is structured as follows. Firstly, the relevant theory is introduced, followed by chapters presenting the papers included in the thesis. Finally, concluding remarks and an outlook are given.

1.1 List of papers included in this thesis

Paper A:
Describing ground and excited state potential energy surfaces for molecular photoswitches using coupled cluster models

A. Hutcheson, A. C. Paul, R. H. Myhre, H. Koch, I-M. Høyvik,
J. Comput. Chem. **2021**, 42(20), 1419, doi: 10.1002/jcc.26553

Paper B:
Convergence of the electronic density for a target region in cluster models of a NH_3 molecular crystal
A. Hutcheson, I-M. Høyvik,
J. Math. Chem. **2022**, doi: 10.1007/s10910-022-01351-w

Paper C:
Coupled cluster core excitation spectra of methane and carbon dioxide clathrates
A. Hutcheson, I-M. Høyvik,
In preparation

1.2 Other publications

Other publications in the field, not included in this thesis.

eT 1.0: An open source electronic structure program with emphasis on coupled cluster and multilevel methods
S. D. Folkestad, E. F. Kjønstad, R. H. Myhre, J. H. Andersen, A. Balbi, S. Coriani, T. Giovannini, L. Goletto, T. S. Haugland, A. Hutcheson, I-M. Høyvik, T. Moitra, A. C. Paul, M. Scavino, A. S. Skeidsvoll, Å. H. Tveten, and H. Koch,
J. Chem. Phys. **2020**, 152, 184103, doi: 10.1063/5.0004713

2 Theoretical background

2.1 Hartree-Fock theory

The Hartree-Fock model (HF)[61,62] is widely used in computational chemistry, on its own and as the foundation of other methods. In HF, the wave function is represented as a single Slater determinant,[63] constructed from molecular orbitals (MOs). The energy is given as the expectation value of the electronic Hamiltonian

$$E = \langle \Psi | H_e | \Psi \rangle \,, \quad H_e = -\frac{1}{2}\sum_{i}^{N_e} \nabla_i^2 - \sum_{i}^{N_e}\sum_{I}^{N_n} \frac{Z_I}{r_{Ii}} + \frac{1}{2}\sum_{i\neq j}^{N_e} \frac{1}{r_{ij}}, \tag{2.1}$$

where $r_{Ii} = |\boldsymbol{R_I} - \boldsymbol{r_i}|$ and $r_{ij} = |\boldsymbol{r_i} - \boldsymbol{r_j}|$, with i, j referring to electrons, capital indices, I, referring to nuclei, and Z_I being the charge of nucleus I. The objective of HF is to minimize the energy with respect to MOs, with the constraint that the MOs remain orthogonal. This can be achieved by the method of Lagrange multipliers, where the Lagrange function is stationary with respect to an orbital variation at convergence

$$\delta L = \delta E - \sum_{ij}^{N_{\text{elec}}} \lambda_{ij}(\langle \delta\phi_i | \phi_j \rangle - \langle \phi_i | \delta\phi_j \rangle) = 0, \tag{2.2}$$

where λ_{ij} are the Lagrange multipliers and ϕ_i and ϕ_j are MOs. The MOs are typically expanded in a finite basis of atomic orbitals (AOs), and the minimization can be formulated as the Roothaan-Hall equations[64,65]

$$\mathbf{F}^{AO}\mathbf{C} = \mathbf{S}\mathbf{C}\boldsymbol{\epsilon}, \tag{2.3}$$

where $\mathbf{C}$ is the orbital coefficients matrix, $\mathbf{S}$ is the overlap matrix, $\mathbf{F}^{AO}$ is the Fock matrix in the AO basis, and $\boldsymbol{\epsilon}$ is a diagonal matrix with the orbital energies on its diagonal. Alternatively, the Fock matrix in the MO basis can be used. Its elements are given as

$$F_{pq} = (\mathbf{C}^{\mathbf{T}}\mathbf{F}^{\mathbf{AO}}\mathbf{C})_{pq} = h_{pq} + G_{pq}(\mathbf{D}), \tag{2.4}$$

where $\mathbf{D}$ is the HF electronic density matrix, which in the orthogonal MO basis is given as

$$\mathbf{D} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \tag{2.5}$$

where the occupied-occupied block is $\mathbf{I}$, and all other blocks are $\mathbf{0}$. The density matrix can be converted to the AO basis by $\mathbf{D}^{AO} = \mathbf{C}\mathbf{D}\mathbf{C}^{\mathbf{T}}$. The elements of the one electron, h_{pq}, and two-electron matrices, G_{pq}, are given by

$$h_{pq} = \int \phi_p^*(\mathbf{r}_1)\Big(-\frac{1}{2}\nabla^2 - \sum_I \frac{Z_n}{r_{1I}}\Big)\phi_q^*(\mathbf{r}_1)d\mathbf{r}_1 \tag{2.6}$$

$$G_{pq} = \sum_{ij}(2g_{pqij} - g_{pjiq})D_{ij}, \tag{2.7}$$

where two-electron integrals are here given in Mulliken notation

$$g_{pqij} = \int\int \phi_p^*(\mathbf{r}_1)\phi_q(\mathbf{r}_1)\frac{1}{r_{12}}\phi_i^*(\mathbf{r}_2)\phi_j(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2. \tag{2.8}$$

Since the Fock matrix depends on the orbital coefficients, the Roothaan-Hall equations must be solved iteratively, which at convergence results in the canonical HF MOs. HF is also called the mean-field approximation since the electrons feel the average field generated by all other electrons. This approximation leads to a deviation from the exact electronic energy, where the energy not accounted for at the HF-limit is termed the correlation energy

$$E_{\text{correlation}} = E_{\text{exact}} - E_{\text{HF}}. \tag{2.9}$$

The HF-limit is the theoretical limit obtained by using an infinite basis set. Therefore, there is an additional error in the energy arising from the use of finite basis sets. The HF-limit can be approached by increasing the size of the basis set, while to account for electron correlation, post-HF methods must be used.

2.2 Coupled cluster theory

In coupled cluster theory, an exponential parametrization is used

$$|\text{CC}\rangle = e^T |\text{HF}\rangle , \quad T = \sum_\mu t_\mu \tau_\mu, \tag{2.10}$$

where T is the cluster operator, which generates excited determinants, $|\mu\rangle$, and is a linear combination of excitation operators, τ_μ, multiplied with associated cluster amplitudes, t_μ.[1]

By including all excitation operators, the exact energy can be obtained. However, including all excitation operators scales factorially with system size. For this reason, the cluster operator is partitioned into classes comprising all single excitations, double excitations, and up to N-electron excitations[1]

$$T = T_1 + T_2 + \cdots + T_N. \tag{2.11}$$

This gives rise to a hierarchy of methods truncated at a given excitation level, where truncating at double excitations gives CCSD,[53] while truncating at triple excitations gives CCSDT,[66] and so on. The hierarchy of methods allows for systematic improvement of the description, but at an increased computational cost. In addition to the standard coupled cluster methods, there also exist perturbative coupled cluster methods, such as CC2[52] and CC3,[54] where the highest order excitation is treated perturbatively. Because of the perturbative treatment, the computational scaling of CC2 and CC3 is one order of magnitude lower than the scaling of CCSD and CCSDT, respectively.

Due to the exponential parametrization of the coupled cluster wave function, variational optimization, which is commonly used for linear parametrizations, becomes complicated.[1] For this reason, subspace projection is used, and the projected coupled cluster equations are solved instead

$$E_{CC} = \langle HF| \bar{H} |HF\rangle \tag{2.12}$$

$$\Omega_\mu = \langle \mu| \bar{H} |HF\rangle , \tag{2.13}$$

where the similarity transformed Hamiltonian, $\bar{H} = e^{-T} H e^T$, is used. By using the exponential parametrization and projecting, CC methods are

size-extensive and have a compact description of electron correlation. The compact description of electron correlation originates from the inclusion of higher-order excitations through disconnected terms of the cluster operator.[1]

However, coupled cluster methods run into problems when states become degenerate or nearly degenerate. In the case of degeneracies between excited states, the problems of coupled cluster methods arise for conical intersections between electronic states of the same symmetry.[67–69] Excluding these features, coupled cluster methods are generally able to describe near-degeneracies between excited states.[70] A more fundamental problem arises when the ground state becomes degenerate, since the HF approximation breaks down. The spin-flip coupled cluster methods were developed to alleviate this problem by exploiting the fact that a high-spin triplet reference may be well behaved, unlike the singlet reference.[71–73] Singlet ground and excited states can then be generated by spin-flip excitations from the triplet reference. However, the more common choice is to use multireference methods. Multireference methods include multiple Slater determinants (or configuration state functions) as the reference, where the expansion coefficients of the determinants and the MOs are optimized simultaneously. This simultaneous optimization is computationally demanding and is therefore often restricted to an active space. The selected active space needs to be able to represent the important determinants for the system, which is difficult to ensure. In coupled cluster theory, all orbitals are included in the parametrization, and coupled cluster methods are therefore considered to be black box methods.

Spectroscopic properties

Coupled cluster methods are well suited to investigate spectroscopic properties, where both core[4–8,74] and valence[9–12] excitations can be accurately represented. One can obtain excitation energies from coupled cluster calculations, which are related to the absorption maxima in experimental spectra, and oscillator strengths, which are comparable to the intensities of experimental spectra.

One of the most common approaches to obtain these quantities is the equation of motion (EOM) formalism.[75–77] In this formalism, the similarity

transformed Hamiltonian is represented in the basis of the HF and excited determinants, and we obtain the Hamiltonian matrix with elements

$$\bar{H}_{\mu\nu} = \langle \mu | \bar{H} | \nu \rangle \,. \tag{2.14}$$

The EOM states and their total energies are obtained as the eigenvectors and eigenvalues of this Hamiltonian matrix. The left and right eigenvectors will not be equal due to $\bar{H}$ being non-symmetric. Therefore, we denote left and right ground states as $\langle \Lambda |$, $|CC\rangle$, and the left and right excited states as $\langle L |$, $|R\rangle$. In the EOM formalism, oscillator strengths, f, are defined as

$$f = \langle \Lambda | d | R \rangle \;\; \langle L | d | CC \rangle \,, \tag{2.15}$$

where d is the dipole operator.

The appropriate choice of coupled cluster method depends on the system and type of spectroscopy of interest. For excited states with doubly excited character, which are important for organic photochemistry, coupled cluster methods including at least triple excitations are required for an accurate description.[78] When using coupled cluster methods including triple excitations, the fraction of singles amplitudes can be used to characterize the excited state and determine whether a lower-level truncation might be sufficient to describe the state.[3] Further, the inclusion of triple excitations has also been shown to be important to account for relaxation effects when describing core excited states.[7,8] In addition to coupled cluster method, the basis set and the geometry of the systems also affect the computed spectroscopic properties. This makes obtaining accurate spectroscopic properties all the more difficult.

There are also complicating factors associated with comparing computed quantities with experiments. Firstly, comparing vertical excitation energies from computations with energies at band maxima from experiments is a comparison with certain flaws.[79–81] One of the flaws is that the experimentally probed excitations might not be vertical. Alternatively, the comparison can be made between 0-0 transition energies,[79,80] which are uniquely defined experimentally and theoretically, allowing for a direct comparison.[3] However, the computational costs of computing 0-0 transition energies are higher than for vertical excitation energies and might even be prohibitively expensive. In addition, the effect of temperature leads to broadening of peaks in experiments, which must be accounted for. Further, some experiments are conducted in a solvent which

can affect excitation energies and intensities directly through electrostatic interaction with the excited state, but also indirectly, e.g., through changes of the geometry.

2.3 Locality and reduced computational scaling

The concept of locality is important in electronic structure theory. The locality of canonical MOs is by construction poor, with the MOs being delocalized across entire molecular systems. However, the canonical MOs are just one choice out of infinitely many possibilities. Other MOs can be obtained by unitary transformations, where occupied-occupied or virtual-virtual rotations can be done freely, since the energy is invariant under such rotations. One class is local MOs, where explicit localization of occupied MOs has been a popular topic for many decades, with seminal contributions such as the widely used Edmiston-Ruedenberg[82], Foster-Boys[83], and Pipek-Mezey[84] localization functions. In addition to explicit localization, several approaches for generating local virtual spaces exist, such as projected atomic orbitals (PAOs)[85], pair-natural orbitals[86–88] and, in the case of local excitations, correlated natural transition orbitals (CNTOs).[89]

Locality is also fundamental in reducing the computational cost of electronic structure methods. In 1996 Kohn[55] introduced the term "nearsightedness" of electrons in many-atom systems, which explains that local electronic properties depend significantly on the effective external potential only at nearby points. The concept is valid at the HF level of theory, where size-intensive properties can be obtained for a target region. However, to compute size-intensive properties, the electronic density of the target region must be converged, which can be achieved if sufficiently large parts of the surroundings are included. Others have also discussed the convergence of the electronic density for a target region.[90–93] An important caveat is that, unlike the electronic density, the energy of the target region converges slowly as larger parts of the surroundings are included. Therefore, the traditional requirement of converging the energy prior to computing properties must be abandoned to make such an approach viable. Being able to compute size-intensive properties of a target region without converging the energy is currently an underexploited feature, as most approaches first target the energy. Conversely, the lo-

cality of electron correlation is exploited to a much larger extent, as it is fundamental in several low-scaling electronic structure models,[14–30,56–60] where local MOs are frequently used.

The ability to obtain size-intensive properties of a target region without converging the energy is particularly useful for molecular crystals, where a cluster approach can be used, thereby avoiding a periodic approach. A cluster approach allows for different basis sets to be used for different regions of the cluster, which can alleviate problems of orthogonalization tails and near-linear dependencies. Orthogonalization tails can lead to slow convergence of the correlation energy in local correlation approaches.[94,95] Near-linear dependencies severely affect the locality of the virtual space.[96] In addition, dealing with near-linear dependencies in a periodic framework may compromise the basis set quality since removing AOs from a unit cell will remove AOs from all unit cells.

Another benefit of a cluster approach is that any of the aforementioned low-scaling electronic structure models can be used to compute size-intensive properties of the target region. The only requirement is that one does not aim to converge the energy with cluster size. In addition to these methods, coupled cluster performed in a reduced orbital space is an attractive option when properties of local excitations are of interest. This can either be done with standard coupled cluster methods or reduced space coupled cluster methods[31,97–103] treating an active orbital space while the remainder of the orbital space is kept frozen at the HF-level. The interactions between the coupled cluster region and the HF regions are then accounted for through contributions to the Fock matrix. Common approaches for selecting the active orbital space are to either use local/semilocal orbitals or to use information from lower-level methods to determine orbitals of importance.

Partitioning of the HF electronic density matrix

To evaluate the HF electronic energy and density of a target region in cluster models, a partitioning of the HF electronic density matrix is useful. For brevity, the HF electronic density matrix will hereafter be referred to as the density matrix. The density matrix of the cluster can be partitioned as follows, $\mathbf{D} = \mathbf{D}_t + \mathbf{D}_r$, where $\mathbf{D}_t$ is the density matrix of a target region, and $\mathbf{D}_r$ is the density matrix of the remainder.[104] All three density

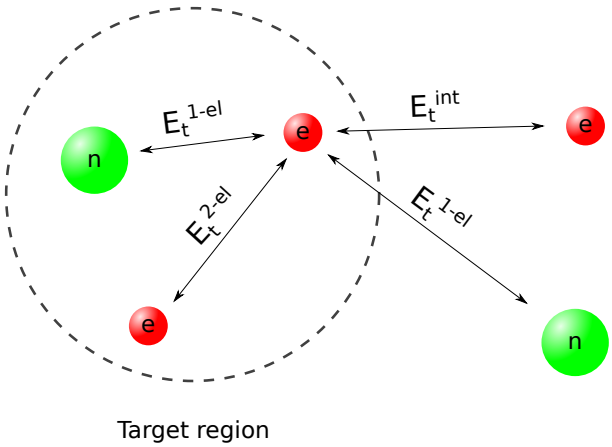

Figure 2.1: A schematic representation of the energy contributions, $E_t^{\text{1-el}}$, $E_t^{\text{2-el}}$, and E_t^{int}, making up the electronic energy of the target region, E_t. In addition to the interactions shown, $E_t^{\text{1-el}}$ also includes the kinetic energy of the electrons. Electrons and nuclei are shown as red and green circles, respectively.

matrices, $\mathbf{D}$, $\mathbf{D}_t$ and $\mathbf{D}_r$ satisfy the trace, symmetry, and idempotence criteria of density matrices representing Slater determinants, i.e., $\mathbf{D}_t$ and $\mathbf{D}_r$ can be viewed as constructed from separate subsets of orthogonal MOs describing $\mathbf{D}$. The electronic energy for the target region described by $\mathbf{D}_t$, interacting with the rest of the cluster described by $\mathbf{D}_r$ is given by

$$\begin{aligned} E_t &= \mathrm{Tr}[\mathbf{h}\mathbf{D}_t] + \frac{1}{4}\mathrm{Tr}[\mathbf{D}_t\mathbf{G}(\mathbf{D}_t)] + \frac{1}{2}\mathrm{Tr}[\mathbf{D}_t\mathbf{G}(\mathbf{D}_r)] \\ &\equiv E_t^{\text{1-el}} + E_t^{\text{2-el}} + E_t^{\text{int}}. \end{aligned} \tag{2.16}$$

We define $E_t^{\text{1-el}}$, $E_t^{\text{2-el}}$, and E_t^{int} as the one-electron, two-electron, and interaction contributions, respectively. A schematic representation of these terms is given in Figure 2.1. The electronic part of the Hartree-Fock energy of the entire cluster is given by

$$E = E_t + \mathrm{Tr}[\mathbf{h}\mathbf{D}_r] + \frac{1}{4}\mathrm{Tr}[\mathbf{D}_r\mathbf{G}(\mathbf{D}_r)], \tag{2.17}$$

i.e., the energy for the target region, E_t, plus the contributions which only depend on $\mathbf{D}_r$. Note that all two-electron interactions between the target region and the rest of the cluster are included in E_t.

As is well-known, the energy of a target region in an infinite (or similarly large) system is not a local quantity, even if $\mathbf{D}_t$ is localized in space. The reason for this is long-range interactions between the local target electronic density and nuclei and between the local target electronic density and the electronic density of the rest of the system.

2.4 Basis set considerations

The choice of basis set is important to achieve accurate results but also to keep the computational cost at an affordable level. The requirements on the basis set depend on the system and properties of interest. The smallest basis sets, so-called minimal basis, include a single basis function for each occupied orbital of the atoms making up the system. A minimal basis is too crude to accurately describe molecular properties, where larger and more flexible basis sets are needed, but may adequately describe long-range electronic interactions. In the case of post-HF methods, the basis sets should be able to account for the correlation energy efficiently. The correlation-consistent basis sets[105] were developed with this goal in mind and are frequently used for post-HF methods. However, the standard correlation-consistent basis sets were designed to accurately represent a valence correlated wave function of neutral ground state systems and therefore lack the flexibility to describe core correlation and delocalized excitations.[1] These deficiencies can both be ameliorated. To describe delocalized excitations, basis sets augmented with diffuse functions can be used.[106] The use of such augmented basis sets have been shown to be important to obtain accurate vertical excitation energies and oscillator strengths.[12] To account for core correlation, basis sets augmented with tight correlating orbitals have been developed.[107]

3 Coupled cluster as a tool for the study of photoswitches

Photoswitches constitute a class of molecular machines that undergo reversible isomerization upon stimulus by light, so-called photoisomerization. Computational approaches for photoisomerization processes are considered to be of great value for the study of photoswitches, where an accurate description of the ground and excited states is essential. In paper A we investigate the merits of using coupled cluster methods to study photoswitches, specifically a retinal model system (penta-2,4-dieniminium cation, PSB3) and azobenzene.

The use of coupled cluster methods is motivated by the attractive properties described in Section 2.2. Their black box nature gives a clear advantage over multireference methods, where user-specified orbital spaces are needed. The main counterargument to using coupled cluster methods is their inability to treat ground state degeneracies and the problems arising at conical intersections between excited states of the same symmetry.[67–69] Although this argument is valid, we assume that such features only make up small regions of potential energy surfaces. Therefore, coupled cluster methods should be able to accurately treat large parts of potential energy surfaces.

This assumption holds for the investigated potential energy surfaces of PSB3 and azobenzene. The performance of CC3 is particularly impressive; generating potential energy surfaces for PSB3 in good agreement with surfaces generated with a high-level multireference method[44] and

giving qualitatively correct potential energy surfaces for azobenzene,[48] even for a state of doubly excited character. In the case of PSB3, the potential energy surfaces are in the vicinity of a conical intersection between the ground and first excited state. This shows that coupled cluster methods including triple excitations can describe potential energy surfaces reasonably close to such features, even if they cannot handle ground state degeneracies. Therefore, coupled cluster methods, and especially methods with triple excitations included, may constitute a valuable tool for investigating photoswitches. In the following sections, relevant background and motivation are presented, followed by the study's main findings. Finally, a summary and discussion are given.

3.1 Background and motivation

In the study of photoswitches, accurate descriptions of the potential energy surfaces are paramount. Due to the occurrence of conical intersections[32] and near-degenerate states, multireference methods[108–119], which are apt at describing such features, have been used extensively.[34–51,120] Single-reference methods have been used to a lesser extent for the same reason.[121–127]

Although degeneracies and near-degeneracies are problematic for coupled cluster methods (see Section 2.2), their compact description of dynamical correlation and black box nature makes them well suited to describe all but these problematic features. For excited electronic states dominated by single-electron excitations, low-level truncations of the hierarchy, e.g., CCSD[53] or approximate versions such as CC2,[52] can be used. However, for electronic states with doubly excited character, coupled cluster models with at least triple excitations[54,66,78,128–130] are needed.

However, due to the high computational cost of including triple (or approximate triple) excitations, there are few such results for potential energy surfaces of photoswitches in the literature. This motivated us to study PSB3. For this system, literature values are available for three paths on the potential energy surface, obtained with spin-flip CCSD with perturbative triples corrections (SF-CCSD(dT))[131] and the multireference configuration interaction with singles, doubles, and a Davidson-type correction (MRCISD+Q).[44] The three paths are called bond length alternation

(BLA), covalent/diradical (DIR) and charge-transfer (CT) paths, with details given in paper A. This allows us to determine how well coupled cluster methods agree with MRCISD+Q. Furthermore, the effect of a high-spin reference could be determined by comparing CC3 with SF-CCSD(dT).

While PSB3 is an ideal system to evaluate the performance of coupled cluster methods, we also want to investigate a more realistic system. Azobenzene is a prime candidate, partly due to the substantial number of computational studies which have been conducted for azobenzene,[35,38–40,42,45,48–51,132–135] but also due to the lack of high-level single-reference calculations including at least triple excitations. From Ref. 50 a complete active space self-consistent field (CASSCF) optimized rotational path for azobenzene is available, which we use to evaluate the performance of coupled cluster methods. The excitation energy presented in Ref. 50 for equilibrium structures deviate significantly from experimental results.[136] This inspired our investigation of the factors affecting the excitation energy of azobenzene equilibrium structures.

3.2 Main findings

For PSB3, the coupled cluster results are evaluated based on the vertical excitation energies and shape of the potential energy surfaces relative to MRCISD+Q results. To assess the shape of surfaces generated by coupled cluster models, the nonparallelity value[131] is computed for the ground (S_0) and first excited singlet state (S_1) surfaces for all paths. The nonparallelity value is given by

$$\Delta E_{max} - \Delta E_{min}, \tag{3.1}$$

where ΔE_{max} is the maximum and ΔE_{min} is the minimum energy difference for a potential energy surface computed in two different manners.

Our study shows that CC2 and CCSD yield an overall poor agreement with MRCISD+Q. For CCSD, large deviations in vertical excitation are observed for all paths, and differences in the shape of the ground and first excited state are present for the CT and DIR paths. This can be seen in Figure 3.1, where the potential energy surfaces for the CT path are shown, with the surfaces for the DIR path being qualitatively the same. The only aspect in which CCSD performs well is in reproducing the shape of the

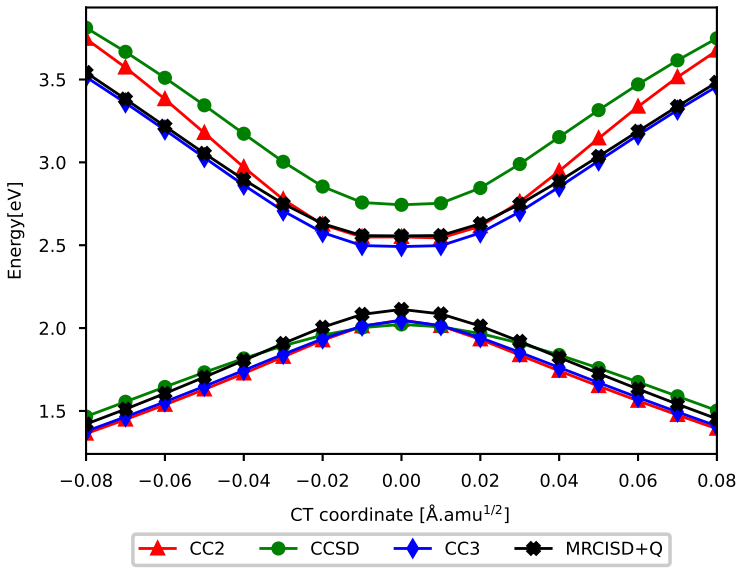

Figure 3.1: S_0 and S_1 energies for CT-path relative to cis-PSB3 computed with CC2, CCSD and CC3 with 6-31G*. MRCISD+Q with 6-31G* benchmark results from Ref. 44 are included for comparison.

ground- and first excited state for the BLA path. This can be seen in Figure 3.2, where the potential energy surfaces for the BLA path are shown. Nevertheless, even with the shape of the surfaces being reproduced, the position of the conical intersection is not. CC2 cannot reproduce the shape of the first excited state, while the shape of the ground state is in good agreement for all paths. However, the improved agreement of CC2 relative to CCSD for the ground state is most likely due to error cancellations since CC2 is an approximation of CCSD. The variable performance of CC2 is also reflected in the vertical excitation energies, which only agree well for parts of the paths.

CC3 yields the best overall agreement with MRCISD+Q, both with respect to the shapes of the surfaces and vertical excitation energies. CC3 also reproduces the conical intersection's position, as seen in Figure 3.2. CC3 is also found to perform similarly to SF-CCSD(dT).[131] Therefore, we conclude that a well-behaved triplet reference state is not required to describe these PSB3 paths accurately.

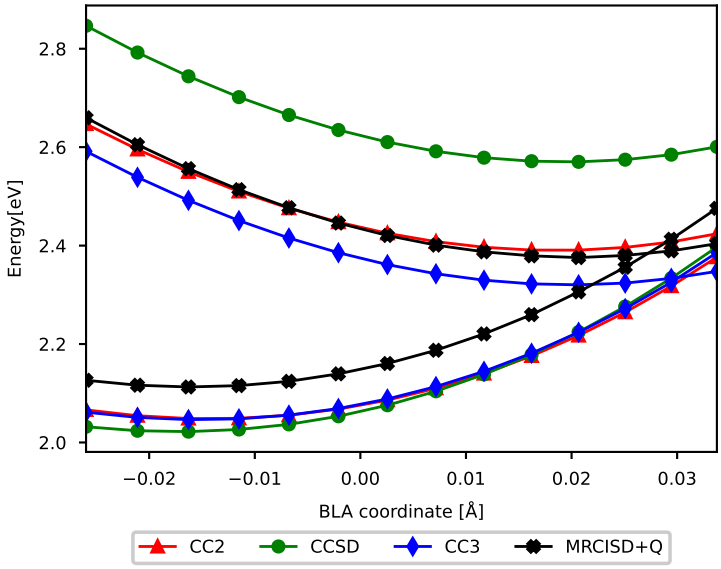

Figure 3.2: S_0 and S_1 energies for BLA-path relative to cis-PSB3 computed with CC2, CCSD and CC3 with 6-31G*. MRCISD+Q with 6-31G* benchmark results from Ref. 44 are included for comparison.

For the cis to trans CNNC-rotational path of azobenzene obtained from Zhu et al.,[50] CC3 is used to generate potential energy surfaces of the ground state (S_0) and two lowest excited singlet states (S_1, S_2). Figure 3.3 shows the resulting potential energy surfaces. Comparing CC3 and the 5SA-CASSCF(6,6) results of Ref 50, there is a qualitative agreement for the ground and lowest excited state, while differences are present for the second excited state. The most striking difference is the energy barrier at a CNNC angle of approximately 150°, which is present for CC3, but absent for CASSCF. CC3 amplitudes indicate that the ordering of states changes along the rotational path. Close to the equilibrium structures, the second excited state is the $\pi\pi^*$ state, while the $n^2\pi^{*2}$ state is the second excited state at twisted geometries. These results are consistent with the CASPT2 study in Ref. 48. The $\pi\pi^*$ state is already well described at the CC2 level of theory, but triple excitations are clearly required to reproduce the $n^2\pi^2$ state qualitatively. This is reflected in the fraction of singles amplitudes, which is high for the $\pi\pi^*$ state and low for the $n^2\pi^{*2}$ state.

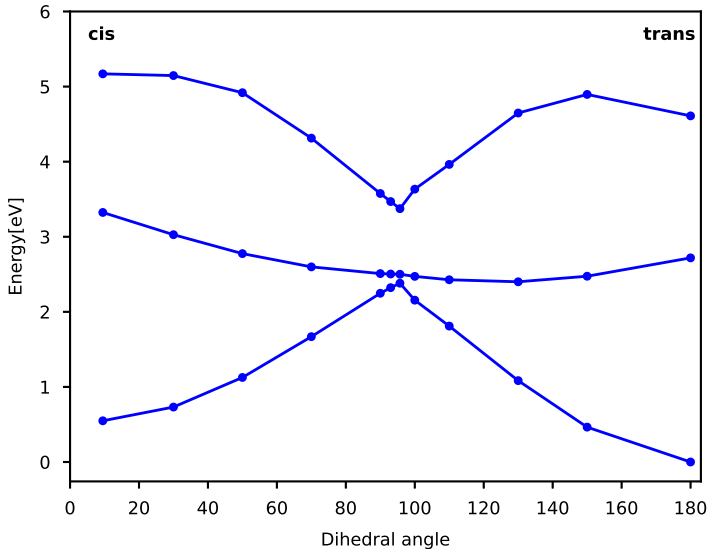

Figure 3.3: Ground state (S_0) and the two lowest excited singlet states (S_1, S_2) calculated with CC3 for a selected set of azobenzene structures along the rotational path from Ref. 50.

CC2 and CC3 are used to investigate the factors influencing the vertical excitation energy for equilibrium structures of azobenzene. Comparisons are also made with absorption maxima obtained from UV-VIS spectroscopy.[136] The factor that has the greatest impact on vertical excitation energies is the molecular geometry, where the comparison is made between the CASSCF optimized structures from Ref 50 and MP2 optimized structures. In addition, the basis set used when calculating the vertical excitation energies must be of sufficient quality. The inclusion of diffuse functions is shown to be of particular importance. The least important factor is the method used, with CC2 and CC3 yielding similar results. However, for the theoretically best result, there are differences compared to experimental absorption maxima, even when a polarizable continuum model (PCM) was used to account for solvent effects. These differences might arise from inaccuracies in the experiment or any of the previously discussed factors affecting the computations. In addition, inherent errors in comparing vertical excitation energies and absorption maxima exist, as

was discussed in Section 2.2. This illustrates how involved obtaining accurate vertical excitation energies is and the difficulties in comparing with experimental results.

3.3 Summary and discussion

Most computational studies of photoswitches are conducted with multireference methods due to the occurrence of conical intersections and near-degenerate states. While coupled cluster methods are unsuitable to describe such features, their compact description of dynamical correlation and black box nature makes them well suited to describe all but these problematic features. The applicability of coupled cluster methods is shown for PSB3 and azobenzene. The inclusion of triple excitations is needed to obtain high-quality potential energy surfaces for PSB3, where CC3 is used in our study. The potential energy surfaces are in the vicinity of a conical intersection, which shows that coupled cluster methods can yield accurate results, even close to problematic features. The inclusion of triple excitations is also important to describe states of doubly excited character, e.g., the $n^2\pi^{*2}$ of azobenzene, where even a qualitatively correct description might be unobtainable without triple excitations. Since states of doubly excited character are important for organic photochemistry, methods used to study photoswitches should be able to handle such states.

In general, CC3 is a robust and versatile method and can be used to gain information about the potential energy surfaces and may indicate prospective isomerization reaction paths of photoswitches. One could envisage that CC3 could help ensure that a suitable orbital space is being used for a multireference method and also determine whether a multireference method includes sufficient dynamical correlation. A preliminary CC3 study could be conducted before more complicated excited state dynamics approaches are pursued. However, to treat photoswitches larger than azobenzene with CC3, reductions in computational costs are needed.

Several factors must be considered to obtain accurate vertical excitation energies. This is investigated for cis- and trans-azobenzene, where the factors in decreasing order of importance are: molecular geometry, basis set, and methods (CC2 or CC3). Neglecting any of these factors may

extend beyond inaccurate vertical excitation energies, e.g., yielding potential energy surfaces of poor quality. In the case of molecular dynamics studies, potential energy surfaces of poor quality may lead to incorrect reaction paths being predicted, as well as inaccurate kinetic parameters. It is particularly problematic when states are of different accuracy. Therefore, such factors should be given sufficient attention also for molecular dynamics studies, where the focus is often on correctly describing the multireference character.

When there is a lack of high-level computational results for comparison, determining the accuracy of computed vertical excitation energies is further complicated. This is the case for azobenzene, and experimentally obtained energies at absorption maxima are used for comparison. However, differences are observed, even for the theoretical best result. Whether the differences arise due to computational factors, experimental factors, or inherent flaws in the comparison is difficult to determine. Regardless of the origin of the differences, the comparison illustrates well the challenges in comparing computed vertical excitation energies with experimentally obtained energies at absorption maxima.

4 Size-intensive properties of molecular crystals

The main obstacle limiting coupled cluster methods is their steep computational scaling. For instance, the computations presented for azobenzene in the previous chapter are already approaching the limits of what CC3 can handle in a reasonable amount of time. Therefore, much effort has been devoted to reduce the computational scaling of coupled cluster methods and electronic structure methods in general.

A central concept for reducing the computational cost of electronic structure methods is the "nearsightedness" of electrons in many-atom systems.[55] The concept explains that local electronic properties depend significantly on the effective external potential only at nearby points. This makes size-intensive properties of a target region obtainable, so long as the electronic density of the target region is converged. The convergence of the electronic density of a target region can be achieved by including sufficiently large parts of the surroundings. However, unlike the electronic density, the energy of the target region converges slowly as more of the surroundings are included. Therefore, the traditional requirement of converging the energy prior to computing properties must be abandoned to make such an approach for obtaining size-intensive properties viable.

This is the foundation for paper B, where we illustrate the advantage of such an approach. We use cluster models of crystalline NH_3, with the target region being a single NH_3 molecule. Upon increasing the cluster

size, the HF electronic density of the target region converges, while the HF energy does not. The converged electronic density allows converged HF dipole moments to be obtained, thereby showing the approach's applicability. Further, any existing fragment or orbital-based correlated wave function models can be used to compute size-intensive properties of a target region. We illustrate this possibility by computing CC2-in-HF vertical excitation energies, which also converge with cluster size. In the following sections, relevant background and motivation are presented, followed by the study's main findings. Finally, a summary and discussion are given.

4.1 Background and motivation

The locality of electron correlation is exploited in several low-scaling electronic structure models.[14–30,56–60] However, they still adhere to the traditional approach of converging the energy, and therefore, do not reap all the benefits of the "nearsightedness" of electrons.[55] In case of molecular crystals, converging the energy of a target region requires the infinite surroundings to be taken into account. This motivates the extension of wave function models originally developed for single molecules to periodic codes.[137–140]

To achieve linear-scaling wave functions for molecular systems, local MO spaces are often used, with approaches existing to generate local occupied[82–84] and virtual[85–89] spaces. However, for the infinite molecular crystals, orbital space locality faces two challenges i) orthogonalization tails and ii) near-linear dependencies.

Orthogonalization tails compromise compactness since local MOs outside a given region must have components inside the region. This is especially problematic when using high-quality basis sets since many MOs are centered far outside the given region. Orthogonalization tails can lead to slow convergence of the correlation energy in local correlation approaches.[94,95]

Near-linear dependencies are especially problematic for large or dense systems. This is due to the concerted effect of AOs on different atomic centers, which greatly enhances near-linear dependencies. Dealing with near-linear dependencies in a periodic framework may

compromise the basis set quality since removing AOs from a unit cell will remove AOs from all unit cells. While near-linear dependencies can be avoided by using plane wave basis sets, this approach has other problems, such as the need for a high energy cut-off to achieve accurate results.

When targeting size-intensive properties of molecular crystals, the "nearsightedness" of electrons may allow the use of cluster models instead of the periodic approach. Cluster models enable different basis sets to be used in different regions of the cluster. Hence one may use smaller basis sets outside a target region. Therefore, the problems of orthogonalization tails and near-linear dependencies may be alleviated by using cluster models. Further, cluster models allow for any molecular wave function based scheme to be used for its description. These aspects motivate the investigation of the convergence of the electronic density of a target region in cluster models of molecular crystals. Convergence of the electronic density allows size-intensive properties to be obtained.

To rationalize why the electronic density of a target region should converge with cluster size, it is instructive to consider an optimization scheme that does not enforce a canonical basis. For illustrative purposes, assume that the optimization scheme generates, or transforms to, an orbital basis of local occupied and virtual MOs in each HF iteration. The resulting new occupied orbital i can be written as

$$\tilde{C}_{\mu i} = C_{\mu i} + \sum_a C_{\mu a}\kappa_{ai} + \mathcal{O}(\boldsymbol{\kappa}^2) \tag{4.1}$$

where $\boldsymbol{\kappa}$ is an orbital rotations matrix. Hence, for each iteration in the energy optimization, an orbital i in the target region will get an amount of virtual MO a mixed in, weighted (to first order) by the magnitude of κ_{ai}. If an occupied MO i, and likewise the electronic density, is to be changed by a virtual orbital a centered far away, κ_{ai} must be of significant size. The size of κ_{ai} is, to a large extent, determined by the size of F_{ai}, and for local orbitals i and a centered far away from each other F_{ai} exhibits a rapid decay.[141] The MOs of the target region, and likewise, the electronic density of the target region will, therefore, at some point be unchanged when increasing the cluster size. The use of local MOs in this analysis is convenient from a conceptual point of view since local MOs also enable partitioning into a target electronic density which is local. However, the electronic density is invariant with respect to redundant rotations, and

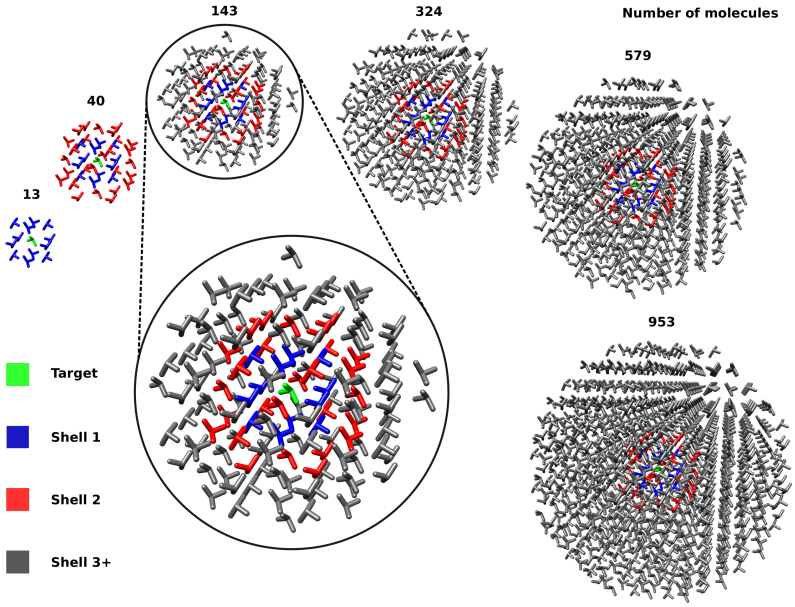

Figure 4.1: NH_3 clusters where the target region and the first surrounding shells have been highlighted. The target region (green) is the central NH_3 molecule, shell 1 (blue) is the first shell of 12 molecules, shell 2 (red) is the second shell of 27 molecules, while shell 3+ (grey) is any shells beyond the second shell.

the convergence properties of a local part of the electronic density with cluster size are indifferent to the choice of basis.

The NH_3 clusters are depicted in Figure 4.1, for which HF calculations are done, followed by localization and partitioning of the occupied orbitals as stated in section 3.3 of paper B. The partitioning of the HF electronic energy presented in Section 2.3 is used to evaluate the energy and electronic density of the target region. The partitioning is not a requirement for computing size-intensive properties of a target region but is useful to illustrate the convergence behaviour of the energy and, indirectly, the electronic density. In addition to investigating the convergence of the electronic density, HF dipole moments and CC2-in-HF vertical excitation energies are also computed. In the case of CC2-in-HF calculations, the virtual space is described by PAOs. These are generated for a region termed the virtual active region, where two regions of different

size are defined: target + shell 1 or target + shell 1-2.

4.2 Main findings

The energy differences between clusters of different sizes are calculated to evaluate how the electronic density and energy of the target region are affected by the cluster size. These are given by

$$\Delta E(m,n) = E^m - E^n \tag{4.2}$$

where E^m is an energy ($E_t^{\text{1-el}}$, $E_t^{\text{2-el}}$, E_t^{int} or E_t) for a cluster containing m molecules and E^n is an energy for a cluster containing n molecules.

Upon increasing the cluster size, the negative one-electron contribution to the target region, $E_t^{\text{1-el}}$, increases in magnitude. This is due to the long-range interaction between the electrons in the target region and all nuclei of the cluster. For the two-electron interaction between the electrons in the target region and the electrons outside the target region, E_t^{int}, we see the same long-range effects. However, here the energy contributions are positive. Hence, these energy contributions do not converge with cluster size.

However, $\Delta E_t^{\text{1-el}}$ and ΔE_t^{int} have similar values, except with opposite signs, and they nearly cancel each other out. This results in small changes to the energy of the target region E_t. Although the changes in E_t are small, convergence cannot be reached. This can be seen in Figure 4.2, where the absolute value of $\Delta E_t(m,n)$ is plotted against m. On the other hand, the two-electron contribution to the target region, $E_t^{\text{2-el}}$, does converge with cluster size. This can be seen from the absolute value of $\Delta E_t^{\text{2-el}}(m,n)$, also plotted in Figure 4.2. This implies that, beyond a given cluster size, $\mathbf{D}_t$ is not significantly affected by extending the cluster further. This numerically illustrates the theoretical discussion in the previous section and originates from the nearsightedness of electrons in many-atom systems. Therefore, one can conclude that a converged electronic density can be obtained for a target region without converging the ground state energy.

The converged electronic density of the target region should result in converged size-intensive properties with respect to cluster size. This is the case for HF dipole moments of the target region, which makes sense since dipole moments only depend on the electronic density. However, CC2-in-HF vertical excitation energies were also found to converge with

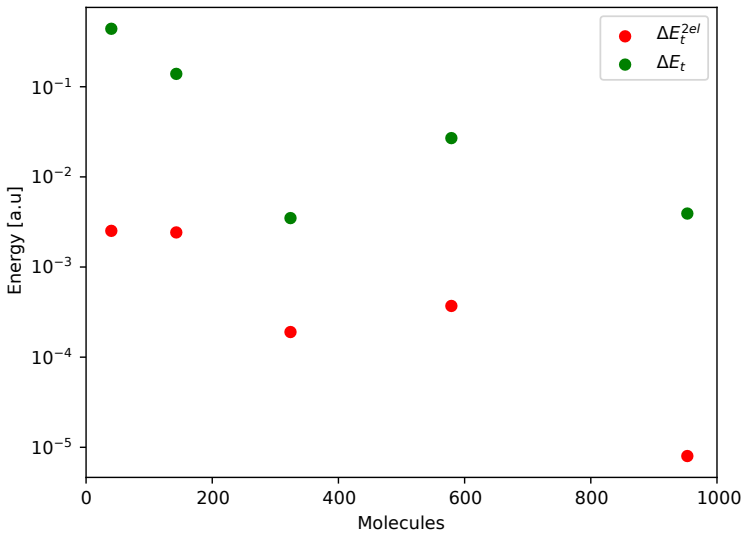

Figure 4.2: The absolute value of $\Delta E_t(m,n)$ and $\Delta E_t^{2\text{-el}}(m,n)$ plotted against number of molecules (m in equation 4.2).

cluster size. The occupied and virtual correlated natural transition orbitals for the excitation are shown in Figure 4.3. While this requires a virtual orbital space extending beyond the target region, it was sufficient to use the target + shell 1 as the virtual active region. Expanding the virtual orbital space beyond shell 1 only leads to minor changes in excitation energies. Therefore, a limited virtual space might be sufficient to obtain accurate excitation energies for local excitations. It is important to note that the CC2-in-HF vertical excitation does not correspond to the lowest excitation in crystalline NH_3 due to the occupied space only including the target region.

Further, only minor changes in dipole moment occurred when changing the basis set outside the target region. Similarly, the vertical excitation energies mainly depend on the basis set of shell 1. This illustrates that smaller basis sets can be used outside the target region, and a minimal basis might be sufficient for the outer regions. Therefore, a cluster approach could alleviate the problem of near-linear dependencies.

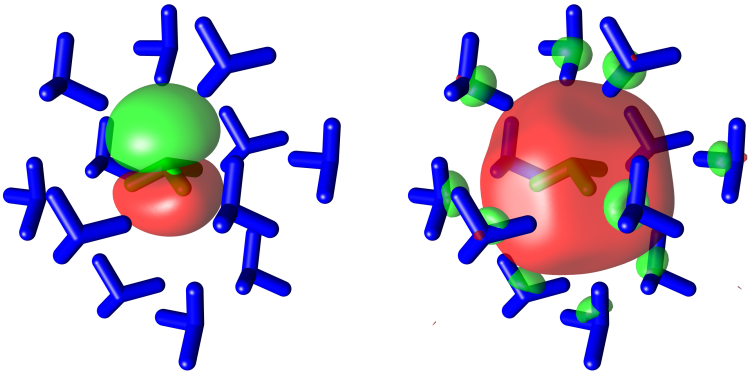

Figure 4.3: Occupied (left) and virtual (right) correlated natural transition orbitals for the investigated excitation. A cluster size of 143 molecules was used, however only the target region and shell 1 are visualized.

4.3 Summary and discussion

The possibility of computing size-intensive properties for a target region using a converged electronic density is currently an underexploited feature. The feature is particularly useful for molecular crystals since periodic approaches and problems associated with their application can be avoided. For cluster models of crystalline NH_3, the HF electronic density of the target region converges with cluster size, while the HF energy does not. The convergence of the electronic density translates to convergence of HF dipole moments of the target region with cluster size. This illustrates the validity of such an approach for computing size-intensive properties for a target region using a converged electronic density.

Further, any existing fragment or orbital-based correlated wave function model can be used to compute size-intensive properties of a target region. The convergence of CC2-in-HF vertical excitation energies of a local excitation illustrates this possibility. Naturally, a virtual space which is allowed to extend outside the target region is needed. However, one does not need to go far outside the target region to obtain convergence of the vertical excitation energy. The important point is to use a high-quality basis set inside and near the target region. The ability to obtain

local excited state properties using coupled cluster for systems as large as molecular crystals has several potential applications. Computing core excitations is one potential application since the occupied orbitals which are excited from are local.

The cluster approach also lends itself to using multilevel coupled cluster methods,[31,97,98,102,103] which allow higher level coupled cluster methods to treat parts of the cluster. In cases where large clusters are needed to converge the electronic density of the target region, HF calculation might become prohibitively expensive. However, computational cost can be reduced by using multilevel HF[104] instead of HF, or by using a QM/MM[142,143] or QM/PCM[144,145] approach.

The ability to use different basis sets in regions of the cluster is another useful feature of cluster approaches. The HF dipole moments and the CC2-in-HF excitation energies show that convergence can be reached with a minimal basis in the outer regions of the cluster. Using smaller basis sets to treat parts of a cluster comes with the benefits of lowering computational costs. Furthermore, it could also alleviate the problems of near-linear dependencies and orthogonalization tails, which can occur in periodic approaches.

5 A computational elucidation of the potential of XAS for gas clathrates

Gas clathrates are structures where small guest molecules are trapped in hydrogen-bonded water cages, where CH_4 and CO_2 clathrate are of particular interest. The interest in CH_4 clathrate can partly be explained by its potential as an energy source, while the interest in CO_2 clathrate is related to carbon capture and storage. Crystal structure, cage occupancy, and formation mechanism are important for gas clathrates. X-ray absorption spectroscopy (XAS) is a promising candidate for studying these properties. However, the method has received limited attention. Therefore, a computational study of gas clathrates may elucidate the potential of characterizing gas clathrates using XAS.

The cluster approach presented in the previous chapter should be well suited for generating core excitation spectra since the excitation is from core orbitals, which are local. In paper C carbon core excitation spectra of CH_4 and CO_2 clathrate are generated using this approach, and the potential of XAS for characterizing these gas clathrates is evaluated.

The spectra generated clearly show that one should be able to distinguish between CH_4 and CO_2 clathrate using XAS. For CH_4 clathrate, different cage types might be distinguishable from XAS, while this is not the case for CO_2 clathrate. Being able to distinguish between cages is

useful since it is required to determine the occupancy of specific cages. In the following sections, relevant background and motivation are presented, followed by the study's main findings. Finally, a summary and discussion are given.

5.1 Background and motivation

Gas clathrates are found in three common crystal structures, where the size of the gas molecules is the main predictor of which structure will be formed. The most common structure found in earth's natural environment is the cubic structure I (sI), where the cages are occupied by small guest molecules, such as CH_4, ethane, CO_2, etc.[146] For larger guest molecules (propane, iso-butane, etc.) the common structure is the cubic structure II (sII). In the case of sI, there are two unique dodecahedral cages (D-cages) and six unique tetradecahedral cages (T-cages), while for sII there are 16 unique D-cages and 8 hexadecahedral cages (H-cages). Being able to determine the total cage occupancy and the cage occupancy of a certain cage type is needed to evaluate the resources found in gas clathrate deposits, as well as the amount of gas that can be stored in clathrates. It should be noted that neither CH_4[147–149] nor CO_2[150] are fixed within the cages, with significant thermal fluctuations for both clathrates even at low temperatures.

There are currently four experimental methods which are widely used to characterize clathrates: Raman spectroscopy,[151–159] nuclear magnetic resonance spectroscopy (NMR),[152,158,160–168] X-ray diffraction (XRD),[153,156,161] and neutron diffraction.[147,150,155,169,170] Conversely, XAS has seen limited use for the study of gas clathrates. To the best of our knowledge, the XAS studies done by Finney and coworkers[171,172] on the formation and pressure dependence of krypton clathrate are the only XAS studies of gas clathrates. XAS has several attractive properties, which makes it a promising technique for studying gas clathrates. Firstly, it is an element-specific technique, which could be useful to differentiate guest molecules containing different atoms. Secondly, XAS is sensitive to the local environment around the absorbing species. For this reason, one should be able to determine the guest molecule, cage type, and possibly even cage vacancies. The sensitivity to the local environment also makes XAS, or more specifically extended X-ray absorption fine

structure (EXAFS), suited to investigate the formation of gas clathrates[171].

However, one of the drawbacks of XAS is that it is an average technique. This means that different species might yield overlapping spectra, which could make it challenging to identify individual species. The identification of individual species is often helped by the use of standards of known species. However, this may not be an option in the case of gas clathrates. A computational study may elucidate the potential of XAS for the study of gas clathrates.

The clusters used in this study were centered on the guest molecules (CH_4 and CO_2) in the unique cages of sI, which constitute 2 D-cages and 6 T-cages, and a single cluster of CH_4 in an H-cage of sII. Carbon core excitation energies and oscillator strengths are calculated using CC2-in-HF, where the guest molecule is chosen as the target region. To reduce the computational cost, the core valence separation approximation is employed, where only excitations involving core orbitals are computed.[6,74,173] The cluster size and virtual active region are chosen based on a preliminary study. A cluster with a radius of 10.2 Å is found to yield a converged spectrum. A virtual active space consisting of the guest molecule and the waters making up the cage is sufficient to obtain a converged spectrum. The basis sets were chosen based on the basis set study in paper B.

5.2 Main findings

For brevity, the spectra are said to be generated for given cages, even though the excitation is from the guest molecules. In the case of CH_4 clathrate there are clear differences between the spectra of all unique cages, even for the same cage type. However, the most significant difference is observed between in D-cages and either T- or H-cages. This can be seen in Figure 5.1, where a representative spectrum of each cage type is plotted. While D-cages produce a spectrum with a single main peak, T- and H-cages produce spectra with a distinct shoulder feature. The spectra of the larger T- and H-cages are also shifted to lower energies. Based on this, one should be able to distinguish between D- or T-type cages. This is an important finding as these cages make up the most common CH_4 crystal structure (sI). Unlike sI CH_4 clathrate, the unique cages of sI CO_2 clathrate yield indistinguishable main peaks, as can be seen from the spectra in Figure 5.2. There are some observable differences in the peaks

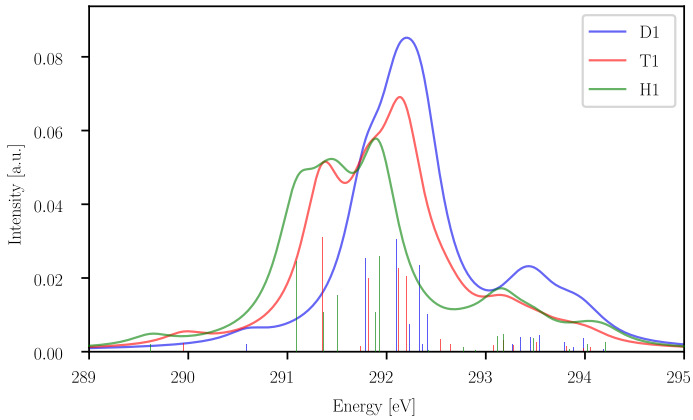

Figure 5.1: Carbon core excitation spectra of CH_4 in D1 (blue), T1 (red) and H1 (green) cages.

following the main peak. However, these have a low intensity and are most likely of limited experimental value. The experimental XAS spectrum for sI CO_2 clathrate, which can be obtained by averaging the spectra of all unique cages, should be quite similar to the spectra of any of the spectra shown in Figure 5.2. For sI CH_4 clathrate, the experimental XAS spectra should appear like the spectra shown in Figure 5.3, which is an averaged spectrum of all unique cages. The position of the main feature differs by approximately 1 eV between CH_4 and CO_2 clathrate. Therefore, XAS should be able to differentiate between the two clathrates.

Since defect-free structures are unlikely to exist, we also investigate how CH_4 vacancies affect the spectra. We focus on the two cage types of sI, specifically the T1 and D1 cage. There are 14 cages and 12 cages surrounding the T1 and D1 cages, respectively. When all surrounding cages are vacant, we observe only minor changes to the spectra of both cages, as seen in Figures 5.4-5.5. Since the changes are minor, it is unlikely that they will be observed experimentally.

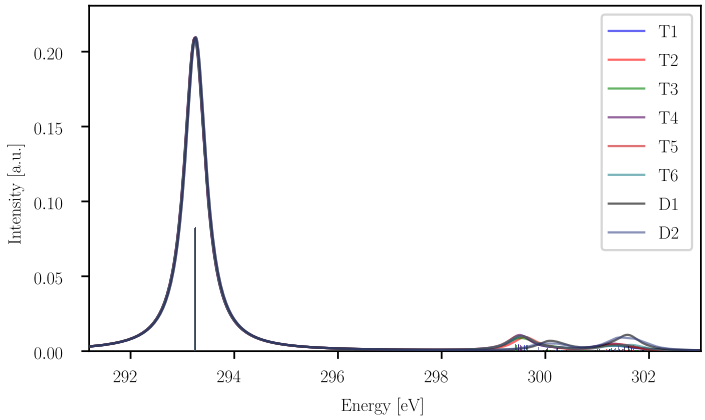

Figure 5.2: Carbon core excitation spectra of CO_2 in all unique cages of sI CO_2 clathrate.

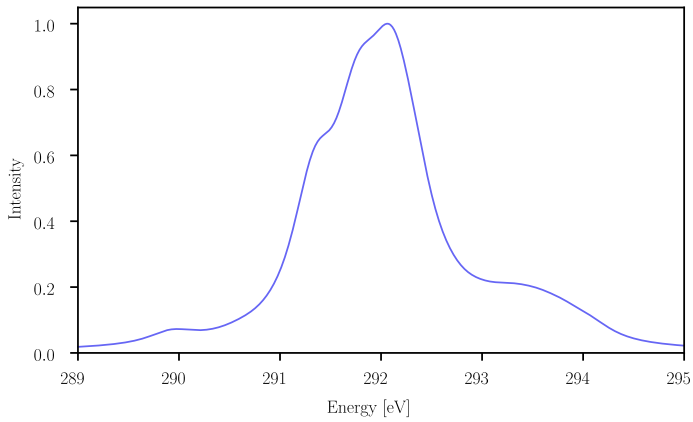

Figure 5.3: Carbon core excitation spectra of CH_4 averaged over all unique cages of sI CH_4 clathrate.

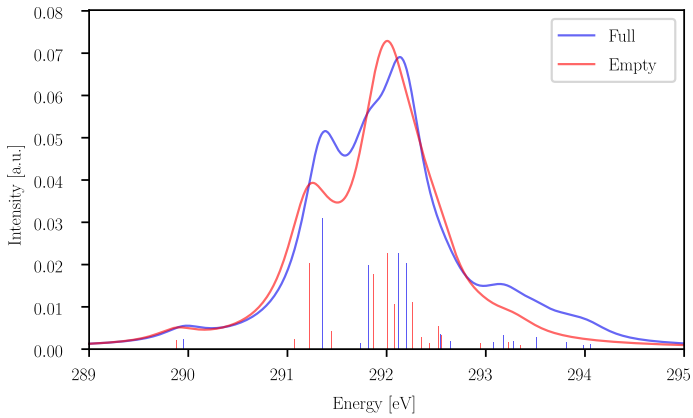

Figure 5.4: Carbon core excitation spectra of CH_4 in the T1 cage, with the all surrounding cages filled (blue) and all surrounding cages empty (red).

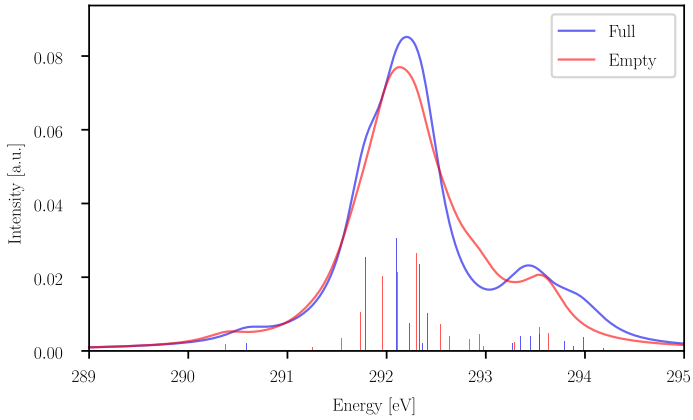

Figure 5.5: Carbon core excitation spectra of CH_4 in the D1, with the all surrounding cages filled (blue) and all surrounding cages empty (red).

5.3 Summary and discussion

The carbon core excitation spectra indicate that XAS can be used to characterize CH_4 and CO_2 clathrate. XAS should be able to distinguish between guest molecules, and for sI CH_4 clathrate, also be able to distinguish between cage types. A potential deficiency of this study could arise due to the guest molecule being fixed, while they are known to have significant thermal fluctuations. A useful continuation of this study would be to investigate the effect that rotation and displacement within the cage have on the carbon core excitation spectra. In addition, a future study using multilevel coupled cluster would be valuable, since CC2 might not be able to accurately represent core excitations. Ideally, MLCC3 would be used, since the inclusion of triple excitations has shown to be important when describing core excited states.[7,8]

Another important aspect not considered in this study is whether it is experimentally feasible to use XAS to study clathrates. Since clathrates require very particular temperature and pressure conditions, it could be challenging. However, since XAS is very sensitive to the local environment, investigation of the formation of clathrates could be an area where XAS would excel.

Upon introducing vacancies in surrounding cages, only limited changes to the spectra are observed. This might mean that XAS is unsuitable to investigate the distribution of vacancies. Including more states in the calculation may capture differences, but the premise for the study is the locality of the excitations and high-lying excitations may require a less local virtual space.

The results indicate that the cluster approach presented in the previous chapter is well suited for investigating core excitations of large systems. For core excitations, a small occupied space is a good approximation, while this was not the case for the valence excitations of crystalline NH_3.

6 Conclusion and outlook

In this thesis, spectroscopic properties of photoswitches and molecular crystals are studied using coupled cluster methods. The studied systems have properties making them challenging to treat with coupled cluster methods, either due to their size or the presence of problematic degeneracies. Despite this, coupled cluster methods are shown to be valuable tools even for these challenging systems, as long as limitations are kept in mind.

The main counterargument to using coupled cluster methods to treat photoswitches is their inability to treat ground state degeneracies and the problems arising at conical intersections between excited states of the same symmetry.[67–69] However, coupled cluster is well suited to treat all but these problematic features. This is shown for the photoswitches PSB3 and azobenzene, where the performance of CC3 is particularly impressive. CC3 was able to reliably generate ground and excited state potential energy surfaces, even for states of doubly excited character and for regions in the vicinity of an isomerization point. Being able to accurately describe potential energy surfaces without having to select an active orbital space, makes coupled cluster methods including triple excitations a valuable tool for investigating photoswitches. One can envisage CC3 as a tool used for preliminary investigation of prospective isomerization paths or to evaluate multireference methods in regards to dynamical correlation or active space selection. However, the system size which can be treated using coupled cluster methods including triple excitations is rather limited. This highlights the value of reducing the computational scaling of

such methods.

In the case of molecular crystals, the steep computational scaling of coupled cluster methods is the obstacle which needs to be overcome. However, local excited state properties can be investigated with coupled cluster in a reduced orbital space, in combination with cluster models of molecular crystals. This possibility arises due to the "nearsightedness" of electrons in many-atom systems[55], which allows for size-intensive properties of a target region to be obtained without requiring the convergence of the energy. This was illustrated for cluster models of crystalline NH_3, where the electronic density of a target region was found to converge with cluster size. As a result, converged HF dipole moments and CC2-in-HF excitation energies of a local excitation could be computed. Currently, this is an underappreciated feature. In principle one can use any existing fragment or orbital correlated wave function model to compute size-intensive properties of the target region. However, due to the energy of the target region not converging with cluster size, the traditional approach of first converging the energy must be abandoned. A cluster approach also gives flexibility in the choice of basis set, where a minimal basis may be sufficient in outer regions of the cluster. This can reduce computational cost and alleviate problems of orthogonalization tails and near-linear dependencies, which can be problematic for periodic approaches.

The locality of core orbitals makes computing core excited properties an interesting application of cluster approaches. Core excitation spectra are generated for CH_4 and CO_2 clathrates, with the goal of evaluating the potential of XAS for the study of such systems. Similar to the other size-intensive properties investigated, the spectra converge with cluster size and only require a moderately sized virtual orbital space. Based on the spectra generated, XAS should be able to differentiate between the CH_4 and CO_2 clathrates and might be able to distinguish between cage types in sI CH_4 clathrate. The theoretical spectra calculated in our study might be of value for interpreting experimental data, once it is available. However, to accurately describe core excitations the inclusion of triple excitations has been shown to be important. Therefore, a continuation study could be to use multilevel CC3, where only the central molecule is treated with CC3. Spectra generated with multilevel CC3 would be valuable in and of itself, but also to evaluate the quality of CC2 spectra. Fur-

ther, since significant thermal fluctuations of the guest molecule within the cage is characteristic for CH_4 and CO_2 clathrates, investigating the effect rotation and displacement of the guest molecule has on the carbon core excitation spectra would be useful.

Bibliography

[1] Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic Structure Theory, First Edition*; Wiley, 2000.

[2] Bartlett, R. J.; Musiał, M. "Coupled-cluster theory in quantum chemistry", *Rev. Mod. Phys.* **2007**, 79, 291–352.

[3] Sneskov, K.; Christiansen, O. "Excited state coupled cluster methods", *WIREs Comput. Mol. Sci.* **2012**, 2, 566–584.

[4] Coriani, S.; Christiansen, O.; Fransson, T.; Norman, P. "Coupled-cluster response theory for near-edge x-ray-absorption fine structure of atoms and molecules", *Phys. Rev.* A **2012**, 85, 022507.

[5] Coriani, S.; Fransson, T.; Christiansen, O.; Norman, P. "Asymmetric-Lanczos-Chain-Driven implementation of electronic resonance convergent coupled-cluster linear response theory", *J. Chem. Theory Comput.* **2012**, 8, 1616–1628.

[6] Coriani, S.; Koch, H. "Communication: X-ray absorption spectra and core-ionization potentials within a core-valence separated coupled cluster framework", *J. Chem. Phys.* **2015**, 143, 181103.

[7] Liu, J.; Matthews, D.; Coriani, S.; Cheng, L. "Benchmark calculations of K-edge ionization energies for first-row elements using scalar-relativistic core–valence-separated equation-of-motion coupled-cluster methods", *J. Chem. Theory Comput.* **2019**, 15, 1642–1651.

[8] Matthews, D. A. "EOM-CC methods with approximate triple excitations applied to core excitation and ionisation energies", *Mol. Phys.* **2020**, *118*, e1771448.

[9] Koch, H.; Christiansen, O.; Jørgensen, P.; Olsen, J. "Excitation energies of BH, CH2 and Ne in full configuration interaction and the hierarchy CCS, CC2, CCSD and CC3 of coupled cluster models", *Chem. Phys. Lett.* **1995**, *244*, 75–82.

[10] Christiansen, O.; Koch, H.; Jørgensen, P.; Olsen, J. "Excitation energies of H2O, N2 and C2 in full configuration interaction and coupled cluster theory", *Chem. Phys. Lett.* **1996**, *256*, 185–194.

[11] Schreiber, M.; Silva-Junior, M. R.; Sauer, S. P. A.; Thiel, W. "Benchmarks for electronically excited states: CASPT2, CC2, CCSD, and CC3", *J. Chem. Phys* **2008**, *128*, 134110.

[12] Silva-Junior, M. R.; Sauer, S. P.; Schreiber, M.; Thiel, W. "Basis set effects on coupled cluster benchmarks of electronically excited states: CC3, CCSDR(3) and CC2", *Mol. Phys.* **2010**, *108*, 453–465.

[13] Loos, P.-F.; Scemama, A.; Blondel, A.; Garniron, Y.; Caffarel, M.; Jacquemin, D. "A Mountaineering Strategy to Excited States: Highly Accurate Reference Energies and Benchmarks", *J. Chem. Theory Comput.* **2018**, *14*, 4360–4379.

[14] Hampel, C.; Werner, H. "Local treatment of electron correlation in coupled cluster theory", *J. Chem. Phys.* **1996**, *104*, 6286–6297.

[15] Scuseria, G. E.; Ayala, P. Y. "Linear scaling coupled cluster and perturbation theories in the atomic orbital basis", *J. Chem. Phys.* **1999**, *111*, 8330–8343.

[16] Christiansen, O.; Manninen, P.; Jørgensen, P.; Olsen, J. "Coupled-cluster theory in a projected atomic orbital basis", *J. Chem. Phys.* **2006**, *124*, 084103.

[17] Flocke, N.; Bartlett, R. J. "A natural linear scaling coupled-cluster method", *J. Chem. Phys.* **2004**, *121*, 10935–10944.

[18] Li, S.; Ma, J.; Jiang, Y. "Linear scaling local correlation approach for solving the coupled cluster equations of large systems", *J. Comput. Chem.* **2002**, *23*, 237–244.

[19] Kobayashi, M.; Nakai, H. “Extension of linear-scaling divide-and-conquer-based correlation method to coupled cluster theory with singles and doubles excitations”, *J. Chem. Phys.* **2008**, *129*, 044103.

[20] Fedorov, D. G.; Kitaura, K. “Coupled-cluster theory based upon the fragment molecular-orbital method”, *J. Chem. Phys.* **2005**, *123*, 134103.

[21] Stoll, H. “Correlation energy of diamond”, *Phys. Rev. B* **1992**, *46*, 6700–6704.

[22] Friedrich, J.; Hanrath, M.; Dolg, M. “Fully automated implementation of the incremental scheme: Application to CCSD energies for hydrocarbons and transition metal compounds”, *J. Chem. Phys.* **2007**, *126*, 154110.

[23] Subotnik, J. E.; Head-Gordon, M. “A local correlation model that yields intrinsically smooth potential-energy surfaces”, *J. Chem. Phys.* **2005**, *123*, 064108.

[24] Maslen, P.; Lee, M.; Head-Gordon, M. “An accurate local model for triple substitutions in fourth order Møller–Plesset theory and in perturbative corrections to singles and doubles coupled cluster methods”, *Chem. Phys. Lett.* **2000**, *319*, 205–212.

[25] Auer, A. A.; Nooijen, M. “Dynamically screened local correlation method using enveloping localized orbitals”, *Chem. Phys.* **2006**, *125*, 024104.

[26] Neese, F.; Wennmohs, F.; Hansen, A. “Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method”, *J. Chem. Phys.* **2009**, *130*, 114108.

[27] Crawford, T.; King, R. A. “Locally correlated equation-of-motion coupled cluster theory for the excited states of large molecules”, *Chem. Phys. Lett.* **2002**, *366*, 611–622.

[28] Hirata, S.; Valiev, M.; Dupuis, M.; Xantheas, S. S.; Sugiki, S.; Sekino, H. “Fast electron correlation methods for molecular clusters in the ground and excited states”, *Mol. Phys.* **2005**, *103*, 2255–2265.

[29] Kats, D.; Korona, T.; Schütz, M. "Local CC2 electronic excitation energies for large molecules with density fitting", *J. Chem. Phys.* **2006**, *125*, 104106.

[30] Ziółkowski, M.; Jansík, B.; Kjærgaard, T.; Jørgensen, P. "Linear scaling coupled cluster method with correlation energy based error control", *J. Chem. Phys.* **2010**, *133*, 014107.

[31] Myhre, R. H.; de Merás, A. M. S.; Koch, H. "The extended CC2 model ECC2", *Mol. Phys.* **2013**, *111*, 1109–1118.

[32] Robb, M.; Bernardi, F.; Olivucci, M. "Conical intersections as a mechanistic feature of organic photochemistry", *Pure Appl. Chem.* **1995**, 67, 783 – 789.

[33] Levine, B. G.; Martínez, T. J. "Isomerization Through Conical Intersections", *Annu. Rev. Phys. Chem.* **2007**, *58*, 613–634.

[34] Garavelli, M.; Vreven, T.; Celani, P.; Bernardi, F.; Robb, M. A.; Olivucci, M. "Photoisomerization Path for a Realistic Retinal Chromophore Model: The Nonatetraeniminium Cation", *J. Am. Chem. Soc.* **1998**, *120*, 1285–1288.

[35] Ishikawa, T.; Noro, T.; Shoda, T. "Theoretical study on the photoisomerization of azobenzene", *J. Chem. Phys.* **2001**, *115*, 7503–7512.

[36] De Vico, L.; Page, C. S.; Garavelli, M.; Bernardi, F.; Basosi, R.; Olivucci, M. "Reaction path analysis of the "tunable" photoisomerization selectivity of free and locked retinal chromophores", *J. Am. Chem. Soc.* **2002**, *124*, 4124–4134.

[37] Page, C. S.; Olivucci, M. "Ground and excited state CASPT2 geometry optimizations of small organic molecules", *J. Comput. Chem.* **2003**, *24*, 298–309.

[38] Wei-Guang Diau, E. "A new trans-to-cis photoisomerization mechanism of azobenzene on the S1(n,π^*) surface", *J. Phys. Chem. A* **2004**, *108*, 950–956.

[39] Cembran, A.; Bernardi, F.; Garavelli, M.; Gagliardi, L.; Orlandi, G. "On the mechanism of the cis-trans isomerization in the lowest electronic states of azobenzene: S0, S1, and T1", *J. Am. Chem. Soc.* **2004**, *126*, 3234–3243.

[40] Gagliardi, L.; Orlandi, G.; Bernardi, F.; Cembran, A.; Garavelli, M. "A theoretical study of the lowest electronic states of azobenzene: The role of torsion coordinate in the cis-trans photoisomerization", *Theor. Chem. Acc.* **2004**, *111*, 363–372.

[41] Wanko, M.; Hoffmann, M.; Strodel, P.; Koslowski, A.; Thiel, W.; Neese, F.; Frauenheim, T.; Elstner, M. "Calculating absorption shifts for retinal proteins: computational challenges", *J. Phys. Chem. B* **2005**, *109*, 3606–3615.

[42] Conti, I.; Garavelli, M.; Orlandi, G. "The different photoisomerization efficiency of azobenzene in the lowest $n\pi^*$ and $\pi\pi^*$ singlets: the role of a phantom state", *J. Am. Chem. Soc.* **2008**, *130*, 5216–5230.

[43] Tiberio, G.; Muccioli, L.; Berardi, R.; Zannoni, C. "How does the trans–cis photoisomerization of azobenzene take place in organic solvents?", *ChemPhysChem* **2010**, *11*, 1018–1028.

[44] Gozem, S.; Huntress, M.; Schapiro, I.; Lindh, R.; Granovsky, A. A.; Angeli, C.; Olivucci, M. "Dynamic electron correlation effects on the ground state potential energy surface of a retinal chromophore model", *J. Chem. Theory Comput.* **2012**, *8*, 4069–4080.

[45] Harabuchi, Y.; Ishii, M.; Nakayama, A.; Noro, T.; Taketsugu, T. "A multireference perturbation study of the NN stretching frequency of trans-azobenzene in $n\pi^*$ excitation and an implication for the photoisomerization mechanism", *J. Chem. Phys* **2013**, *138*, 064305.

[46] Walczak, E.; Szefczyk, B.; Andruniów, T. "Geometries and vertical excitation energies in retinal analogues resolved at the CASPT2 level of theory: critical assessment of the performance of CASSCF, CC2, and DFT methods", *J. Chem. Theory Comput.* **2013**, *9*, 4915–4927.

[47] Quick, M.; Dobryakov, A. L.; Gerecke, M.; Richter, C.; Berndt, F.; Ioffe, I. N.; Granovsky, A. A.; Mahrwald, R.; Ernsting, N. P.; Kovalenko, S. A. "Photoisomerization dynamics and pathways of trans- and cis-azobenzene in solution from broadband femtosecond spectroscopies and calculations", *J. Phys. Chem. B* **2014**, *118*, 8756–8771.

[48] Casellas, J.; Bearpark, M. J.; Reguero, M. "Excited-state decay in the photoisomerisation of azobenzene: a new balance between mechanisms", *ChemPhysChem* **2016**, *17*, 3068–3079.

[49] Yu, L.; Xu, C.; Zhu, C. "Probing the $\pi \rightarrow \pi^*$ photoisomerization mechanism of cis-azobenzene by multi-state ab initio on-the-fly trajectory dynamics simulation", *Phys. Chem. Chem. Phys.* **2015**, *17*, 17646–17660.

[50] Xu, C.; Yu, L.; Gu, F. L.; Zhu, C. "Probing the $\pi \rightarrow \pi^*$ photoisomerization mechanism of trans-azobenzene by multi-state ab initio on-the-fly trajectory dynamics simulations", *Phys. Chem. Chem. Phys.* **2018**, *20*, 23885–23897.

[51] Aleotti, F.; Soprani, L.; Nenov, A.; Berardi, R.; Arcioni, A.; Zannoni, C.; Garavelli, M. "Multidimensional potential energy surfaces resolved at the RASPT2 level for accurate photoinduced isomerization dynamics of azobenzene", *J. Chem. Theory Comput.* **2019**, *15*, 6813–6823.

[52] Christiansen, O.; Koch, H.; Jørgensen, P. "The second-order approximate coupled cluster singles and doubles model CC2", *Chem. Phys. Lett.* **1995**, *243*, 409 – 418.

[53] Purvis III, G.; Bartlett, R. "A full coupled-cluster singles and doubles model: The inclusion of disconnected triples", *J. Chem. Phys.* **1982**, *76*, 1910–1918.

[54] Koch, H.; Christiansen, O.; Jørgensen, P.; Merás, A.; Helgaker, T. "The CC3 model : An iterative coupled cluster approach including connected triples", *J. Chem. Phys.* **1997**, *106*, 1808–1818.

[55] Kohn, W. "Density Functional and Density Matrix Method Scaling Linearly with the Number of Atoms", *Phys. Rev. Lett.* **1996**, *76*, 3168–3171.

[56] Saebo, S.; Pulay, P. "Local Treatment of Electron Correlation", *Annu. Rev. Phys. Chem.* **1993**, *44*, 213–236.

[57] Schütz, M.; Hetzer, G.; Werner, H.-J. "Low-order scaling local electron correlation methods. I. Linear scaling local MP2", *J. Chem. Phys.* **1999**, *111*, 5691–5705.

[58] Ayala, P. Y.; Scuseria, G. E. "Linear scaling second-order Møller–Plesset theory in the atomic orbital basis for large molecular systems", *J. Chem. Phys.* **1999**, *110*, 3660–3671.

[59] Doser, B.; Lambrecht, D. S.; Ochsenfeld, C. "Tighter multipole-based integral estimates and parallel implementation of linear-scaling AO–MP2 theory", *Phys. Chem. Chem. Phys.* **2008**, 10, 3335–3344.

[60] Yang, J.; Kurashige, Y.; Manby, F. R.; Chan, G. K. "Tensor factorizations of local second-order Møller–Plesset theory", *J. Chem. Phys.* **2011**, 134, 044123.

[61] Hartree, D. R. "The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part II. Some Results and Discussion", *Math. Proc. Cambridge Philos. Soc.* **1928**, 24, 111–132.

[62] Fock, V. "Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems", *Z. Physik* **1930**, 61, 126–148.

[63] Slater, J. C. "The Theory of Complex Spectra", *Phys. Rev.* **1929**, 34, 1293–1322.

[64] Roothaan, C. C. J. "New Developments in Molecular Orbital Theory", *Rev. Mod. Phys.* **1951**, 23, 69–89.

[65] Hall, G. "The molecular orbital theory of chemical valency VIII. A method of calculating ionization potentials", *Proc. R. Soc. Lond. A. Math. Phys. Sci.* **1951**, 205, 541–552.

[66] Noga, J.; Bartlett, R. J. "The full CCSDT model for molecular electronic structure", *J. Chem. Phys.* **1987**, 86, 7041–7050.

[67] Hättig, C. *Response Theory and Molecular Properties: (A Tribute to Jan Linderberg and Poul Jørgensen)*; Jensen, H., Ed.; Advances in Quantum Chemistry, Vol. 50; Academic Press, 2005; pp 37–60.

[68] Köhn, A.; Tajti, A. "Can coupled-cluster theory treat conical intersections?", *J. Chem. Phys.* **2007**, 127, 044105.

[69] Kjønstad, E. F.; Myhre, R. H.; Martínez, T. J.; Koch, H. "Crossing conditions in coupled cluster theory", *J. Chem. Phys* **2017**, 147, 164105.

[70] Krylov, A. I. "Equation-of-Motion Coupled-Cluster Methods for Open-Shell and Electronically Excited Species: The Hitchhiker's Guide to Fock Space", *Annu. Rev. Phys. Chem.* **2008**, 59, 433–462.

[71] Krylov, A. I. "Size-consistent wave functions for bond-breaking: the equation-of-motion spin-flip model", *Chem. Phys. Lett.* **2001**, *338*, 375–384.

[72] Levchenko, S. V.; Krylov, A. I. "Equation-of-motion spin-flip coupled-cluster model with single and double substitutions: Theory and application to cyclobutadiene", *J. Chem. Phys.* **2004**, *120*, 175–185.

[73] Krylov, A. I. "Spin-Flip Equation-of-Motion Coupled-Cluster Electronic Structure Method for a Description of Excited States, Bond Breaking, Diradicals, and Triradicals", *Acc. Chem. Res.* **2006**, *39*, 83–91.

[74] Coriani, S.; Koch, H. "Erratum: "Communication: X-ray absorption spectra and core-ionization potentials within a core-valence separated coupled cluster framework" [J. Chem. Phys. 143, 181103 (2015)]", *J. Chem. Phys.* **2016**, *145*, 149901.

[75] Sekino, H.; Bartlett, R. J. "A linear response, coupled-cluster theory for excitation energy", *Int. J. Quantum Chem.* **1984**, *26*, 255–265.

[76] Geertsen, J.; Rittby, M.; Bartlett, R. J. "The equation-of-motion coupled-cluster method: Excitation energies of Be and CO", *Chem. Phys. Lett.* **1989**, *164*, 57–62.

[77] Stanton, J. F.; Bartlett, R. J. "The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties", *J. Chem. Phys.* **1993**, *98*, 7029–7039.

[78] Manohar, P. U.; Krylov, A. I. "A noniterative perturbative triples correction for the spin-flipping and spin-conserving equation-of-motion coupled-cluster methods with single and double substitutions", *J. Chem. Phys* **2008**, *129*, 194105.

[79] Öhrn, A.; Christiansen, O. "Electronic excitation energies of pyrimidine studied using coupled cluster response theory", *Phys. Chem. Chem. Phys.* **2001**, *3*, 730–740.

[80] Bomble, Y. J.; Sattelmeyer, K. W.; Stanton, J. F.; Gauss, J. "On the vertical excitation energy of cyclopentadiene", *J. Chem. Phys.* **2004**, *121*, 5236–5240.

[81] Bai, S.; Mansour, R.; Stojanović, L.; Toldo, J.; Barbatti, M. "On the Origin of the Shift Between Vertical Excitation and Band Maximum in Molecular Photoabsorption", *J. Mol. Model.* **2020**, *26*, 107.

[82] Edmiston, C.; Ruedenberg, K. "Localized Atomic and Molecular Orbitals", *Rev. Mod. Phys.* **1963**, 35, 457–464.

[83] Boys, S. F. "Construction of Some Molecular Orbitals to Be Approximately Invariant for Changes from One Molecule to Another", *Rev. Mod. Phys.* **1960**, *32*, 296–299.

[84] Pipek, J.; Mezey, P. G. "A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions", *J. Chem. Phys.* **1989**, 90, 4916–4926.

[85] Pulay, P. "Localizability of dynamic electron correlation", *Chem. Phys. Lett.* **1983**, *100*, 151–154.

[86] Edmiston, C.; Krauss, M. "Pseudonatural Orbitals as a Basis for the Superposition of Configurations. I. He2+", *J. Chem. Phys.* **1966**, 45, 1833–1839.

[87] Meyer, W. "Ionization energies of water from PNO-CI calculations", *Int. J. Quantum Chem.* **1971**, 5, 341–348.

[88] Meyer, W. "PNO–CI Studies of electron correlation effects. I. Configuration expansion by means of nonorthogonal orbitals, and application to the ground state and ionized states of methane", *J. Chem. Phys.* **1973**, 58, 1017–1035.

[89] Høyvik, I.-M.; Myhre, R. H.; Koch, H. "Correlated natural transition orbitals for core excitation energies in multilevel coupled cluster models", *J. Chem. Phys.* **2017**, *146*, 144109.

[90] Yang, W. "Direct calculation of electron density in density-functional theory", *Phys. Rev. Lett.* **1991**, *66*, 1438–1441.

[91] Prodan, E.; Kohn, W. "Nearsightedness of electronic matter", *Proc. Natl. Acad. Sci.* **2005**, *102*, 11635–11638.

[92] Zhang, I.; Jiang, J.; Gao, B.; Xu, X.; Luo, Y. “RRS-PBC: A molecular approach for periodic systems”, *Sci. China Chem* **2014**, 57, 1–6.

[93] Huang, C. “Embedded Cluster Density Approximation for Exchange–Correlation Energy: A Natural Extension of the Local Density Approximation”, *J. Chem. Theory Comput.* **2018**, 14, 6211–6225.

[94] Krause, C.; Werner, H.-J. “Comparison of explicitly correlated local coupled-cluster methods with various choices of virtual orbitals”, *Phys. Chem. Chem. Phys.* **2012**, 14, 7591–7604.

[95] Hansen, A. S.; Baardsen, G.; Rebolini, E.; Maschio, L.; Pedersen, T. B. “Representation of the virtual space in extended systems – a correlation energy convergence study”, *Mol. Phys.* **2020**, 118, e1733118.

[96] Høyvik, I.-M. “The spectrum of the atomic orbital overlap matrix and the locality of the virtual electronic density matrix”, *Mol. Phys.* **2020**, 118, e1765034.

[97] Myhre, R. H.; Sánchez de Merás, A. M. J.; Koch, H. “Multi-level coupled cluster theory”, *J. Chem. Phys.* **2014**, 141, 224105.

[98] Myhre, R. H.; Koch, H. “The multilevel CC3 coupled cluster model”, *J. Chem. Phys.* **2016**, 145, 044111.

[99] Baudin, P.; Kristensen, K. “LoFEx — A local framework for calculating excitation energies: Illustrations using RI-CC2 linear response theory”, *J. Chem. Phys.* **2016**, 144, 224106.

[100] Baudin, P.; Bykov, D.; Liakh, D.; Ettenhuber, P.; Kristensen, K. “A local framework for calculating coupled cluster singles and doubles excitation energies (LoFEx-CCSD)”, *Mol. Phys.* **2017**, 115, 2135–2144.

[101] Baudin, P.; Kristensen, K. “Correlated natural transition orbital framework for low-scaling excitation energy calculations (CorNFLEx)”, *J. Chem. Phys.* **2017**, 146, 214114.

[102] Folkestad, S. D.; Koch, H. “Multilevel CC2 and CCSD Methods with Correlated Natural Transition Orbitals”, *J. Chem. Theory Comput.* **2020**, 16, 179–189.

[103] Folkestad, S. D.; Kjønstad, E. F.; Goletto, L.; Koch, H. "Multilevel CC2 and CCSD in Reduced Orbital Spaces: Electronic Excitations in Large Molecular Systems", *J. Chem. Theory Comput.* **2021**, 17, 714–726.

[104] Sæther, S.; Kjærgaard, T.; Koch, H.; Høyvik, I.-M. "Density-Based Multilevel Hartree–Fock Model", *J. Chem. Theory Comput.* **2017**, 13, 5282–5290.

[105] Dunning, T. H. "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen", *J. Chem. Phys.* **1989**, 90, 1007–1023.

[106] Kendall, R. A.; Dunning, T. H.; Harrison, R. J. "Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions", *J. Chem. Phys.* **1992**, 96, 6796–6806.

[107] Woon, D. E.; Dunning, T. H. "Gaussian basis sets for use in correlated molecular calculations. V. Core-valence basis sets for boron through neon", *J. Chem. Phys.* **1995**, 103, 4572–4585.

[108] Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. "A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach", *Chem. Phys.* **1980**, 48, 157 – 173.

[109] Buenker, R. J.; Peyerimhoff, S. D.; Butscher, W. "Applicability of the multi-reference double-excitation CI (MRD-CI) method to the calculation of electronic wavefunctions and comparison with related techniques", *Mol. Phys.* **1978**, 35, 771–791.

[110] Langhoff, S. R.; Davidson, E. R. "Configuration interaction calculations on the nitrogen molecule", *Int. J. Quantum Chem.* **1974**, 8, 61–72.

[111] Duch, W.; Diercksen, G. H. F. "Size-extensivity corrections in configuration interaction methods", *J. Chem. Phys.* **1994**, 101, 3018–3030.

[112] Andersson, K.; Malmqvist, P.; Roos, B. O. "Second-order perturbation theory with a complete active space self-consistent field reference function", *J. Chem. Phys.* **1992**, 96, 1218–1226.

[113] Finley, J.; Åke Malmqvist, P.; Roos, B. O.; Serrano-Andrés, L. "The multi-state CASPT2 method", *Chem. Phys. Lett.* **1998**, 288, 299 – 306.

[114] Ghigo, G.; Roos, B. O.; Åke Malmqvist, P. "A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2)", *Chem. Phys. Lett.* **2004**, 396, 142 – 149.

[115] Malmqvist, P. Å.; Pierloot, K.; Shahi, A. R. M.; Cramer, C. J.; Gagliardi, L. "The restricted active space followed by second-order perturbation theory method: Theory and application to the study of CuO2 and Cu2O2 systems", *J. Chem. Phys* **2008**, 128, 204109.

[116] Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. "Introduction of n-electron valence states for multireference perturbation theory", *J. Chem. Phys.* **2001**, 114, 10252–10264.

[117] Angeli, C.; Borini, S.; Cestari, M.; Cimiraglia, R. "A quasidegenerate formulation of the second order n-electron valence state perturbation theory approach", *J. Chem. Phys.* **2004**, 121, 4043–4049.

[118] Angeli, C.; Pastore, M.; Cimiraglia, R. "New perspectives in multireference perturbation theory: The n-electron valence state approach", *Theor. Chem. Acc.* **2007**, 117, 743–754.

[119] Granovsky, A. A. "Extended multi-configuration quasi-degenerate perturbation theory: The new approach to multi-state multi-reference perturbation theory", *J. Chem. Phys* **2011**, 134, 214113.

[120] Sobolewski, A. L. "Reversible molecular switch driven by excited-state hydrogen transfer", *Phys. Chem. Chem. Phys.* **2008**, 10, 1243–1247.

[121] Crecca, C. R.; Roitberg, A. E. "Theoretical Study of the Isomerization Mechanism of Azobenzene and Disubstituted Azobenzene Derivatives", *J. Phys. Chem. A* **2006**, 110, 8188–8203.

[122] Send, R.; Sundholm, D. "Stairway to the Conical Intersection: A Computational Study of the Retinal Isomerization", *J. Phys. Chem. A* **2007**, 111, 8766–8773.

[123] Chmura, B.; Rode, M. F.; Sobolewski, A. L.; Lapinski, L.; Nowak, M. J. "A Computational Study on the Mechanism of Intramolecular Oxo–Hydroxy Phototautomerism Driven by Repulsive $\pi\sigma^*$ State", *J. Phys. Chem. A* **2008**, 112, 13655–13661.

[124] Send, R.; Sundholm, D.; Johansson, M. P.; Pawłowski, F. "Excited State Potential Energy Surfaces of Polyenes and Protonated Schiff Bases", *J. Chem. Theory Comput.* **2009**, 5, 2401–2414.

[125] Plötner, J.; Dreuw, A. "Molecular Mechanism of the Z/E-Photoisomerization of Hemithioindigo Hemistilbene", *The Journal of Physical Chemistry A* **2009**, 113, 11882–11887.

[126] Rode, M. F.; Sobolewski, A. L.; Dedonder, C.; Jouvet, C.; Dopfer, O. "Computational Study on the Photophysics of Protonated Benzene", *J. Phys. Chem. A* **2009**, 113, 5865–5873.

[127] Jiang, C.-W.; Xie, R.-H.; Li, F.-L.; Allen, R. E. "Ultrafast cis-to-trans photoisomerization of a bridged azobenzene through nππ excitation: Rotational pathway is not restricted", *Chem. Phys. Lett.* **2012**, 521, 107 – 112.

[128] Piecuch, P.; Włoch, M. "Renormalized coupled-cluster methods exploiting left eigenstates of the similarity-transformed Hamiltonian", *J. Chem. Phys.* **2005**, 123, 224105.

[129] Manohar, P. U.; Stanton, J. F.; Krylov, A. I. "Perturbative triples correction for the equation-of-motion coupled-cluster wave functions with single and double substitutions for ionized states: Theory, implementation, and examples", *J. Chem. Phys* **2009**, 131, 114112.

[130] Matthews, D. A.; Stanton, J. F. "A new approach to approximate equation-of-motion coupled cluster with triple excitations", *J. Chem. Phys.* **2016**, 145, 124102.

[131] Gozem, S.; Krylov, A. I.; Olivucci, M. "Conical Intersection and Potential Energy Surface Features of a Model Retinal Chromophore: Comparison of EOM-CC and Multireference Methods", *J. Chem. Theory Comput.* **2013**, 9, 284–292.

[132] Hättig, C.; Hald, K. "Implementation of RI-CC2 triplet excitation energies with an application to trans-azobenzene", *Phys. Chem. Chem. Phys.* **2002**, 4, 2111–2118.

[133] Fliegl, H.; Köhn, A.; Hättig, C.; Ahlrichs, R. "Ab Initio Calculation of the Vibrational and Electronic Spectra of trans- and cis-Azobenzene", *J. Am. Chem. Soc.* **2003**, 125, 9821–9827.

[134] Wang, L.; Xu, W.; Yi, C.; Wang, X. “Isomerization and electronic relaxation of azobenzene after being excited to higher electronic states”, *J. Mol. Graphics Modell.* **2009**, *27*, 792 – 796.

[135] Nenov, A.; et al.. “UV-Light-Induced Vibrational Coherences: The Key to Understand Kasha Rule Violation in trans-Azobenzene”, *J. Phys. Chem. Lett.* **2018**, *9*, 1534–1541.

[136] Vetráková, L.; Ladányi, V.; Al Anshori, J.; Dvořák, P.; Wirz, J.; Heger, D. “The Absorption Spectrum of cis-Azobenzene”, *Photochem. Photobiol. Sci.* **2017**, *16*, 1749–1756.

[137] Marsman, M.; Grüneis, A.; Paier, J.; Kresse, G. “Second-order Møller–Plesset perturbation theory applied to extended systems. I. Within the projector-augmented-wave formalism using a plane wave basis set”, *J. Chem. Phys.* **2009**, *130*, 184103.

[138] Rebolini, E.; Baardsen, G.; Hansen, A. S.; Leikanger, K. R.; Pedersen, T. B. “Divide–Expand–Consolidate Second-Order Møller–Plesset Theory with Periodic Boundary Conditions”, *J. Chem. Theory Comput.* **2018**, *14*, 2427–2438.

[139] McClain, J.; Sun, Q.; Chan, G. K.-L.; Berkelbach, T. C. “Gaussian-Based Coupled-Cluster Theory for the Ground-State and Band Structure of Solids”, *J. Chem. Theory Comput.* **2017**, *13*, 1209–1218.

[140] Pisani, C.; Busso, M.; Capecchi, G.; Casassa, S.; Dovesi, R.; Maschio, L.; Zicovich-Wilson, C.; Schütz, M. “Local-MP2 electron correlation method for nonconducting crystals”, *J. Chem. Phys.* **2005**, *122*, 094113.

[141] Høyvik, I.-M.; Kristensen, K.; Kjaergaard, T.; Jørgensen, P. “A perspective on the localizability of Hartree–Fock orbitals”, *Theor. Chem. Acc.* **2014**, *133*.

[142] Warshel, A.; Karplus, M. “Calculation of ground and excited state potential surfaces of conjugated molecules. I. Formulation and parametrization”, *J. Am. Chem. Soc.* **1972**, *94*, 5612–5625.

[143] Levitt, M.; Warshel, A. “Computer simulation of protein folding”, *Nature* **1975**, *253*, 694–698.

[144] Tomasi, J.; Mennucci, B.; Cammi, R. "Quantum Mechanical Continuum Solvation Models", *Chem. Rev.* **2005**, *105*, 2999–3094.

[145] Mennucci, B. "Polarizable continuum model", *WIREs Comput. Mol. Sci.* **2012**, *2*, 386–404.

[146] Sloan, E. D. "Fundamental principles and applications of natural gas hydrates", *Nature* **2003**, *426*, 353—363.

[147] Tse, J. S.; Ratcliffe, C. I.; Powell, B. M.; Sears, V. F.; Handa, Y. P. "Rotational and Translational Motions of Trapped Methane. Incoherent Inelastic Neutron Scattering of Methane Hydrate", *J. Phys. Chem.* A **1997**, *101*, 4491–4495.

[148] Gutt, C.; Asmussen, B.; Press, W.; Merkl, C.; Casalta, H.; Greinert, J.; Bohrmann, G.; Tse, J. S.; Hüller, A. "Quantum rotations in natural methane-clathrates from the Pacific sea-floor", *Europhys. Lett.* **1999**, *48*, 269–275.

[149] Kamiyama, T.; Seki, N.; Iwasa, H.; Uchida, T.; Ebinuma, T.; Narita, H.; Igawa, N.; Ishii, Y.; Bennington, S.; Kiyanagi, Y. "Methane molecular motion in clathrate hydrate host framework", *Phys. B: Condens. Matter.* **2006**, *385-386*, 202–204.

[150] Ikeda, T.; Yamamuro, O.; Matsuo, T.; Mori, K.; Torii, S.; Kamiyama, T.; Izumi, F.; Ikeda, S.; Mae, S. "Neutron diffraction study of carbon dioxide clathrate hydrate", *J. Phys. Chem. Solids* **1999**, *60*, 1527–1529.

[151] Sum, A. K.; Burruss, R. C.; Sloan, E. D. "Measurement of Clathrate Hydrates via Raman Spectroscopy", *J. Phys. Chem. B.* **1997**, *101*, 7371–7377.

[152] Subramanian, S.; Kini, R.; Dec, S.; Sloan, E. "Evidence of structure II hydrate formation from methane+ethane mixtures", *Chem. Eng. Sci.* **2000**, *55*, 1981–1999.

[153] Uchida, T.; Takeya, S.; Kamata, Y.; Ohmura, R.; Narita, H. "Spectroscopic Measurements on Binary, Ternary, and Quaternary Mixed-Gas Molecules in Clathrate Structures", *Ind. Eng. Chem. Res.* **2007**, *46*, 5080–5087.

[154] Ota, M.; Saito, T.; Aida, T.; Watanabe, M.; Sato, Y.; Smith Jr., R. L.; Inomata, H. "Macro and microscopic CH4–CO2 replacement in CH4 hydrate under pressurized CO2", *AIChE J.* **2007**, 53, 2715–2721.

[155] Murshed, M. M.; Kuhs, W. F. "Kinetic Studies of Methane–Ethane Mixed Gas Hydrates by Neutron Diffraction and Raman Spectroscopy", *J. Phys. Chem. B.* **2009**, 113, 5172–5180.

[156] Klapp, S. A.; Murshed, M. M.; Pape, T.; Klein, H.; Bohrmann, G.; Brewer, P. G.; Kuhs, W. F. "Mixed gas hydrate structures at the Chapopote Knoll, southern Gulf of Mexico", *Earth Planet. Sci. Lett.* **2010**, 299, 207–217.

[157] Ohno, H.; Kida, M.; Sakurai, T.; Iizuka, Y.; Hondoh, T.; Narita, H.; Nagao, J. "Symmetric Stretching Vibration of CH4 in Clathrate Hydrate Structures", *ChemPhysChem* **2010**, 11, 3070–3073.

[158] Lee, H.-H.; Ahn, S.-H.; Nam, B.-U.; Kim, B.-S.; Lee, G.-W.; Moon, D.; Shin, H. J.; Han, K. W.; Yoon, J.-H. "Thermodynamic Stability, Spectroscopic Identification, and Gas Storage Capacity of CO2–CH4–N2 Mixture Gas Hydrates: Implications for Landfill Gas Hydrates", *Environ. Sci. Technol.* **2012**, 46, 4184–4190.

[159] Chazallon, B.; Pirim, C. "Selectivity and CO2 capture efficiency in CO2-N2 clathrate hydrates investigated by in-situ Raman spectroscopy", *Chem. Eng. J.* **2018**, 342, 171–183.

[160] Ripmeester, J. A.; Ratcliffe, C. I. "Low-temperature cross-polarization/magic angle spinning carbon-13 NMR of solid methane hydrates: structure, cage occupancy, and hydration number", *J. Phys. Chem.* A **1988**, 92, 337–339.

[161] Seo, Y.-T.; Lee, H. "Structure and Guest Distribution of the Mixed Carbon Dioxide and Nitrogen Hydrates As Revealed by X-ray Diffraction and 13C NMR Spectroscopy", *J. Phys. Chem. B.* **2004**, 108, 530–534.

[162] Seo, Y.; Kang, S.-P.; Jang, W. "Structure and Composition Analysis of Natural Gas Hydrates: 13C NMR Spectroscopic and Gas Uptake Measurements of Mixed Gas Hydrates", *J. Phys. Chem.* A **2009**, 113, 9641–9649.

[163] Lee, S.; Seo, Y. "Experimental Measurement and Thermodynamic Modeling of the Mixed CH4 + C3H8 Clathrate Hydrate Equilibria in Silica Gel Pores: Effects of Pore Size and Salinity", *Langmuir* **2010**, *26*, 9742–9748.

[164] Lee, J.-W.; Lu, H.; Moudrakovski, I. L.; Ratcliffe, C. I.; Ohmura, R.; Alavi, S.; Ripmeester, J. A. "13C NMR Studies of Hydrocarbon Guests in Synthetic Structure H Gas Hydrates: Experiment and Computation", *J. Phys. Chem.* A **2011**, *115*, 1650–1657.

[165] Siuda, P.; Sadlej, J. "Nuclear Magnetic Resonance Parameters for Methane Molecule Trapped in Clathrate Hydrates", *J. Phys. Chem.* A **2011**, *115*, 612–619.

[166] Kida, M.; Hori, A.; Sakagami, H.; Takeya, S.; Kamata, Y.; Takahashi, N.; Ebinuma, T.; Narita, H. "13C Chemical Shifts of Propane Molecules Encaged in Structure II Clathrate Hydrate", *J. Phys. Chem.* A **2011**, *115*, 643–647.

[167] Ohno, H.; Moudrakovski, I.; Gordienko, R.; Ripmeester, J.; Walker, V. K. "Structures of Hydrocarbon Hydrates during Formation with and without Inhibitors", *J. Phys. Chem.* A **2012**, *116*, 1337–1343.

[168] Lee, J.-W.; Lee, J.; Kang, S.-P. "13C NMR spectroscopies and formation kinetics of gas hydrates in the presence of monoethylene glycol as an inhibitor", *Chem. Eng. Sci.* **2013**, *104*, 755–759.

[169] Halpern, Y.; Thieu, V.; Henning, R. W.; Wang, X.; Schultz, A. J. "Time-Resolved in Situ Neutron Diffraction Studies of Gas Hydrate: Transformation of Structure II (sII) to Structure I (sI)", *J. Am. Chem. Soc.* **2001**, *123*, 12826–12831.

[170] Everett, S.; Rawn, C.; Chakoumakos, B.; Keffer, D.; Huq, A.; Phelps, T. "Insights into the structure of mixed CO2/CH4 in gas hydrates", *Am. Mineral.* **2015**, *100*, 1203–1208.

[171] Bowron, D. T.; Filipponi, A.; Roberts, M. A.; Finney, J. L. "Hydrophobic Hydration and the Formation of a Clathrate Hydrate", *Phys. Rev. Lett.* **1998**, *81*, 4164–4167.

[172] Bowron, D. T.; Weigel, R.; Filipponi, A.; Roberts, M. A.; Finney, J. L. “X-Ray absorption spectroscopy investigations of the hydrophobic hydration of krypton at high pressure”, *Mol. Phys.* **2001**, 99, 761–765.

[173] Cederbaum, L. S.; Domcke, W.; Schirmer, J. “Many-body theory of core holes”, *Phys. Rev.* A **1980**, *22*, 206–222.

Paper A

Describing ground and excited state potential energy surfaces for molecular photoswitches using coupled cluster models

A. Hutcheson, A. C. Paul, R. H. Myhre, I-M. Høyvik

J. Comput. Chem. **2021**, 42(20), 1419;
doi: 10.1063/5.0004713

Received: 16 December 2020 | Revised: 19 March 2021 | Accepted: 20 April 2021

DOI: 10.1002/jcc.26553

FULL PAPER

Journal of COMPUTATIONAL CHEMISTRY WILEY

Describing ground and excited state potential energy surfaces for molecular photoswitches using coupled cluster models

Anders Hutcheson[1] | Alexander Christian Paul[1] | Rolf H. Myhre[1] | Henrik Koch[2] | Ida-Marie Høyvik[1]

[1]Department of Chemistry, The Norwegian University of Science and Technology, Trondheim, Norway

[2]Scuola Normale Superiore, Pisa, Italy

Correspondence
Ida-Marie Høyvik, Department of Chemistry, The Norwegian University of Science and Technology, Høgskoleringen 5, 7491 Trondheim, Norway.
Email: ida-marie.hoyvik@ntnu.no

Funding information
Marie Sklodowska-Curie European Training Network, Grant/Award Number: 765739; Norges Forskningsråd, Grant/Award Numbers: 263110, 275506

Abstract

In this article, we use two extensively studied systems, a retinal model system and azobenzene, to explore the use of coupled cluster models for describing ground and singlet excited state potential energy surfaces of photoswitchable systems. While not being suitable for describing nuclear dynamics of photoisomerization, coupled cluster models have useful attributes, such as the inclusion of dynamical correlation, their black box nature, and the systematic improvement offered by truncation level. Results for the studied systems show that when triple excitations (here through the CC3 model) are included, ground and excited state potential energy surfaces for isomerization paths may reliably be generated, also for states of doubly excited character. For ground state equilibrium cis- and trans-azobenzene, the molecular geometry and basis set is seen to significantly impact the vertical excitation energies for the two lowest excited states. Efficient implementations of coupled cluster models can therefore constitute valuable tools for investigating photoswitchable systems and can be used for preliminary black box studies to gather information before more complicated excited state dynamics approaches are pursued.

KEYWORDS
azobenzene, coupled cluster, excited states, penta-2,4-dieniminium cation, photoisomerization

1 | INTRODUCTION

Photoswitches constitute a class of molecular machines that undergo reversible isomerization upon stimulus by light, so-called photoisomerization. Computational approaches for photoisomerization processes are considered to be of great value for the study of photoswitches, especially in conjunction with spectroscopic techniques.[1–3] To get a more complete overview of the isomerization mechanism itself and the interplay of several electronic states, accurate ab initio molecular dynamics are needed. However, the computational costs of such calculations are high, and they are complicated by the frequent occurrence of conical intersections,[4] which require special attention to the chosen method. From ground and excited state electronic structure calculations, the mechanism itself cannot be fully elucidated, but electronic structure theory provides valuable information about the potential energy surfaces and may indicate possible reaction paths for the isomerization process.

Many features of the potential energy surface are important for the machinery of photoswitches. Ground state surfaces give information on relative stability of isomers, relaxation processes following decay from excited state and energy barriers for thermal isomerization. The energy barriers will indicate whether a molecule may

function as a photoswitch at all (high enough energy barrier in one direction) and if the process is thermally reversible. Excited states naturally play a key role in the mechanisms of photoisomerization, with curvature and shapes giving information on the efficiency of possible decay pathways to the ground state and whether the process is photochemically reversible. The energy of the excited states relative to the ground state energy will give information on wavelengths needed for initiating (and for certain photoswitches, reversing) the isomerization process. Mapping out potential energy surfaces for ground and excited states can also yield information on and characterize conical intersections, depending on the choice of method.

Due to the occurrence of conical intersections and near-degenerate states in photoisomerization processes, multiconfigurational self-consistent field methods, such as the complete active space SCF (CASSCF)[5] methods, as well as multireference methods including dynamical electron correlation, such as multireference configuration interaction with singles and doubles (MRCISD),[6] MRCISD with a Davidson-type correction (MRCISD+Q),[7,8] (multi-state) complete active space second order perturbation theory (CASPT2)[9,10] with and without a ionization potential - electron affinity shift (IPEA),[11] restricted active space second order perturbation theory (RASPT2),[12] (quasi-degenerate) n-electron valence state second order perturbation theory (NEVPT2),[13–15] and extended multiconfiguration quasi degenerate second order perturbation theory (XMCQDPT2),[16] have been used extensively.[1,17–33] Single-reference methods such as time dependent density functional theory (TDDFT),[34] Møller–Plesset perturbation theory,[35] and various members of the coupled cluster hierarchy of wave function models, have also been used.[24,36–44] One advantage of single-reference methods compared to multireference methods is that no active space selection is necessary. Hence, the quality of the results will not depend on whether the appropriate orbital space has been chosen by the user. Furthermore, although single-reference coupled cluster methods cannot describe near-degeneracies, coupled cluster models with at least triples excitations can describe electronic states with doubly excited character that are important for organic photochemistry. For this purpose, approximate triples schemes such as the iterative CC3 model,[45] or the non-iterative equation-of-motion (EOM) CCSD(dT), EOM-CCSD(fT)[46–48] and EOM-CCSD(T)(a)/EOM-CCSD(T)(a)*[49] models can be used in place of CCSDT.[50] However, for excited electronic states dominated by single-electron excitations, low-level truncations of the hierarchy, for example, CCSD[51] or approximate versions such as CC2,[52] can be used. Hence, the adequacy of coupled cluster models for photoswitches will depend on the chosen truncation level.

The coupled cluster Jacobian, which is diagonalized to obtain excited states, is nonsymmetric and may become defective at conical intersections.[53] This precludes the standard models from being used for ab initio excited state dynamics of photochemical systems. However, coupled cluster has advantages for the characterization of ground and excited state surfaces, such as the inclusion of dynamical correlation and the possibility for systematic improvement offered by choice of truncation level. In addition, coupled cluster models are black box in nature and the user does not have to define active spaces or parameters. However, due to the computational cost of including triple (or approximate triple) excitations to the hierarchy, there are few such results for potential energy surfaces of photoswitches in the literature. A notable exception is the use of EOM spin-flip (SF) CCSD with perturbative triples corrections (EOM-SF-CCSD(dT) and EOM-SF-CCSD(fT)) for the penta-2,4-dieniminium cation (PSB3) model system of retinal, which shows excellent agreement with MRCISD+Q results.[54] The spin-flip approach exploits the fact that a high-spin triplet reference is well behaved across the bond torsion and generates singlet ground and excited states as spin-flip excitations from the triplet reference. From this study, it is clear that inclusion of triples leads to a significantly better agreement with MRCISD+Q compared to EOM-SF-CCSD; however, it is unclear whether the high-spin reference is necessary to obtain this accuracy. CC3 has also been used for the PSB3 system,[44] but without comparison to high-level multireference methods.

For PSB3, the extensive set of MRCISD+Q calculations by Olivucci and collaborators[27] have proven indispensable for benchmarking other computational methods.[54,55] Comparisons have been made between these MRCISD+Q results and EOM variants of coupled cluster such as EOM-CCSD, EOM-SF-CCSD, EOM-SF-CCSD (dT), and EOM-SF-CCSD(fT).[54] Based on this comparison, EOM-SF-CCSD(dt/ft) was shown to agree well with MRCISD+Q. The importance of dynamical electron correlation has also been shown for PSB3, where CASSCF gave a barrierless first singlet excited state (S_1) potential energy surface while CASPT2 yields an S_1 surface with a small barrier.[56] Introduction of dynamical correlation by performing single point calculations along the CASSCF S_1 surface did not correct the shape of the potential energy surface. In addition, Send et al.[44] investigated excited state surfaces generated by rotation around all bonds in PSB3 using TDDFT, CC2, and CC3. From their results, it appears that there are significant differences between the coupled cluster methods and TDDFT, particularly for rotations around single bonds.

Another photoswitch that has garnered much attention is azobenzene and its derivatives. Due to the size of azobenzene, limitations are imposed on choice of computational method and/or basis set. Zhu et al.[31,32] optimized equilibrium states, transition states and conical intersections with five state averaged (5SA) CASSCF (6,6)/6-31G,[57] for the ground state and the first three singlet excited states. The focus of the work in Reference 32 was primarily to carry out molecular dynamics, however, a rotational path was generated by interpolation, which can be used to compare other computational methods with CASSCF. It should be noted that the CASSCF excitation energies of the cis and trans structures in Refs. 31 and 32 deviate significantly from experimental values, which can be explained by the fact that CASSCF excitation energies have been found to have a large deviation from more accurate methods.[58] Recently, a thorough mapping of the ground state and first excited state potential energy surfaces of azobenzene was performed by Aleotti et al.[33] using RASPT2. In this case, the excitation energies of the cis and trans structures are in much better agreement with experimental values. This shows the importance of dynamical electron correlation when computing

excitation energies for comparison with experimental results. Casellas et al.[30] further argue that dynamical electron correlation must also be included for geometry optimization of critical points on the lowest singlet excited states. In addition to the studies mentioned, there are several studies of azobenzene using CASSCF and other multireference techniques containing dynamical electron correlation.[3,18,21–23,25,28,59] With respect to single-reference methods, CC2 and CCSD calculations of excitation energies have been performed for azobenzene by Hättig and coworkers.[38,39] However, high-level single-reference calculations with at least triples corrections are not presented in literature for azobenzene, and hence experimental results[60] provide the basis for evaluating computational results.

The focus of this article is to explore the capabilities of the coupled cluster models for mapping singlet potential energy surfaces of photoswitches. We utilize the newly released and efficient coupled cluster code in the eT program[61] to generate potential energy surfaces for CC2, CCSD, and CC3. As already mentioned, CC2 and CCSD will not give good results for excitations of strong doubly excited character, but for single-excitation dominated processes they are expected to give good results.[62] In this study, the performance of CC2, CCSD, and CC3 are presented for torsional angles that include potentially problematic regions such as isomerization points. The photoswitches studied are PSB3, which displays torsion of a C—C double bond, and azobenzene where we study the torsion about the N—N double bond. The large number of studies facilitates evaluation of the performance of the coupled cluster models. The advantage of CC3 is the fact that it can describe excitations of doubly excited character and that results are not complicated by choice of an active orbital space. With the new CC3 implementation (see, Reference 63) photoswitches of about 20 non-hydrogen atoms can be treated using augmented basis sets, such as aug-cc-pVDZ, using high performance computing resources. The inclusion of diffuse functions has been shown to significantly impact vertical excitation energies,[64] hence being able to perform CC3 calculations with augmented basis sets is valuable.

The article is structured as follows. In Section 2.1, we provide computational details. In Section 2.2, we present potential energy surfaces of PSB3 generated by CC2, CCSD, and CC3 and compare to MRCISD+Q[27] and spin-flip coupled cluster methods.[54] Further, we explore the basis set effect for the potential energy surfaces for PSB3. In Section 2.3, we explore how coupled cluster model, basis set, molecular geometry, and solvent effects affect vertical excitation energies for cis and trans ground state equilibrium azobenzene structures. Finally, CC3 potential energy surfaces for isomerization through CNNC torsion for azobenzene are presented in Section 2.4. Concluding remarks are given in Section 3.

2 | RESULTS AND DISCUSSION

2.1 | Computational details

In this section, the software and settings used for computations are presented. CC2, CCSD, and CC3 excitation energies were computed with the eT program.[61] All coupled cluster calculations presented are computed using frozen core. The decomposition of the electron repulsion integrals[65] threshold was set to 10^{-10}. The energy and gradient thresholds for Hartree–Fock and the energy and residual thresholds for the coupled cluster ground state were all set to 10^{-8}. The energy and residual threshold for the coupled cluster excited states were set to 10^{-4}. In coupled cluster calculations where the solvent effects of methanol were accounted for by a polarizable continuum model (PCM),[66] the static permittivity was set to 32.63, the optical permittivity was set to 1.758 and the probe radius was set to 1.855 Å.

2.2 | The penta-2,4-dieniminium cation (PSB3)

In this section, we discuss CC2, CCSD, and CC3 results for CASSCF optimized penta-2,4-dieniminium cation (PSB3) structures presented by Olivucci and collaborators.[27] The three relevant isomerization paths are depicted schematically in Figure 1. The charge transfer (CT) path and covalent/diradical (DIR) path are both rotational minimum energy paths connecting the transition states (TS_{CT} and TS_{DIR}) to the cis-reactant and trans-product. The paths differ in the character of the wave functions, with the DIR path being dominated by a covalent/diradical wave function along the entire isomerization path and the CT path being dominated by a charge transfer wave function close to TS_{CT}, but otherwise being dominated by a covalent/diradical wave function. For these two paths, calculations are performed for the transition states and eight geometries on either side of the transition state, with the dihedral angle in the range $\sim 82^{\circ}$–103°. The bond length alternation (BLA) path corresponds to a coordinate obtained by linear interpolation and extrapolation of the two transition state structures (TS_{CT} and TS_{DIR}) and intercepts a single conical intersection. The bond length alternation coordinate is defined as the difference between the average bond lengths of formal double and single bonds, and has a positive value at TS_{CT} and negative values at TS_{DIR}. For details, see Reference 27. These structures are well suited for testing the capabilities of the coupled cluster methods, since extensive literature on both multireference and equation of motion-coupled cluster (with and without spin-flip) exists. The MRCISD+Q/6-31G* results from Reference 27 are used as reference for evaluating the coupled cluster methods. Although 6-31G* is a rather small basis set for describing excited states, results in Reference 27 are presented for all paths using 6-31G*. It is, therefore, useful in the context of evaluating whether coupled cluster models can perform on par with MRCISD +Q. To assess the shape of surfaces generated by coupled cluster models, the nonparallelity value is computed for the S_0 and S_1 surfaces for all paths. The nonparallelity value is given by,

$$\Delta E_{max} - \Delta E_{min} \tag{1}$$

where ΔE_{max} and ΔE_{min} is the maximum and minimum energy difference for a potential energy surface computed in two different manners, respectively. Nonparallelity values are in Sections 2.2.1–2.2.3 used to evaluate the shape of the coupled cluster surfaces relative to the benchmark results. In Section 2.2.4, we use nonparallelity values

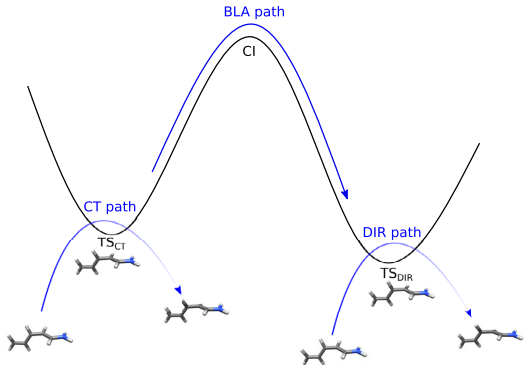

FIGURE 1 Schematic representation of the CT and DIR path, which are minimum energy paths in the cis → trans isomerization, and the BLA path, which is a path that interpolates the transition states of the CT and DIR paths

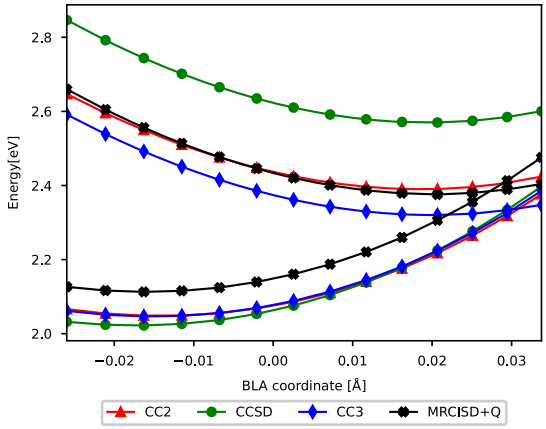

FIGURE 2 S_0 and S_1 energies for BLA-path relative to cis-PSB3 computed with CC2, CCSD and CC3 with 6-31G*. MRCISD+Q with 6-31G* benchmark results from Reference 27 are included for comparison

to evaluate how coupled cluster surfaces change upon changing the basis set.

2.2.1 | BLA path

In this section, we consider results for the BLA path using 6-31G*. CC2, CCSD, and CC3 results are plotted in Figure 2 together with the MRCISD+Q benchmark results of Reference 27, and energies are relative to cis-PSB3. Additionally, nonparallelity values of the coupled cluster models relative to MRCISD+Q are presented in Table 1. From Figure 2, we first of all see that CCSD cannot describe the BLA path. The deviation in vertical excitation energies relative to MRCISD+Q is found to vary between 0.13 and 0.28 eV. The conical intersection cannot be seen for the CCSD surfaces, as it is shifted beyond TS_{DIR} and outside the range of the plot. However, the nonparallelity values for the S_0 and S_1 surfaces for the BLA path computed with CCSD are only of 0.02 eV and 0.01 eV (see Table 1), so the shape of the S_0 and S_1 surfaces are similar to those from the MRCISD+Q benchmark results. This is consistent with EOM-CCSD results from Reference 54, where the nonparallelity values for the S_0 and S_1 surfaces where 0.02 eV and 0.01 eV, respectively. CC2 energies are closer to the MRCISD+Q reference values than CCSD, as is seen from both Figure 2 and by comparing vertical excitation energies. The conical intersection for CC2 is shifted closer to that of the MRCISD+Q relative to CCSD, although still beyond TS_{DIR} and outside of the plot. With respect to the vertical excitation energies, the CC2 results deviate between 0.02 and 0.11 eV from the MRCISD+Q results, which is smaller than the deviations for CCSD. However, the nonparallelity values are larger for CC2 than for CCSD, with a value of 0.04 eV for the S_0 surface and 0.03 eV for the S_1 surface. We see from Figure 2 that the CC3 results for the S_0 and S_1 surfaces appear to be quite similar to the MRCISD+Q results, except that they are shifted downward in energy by approximately 0.07 eV. The CC3 vertical excitation energies differ by only 0.01–0.03 eV compared to the MRCISD+Q results and the nonparallelity values are small, with a value of 0.03 eV for S_0 and 0.01 eV for S_1. Hence, the CC3 surfaces are similar to the MRCISD+Q surfaces, also with respect to the position of the conical intersection. CC3 also performs similarly to EOM-SF-CCSD(dT) from Reference 54, as can be seen from the nonparallelity values given in Table 1. CC3 shows a slightly better agreement for the S_0 surface, while for the S_1 surface the agreement depends on whether the nonparallelity value of EOM-SF-CCSD(dT) with or without kinks is used (see Reference 54). Small differences are also seen when comparing CC3 vertical excitation energy of TS_{CT} and TS_{DIR} with either EOM-SF-CCSD(dT) or EOM-SF-CCSD(fT) from Reference 54. For TS_{CT} the vertical excitation energy is 0.45 eV, 0.48 eV, and 0.43 eV, for CC3, EOM-SF-CCSD(dT), and EOM-SF-CCSD(fT), respectively. For TS_{DIR}, the vertical excitation energy is 0.05 eV, 0.03 eV, and 0.03 eV, for CC3, EOM-SF-CCSD(dT), and EOM-SF-CCSD(fT), respectively. Hence, we see that accurate results may be obtained without using a well-behaved high-spin triplet reference.

2.2.2 | CT path

In this section, we consider results for the CT path using 6-31G*. CC2, CCSD, and CC3 results are plotted in Figure 3 together with the MRCISD+Q benchmark results of Reference 27, and energies are relative to cis-PSB3. From Figure 3, we see that CCSD provides too large excitation energies, as was the case for the BLA path. The difference in vertical excitation energies between CCSD and MRCISD+Q is in the range 0.23–0.28 eV, which is similar to that of the BLA path (0.13–0.28 eV). However, CCSD is doing a worse job reproducing the

TABLE 1 The nonparallelity values for CC2, CCSD, CC3 for the S_0 and S_1 states along the DIR, CT and BLA paths, given in eV

Method	DIR S_0	CT S_0	BLA S_0	DIR S_1	CT S_1	BLA S_1
CC2	0.06	0.03	0.04	0.22	0.22	0.03
CCSD	0.16	0.14	0.02	0.11	0.10	0.01
CC3	0.06	0.03	0.03	0.03	0.04	0.01
EOM-SF-CCSD(dT)	0.06	0.02	0.06 (0.04)	0.03	0.02	0.04 (0.01)

Note: The nonparallelity values are computed relative to the MRCISD+Q/6-31G* benchmark results from Reference 27. Nonparallelity values for EOM-SF-CCSD(dT) are reproduced from Reference 54, where the values in parentheses are calculated by ignoring the kinks along the BLA path. The EOM-SF-CCSD(dT) used a restricted open shell Hartree–Fock reference state.

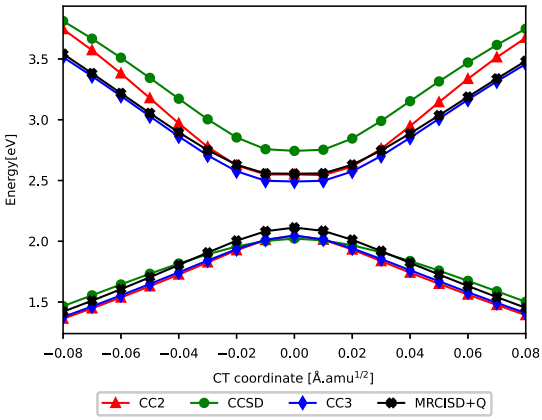

FIGURE 3 S_0 and S_1 energies for CT-path relative to cis-PSB3 computed with CC2, CCSD and CC3 with 6-31G*. MRCISD+Q with 6-31G* benchmark results from Reference 27 are included for comparison

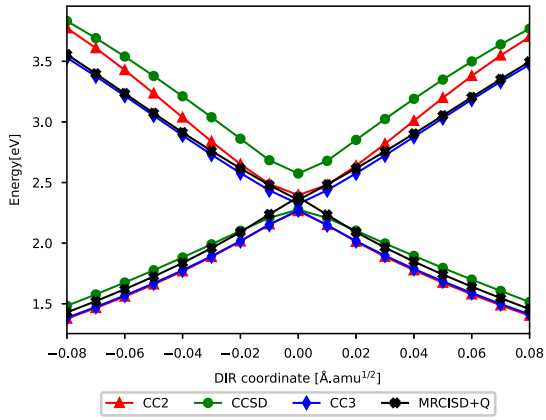

FIGURE 4 S_0 and S_1 energies for the DIR-path relative to cis-PSB3 computed with CC2, CCSD and CC3 with 6-31G*. MRCISD+Q with 6-31G* benchmark results from Reference 27 are included for comparison

shapes of the S_0 and S_1 surfaces than for the BLA path, as can be seen from the nonparallelity values (see Table 1). The nonparallelity values of the CT path with CCSD are one order of magnitude larger than for the BLA path for both surfaces. Considering the CC2 surfaces, we see from Figure 3 that CC2 provides a better agreement with MRCISD+Q for some of the vertical excitation energies than CCSD, with deviations ranging from 0.06 eV (close to CT coordinate 0.0) to 0.26 eV at CT coordinates ±0.08 Å·amu$^{1/2}$. CC2 has a nonparallelity value for the S_0 surface of 0.03 eV, whereas the nonparallelity value for S_1 surface is 0.22 eV. Hence, CC2 can reproduce the shape of the ground state, but not the first excited state. The improved performance of CC2 relative to CCSD for the S_0 surface is most likely due to error cancellations, since CC2 is an approximation of CCSD. CC3 yields results that are consistent with MRCISD+Q. As can be seen from Figure 3, both the S_0 and S_1 CC3 surfaces are close to the MRCISD+Q surfaces. The vertical excitation energies differ from the MRCISD+Q results by 0.01–0.03 eV, which is significantly less than the deviations for CCSD (0.23–0.28 eV) and CC2 (0.06–0.26 eV). The nonparallelity value of CC3 for the CT path is also small, with values of 0.03 eV for S_0 and 0.04 eV for S_1. Based on the nonparallelity value, CC3 performs slightly worse than EOM-SF-CCSD(dT) from Reference 54 for both states, but the nonparallelity values are still the same order of magnitude.

2.2.3 | DIR path

The results for the DIR path for CC2, CCSD and CC3 using 6-31G* are plotted in Figure 4 together with the MRCISD+Q benchmark results of Reference 27, and energies are relative to cis-PSB3. From Figure 4 we see that the CC2, CCSD, and CC3 results for the DIR path show much of the same characteristics as the CT path. CCSD displays too large vertical excitation energies, deviating between 0.22–0.28 eV, and surfaces, which deviate from the shapes of the MRCISD+Q surfaces (see Table 1). The CC2 S_0 surface has approximately the same shape as the MRCISD+Q S_0 surface, whereas larger deviations are found for the S_1 surface. The deviation for CC2 vertical excitation energies relative to MRCISD+Q varies between 0.08 eV (close to DIR coordinate 0.0) to 0.27 eV at DIR coordinates ±0.08 Å·amu$^{1/2}$. CC3 shows small deviations in excitation energies, between 0.02 and 0.04 eV, compared to MRCISD+Q, as well as small nonparallelity values for both

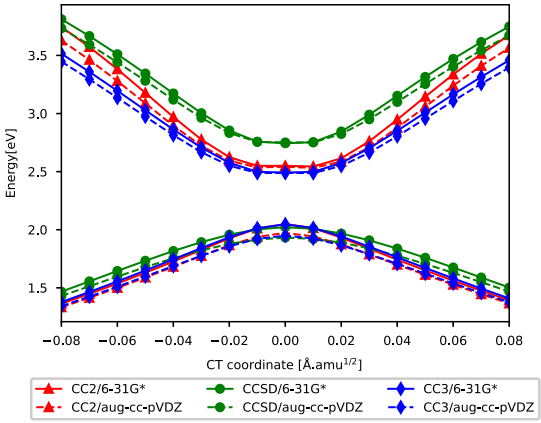

FIGURE 5 S_0 and S_1 energies for the CT path relative to cis-PSB3 computed with CC2, CCSD and CC3 with the 6-31G* and aug-cc-pVDZ basis sets

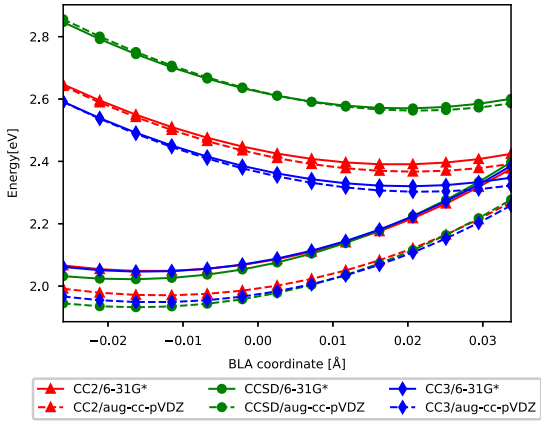

FIGURE 7 S_0 and S_1 energies for the BLA path relative to cis-PSB3 computed with CC2, CCSD and CC3 with the 6-31G* and aug-cc-pVDZ basis sets

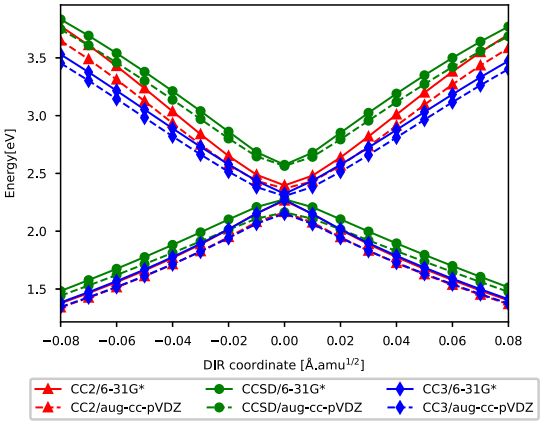

FIGURE 6 S_0 and S_1 energies for the DIR path relative to cis-PSB3 computed with CC2, CCSD and CC3 with the 6-31G* and aug-cc-pVDZ basis sets

TABLE 2 The nonparallelity values for CC2, CCSD, CC3 for the S_0 and S_1 states along the DIR, CT and BLA paths, using aug-cc-pVDZ. All energies are given in eV

Method	DIR S_0	CT S_0	BLA S_0	DIR S_1	CT S_1	BLA S_1
CC2	0.07	0.05	0.03	0.10	0.11	0.03
CCSD	0.07	0.06	0.03	0.08	0.09	0.02
CC3	0.08	0.06	0.03	0.06	0.06	0.02

Note: The nonparallelity values are computed relative to the CC2, CCSD, and CC3 results using 6-31G*.

surfaces. The nonparallelity values of CC3 and EOM-SF-CCSD (dT) from Reference 54 are equal for both surfaces.

2.2.4 | Basis set considerations

Most results presented for PSB3 in the literature are generated using 6-31G*. In the previous sections, we presented the performance of CC2, CCSD, and CC3 relative to MRCISD+Q, all using 6-31G*. We note that basis set effects for MRCISD+Q was explored in Reference 27 for the BLA path using Pople basis sets. In this section, we study the basis set dependence for coupled cluster calculations. We choose to extend from using 6-31G* to the correlation consistent basis set aug-cc-pVDZ rather than staying within the Pople basis set hierarchy, since Dunning's correlation consistent basis sets are seen to converge faster with respect to correlation energy than the Pople basis sets.[67] In Figures 5–7, we present the 6-31G* coupled cluster results from Figures 2–4 together with results generated using aug-cc-pVDZ. The effects of increasing the basis set on the CT path (Figure 5) and the DIR path (Figure 6) are similar. The excited state is shifted to lower energies at large absolute coordinates, whereas the ground state is shifted to lower energies around coordinate 0.0 Å·amu$^{1/2}$. For CC2 the largest shift is found close to 0.0 Å·amu$^{1/2}$, with the ground state being shifted 0.08 eV for the CT path and 0.10 eV for the DIR path. For the BLA path (Figure 7), the change of basis to aug-cc-pVDZ shifts the ground state energy (relative to cis-PSB3) for all three methods by 0.08–0.11 eV. For the excited state, the shift is smaller, as seen from Figure 7. In Table 2, the nonparallelity values of the aug-cc-pVDZ calculations relative to the 6-31G* coupled cluster calculations are presented and demonstrate how much the shape of the surface is affected by choice of basis set. We see from Table 2 that, although the ground state is lowered in energy, the shape of the ground and excited state for the BLA path is least affected by change of basis. The effect is similar for all coupled cluster models used, with the nonparallelity values being 0.03 eV for the ground state and 0.02–

0.03 eV for the excited state. This is also what is seen for MRCISD+Q results in Reference 27. The ground and excited states for CT and DIR paths are more affected by the choice of basis than the BLA path, as seen from Figures 5 and 6 and the nonparallelity values.

2.3 | Exploring vertical excitation energies for ground state equilibrium cis- and trans-azobenzene

In this section we investigate how choice of coupled cluster method and basis set, as well as accounting for solvent, affects computed $S_0 \rightarrow S_1$ and $S_0 \rightarrow S_2$ vertical excitation energies for the azobenzene molecule (see Figure 8). CC2 and CC3 vertical excitation energies are computed for two sets of cis and trans equilibrium structures to also illustrate the effect the molecular geometry has on the vertical excitation energies. The structures used are the 5SA-CASSCF(6,6)/6-31G optimized structures from Reference 32 (denoted cis[a] and trans[a]) and MP2/aug-cc-pVDZ optimized structures (denoted cis[b] and trans[b]). CC2 results are computed using cc-pVDZ, aug-cc-pVDZ, cc-pVTZ and aug-cc-pVTZ, whereas CC3 results are computed using cc-pVDZ and aug-cc-pVDZ. For calculations using aug-cc-pVDZ and aug-cc-pVTZ augmentation is only added to non-hydrogen atoms (using cc-pVDZ and cc-pVTZ on hydrogen). The vertical excitation energies are presented in Table 3 together with experimental UV–VIS results.[60] A discussion in connection to the experimental results is provided at the end of this section.

We first compare CC2 and CC3 results. The absolute difference between CC2 and CC3 $S_0 \rightarrow S_1$ excitation energies when the same basis set is used is 0.00–0.02 eV, whereas the absolute difference between CC2 and CC3 $S_0 \rightarrow S_2$ excitation energies is 0.04–0.10 eV. The slightly larger difference between CC2 and CC3 for the $S_0 \rightarrow S_2$ transition relative to the $S_0 \rightarrow S_1$ transition, may be due to the S_2 state having a lower fraction of single amplitudes and hence CC3 provides a better description. The fraction of singles in the CC3 calculations for cis[a] and trans[a] are listed in Table 4 (together with fraction of singles for the rotational path discussed in Section 2.4). However, the differences between CC2 and CC3 are still small for the structures studied in this section, and this is consistent with available benchmark literature.[62]

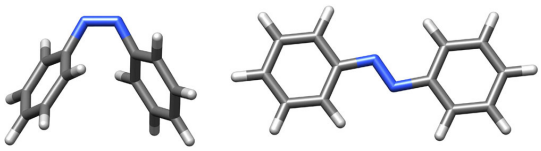

FIGURE 8 Cis (left) and trans (right) isomers of azobenzene

From Table 3, we see that the choice of basis set impacts excitation energies for $S_0 \rightarrow S_1$ differently than $S_0 \rightarrow S_2$ for both sets of structures. Improving the basis set from cc-pVDZ to aug-cc-pVDZ lowers both excitation energies for all structures and methods. For CC2, the $S_0 \rightarrow S_1$ excitation energies are lowered by between 0.04 eV (for trans[a]) and up to 0.06 eV (for cis[b]). Changes in $S_0 \rightarrow S_1$ excitation energies for CC3 when improving from cc-pVDZ to aug-cc-pVDZ are seen to be in the same range (0.02–0.06 eV). In contrast to the $S_0 \rightarrow S_1$ excitation energies, the $S_0 \rightarrow S_2$ excitation energies are seen to change significantly upon improving the basis set from cc-pVDZ to aug-cc-pVDZ. The largest change for CC2 is 0.22 eV (for trans[a] and trans[b]) and the largest change for CC3 is 0.23 eV (trans[b]). Improving the basis set from cc-pVDZ to cc-pVTZ also lowers the excitation energies for all structures. The $S_0 \rightarrow S_1$ excitation is lowered in the range 0.06–0.07 eV, whereas the $S_0 \rightarrow S_2$ excitation energy is lowered in the range 0.13–0.17 eV. Hence, cc-pVTZ lowers the $S_0 \rightarrow S_1$ transition more and $S_0 \rightarrow S_2$ transition less than does aug-cc-pVDZ. We see from the CC2 results that increasing the basis set from aug-cc-pVDZ to aug-cc-pVTZ lowers the $S_0 \rightarrow S_1$ excitation energies by 0.03–0.05 eV, whereas the $S_0 \rightarrow S_2$ excitation energies are lowered by 0.01–0.04 eV. In contrast, increasing the basis set from cc-pVTZ to aug-cc-pVTZ, the $S_0 \rightarrow S_1$ excitation energy is lowered by 0.02–0.04 eV, whereas the $S_0 \rightarrow S_2$ excitation energy is lowered by 0.06–0.09 eV. Hence, $S_0 \rightarrow S_1$ and $S_0 \rightarrow S_2$ excitations seems to have different basis set requirements, with $S_0 \rightarrow S_2$ excitations being more sensitive to the inclusion of diffuse functions.

TABLE 3 CC2 and CC3 vertical excitation energies calculated for cis- and trans-azobenzene structures optimized with CASSCF/6-31G[32] (superscript *a*) and cis- and trans-azobenzene structures optimized with MP2/aug-cc-pVDZ (superscript *b*). All energies are given in eV

		CC2				CC3		
Structure	Transition	DZ	aDZ	TZ	aTZ	DZ	aDZ	UV–VIS[60]
Cis[a]	$S_0 \rightarrow S_1$	2.78	2.73	2.72	2.70	2.77	2.73	2.88
	$S_0 \rightarrow S_2$	4.71	4.52	4.58	4.51	4.62	4.48	4.43
Trans[a]	$S_0 \rightarrow S_1$	2.72	2.68	2.66	2.64	2.72	2.70	2.80
	$S_0 \rightarrow S_2$	4.56	4.34	4.39	4.30	4.61	4.42	3.92
Cis[b]	$S_0 \rightarrow S_1$	3.18	3.12	3.11	3.08	3.17	3.11	2.88
	$S_0 \rightarrow S_2$	4.61	4.45	4.47	4.41	4.55	4.41	4.43
Trans[b]	$S_0 \rightarrow S_1$	2.94	2.89	2.87	2.84	2.94	2.91	2.80
	$S_0 \rightarrow S_2$	4.20	3.98	4.03	3.94	4.30	4.07	3.92

Note: The basis set used for CC2 are cc-pVDZ, cc-pVTZ, aug-cc-pVDZ and aug-cc-pVTZ (cc-pVDZ and aug-cc-pVDZ for CC3). In addition, energies from UV–VIS spectroscopy performed in methanol[60] for cis- and trans-azobenzene are presented.

TABLE 4 Fraction of singles amplitudes, $|\mathbf{R}_1|/|\mathbf{R}|$, for the first and second excited states calculated with CC3/cc-pVDZ for the CASSCF/6-31G azobenzene structures from Zhu et al.[32]

Dihedral angle (degrees)	S_1	S_2
9.5 (cis[a])	0.96	0.92
30	0.96	0.91
50	0.96	0.68
70	0.96	0.38
90	0.95	0.05
110	0.93	0.13
130	0.96	0.44
150	0.95	0.91
180 (trans[a])	0.95	0.93

TABLE 5 CC2 and CC3 vertical excitation energies calculated with cc-pVDZ and aug-cc-pVDZ, where methanol solvent effects are accounted for by the PCM model (see Section 2.1 for PCM specifications). All energies are given in eV

		CC2		CC3
Structure	Transition	DZ	aDZ	DZ
Cis[a]	$S_0 \rightarrow S_1$	2.90	2.85	2.90
	$S_0 \rightarrow S_2$	4.72	4.54	4.63
Trans[a]	$S_0 \rightarrow S_1$	2.78	2.74	2.78
	$S_0 \rightarrow S_2$	4.51	4.28	4.58
Cis[b]	$S_0 \rightarrow S_1$	3.26	3.18	3.25
	$S_0 \rightarrow S_2$	4.55	4.36	4.51
Trans[b]	$S_0 \rightarrow S_1$	2.98	2.93	2.98
	$S_0 \rightarrow S_2$	4.16	3.93	4.26

From the results in Table 3 we can also discuss excitation energies for two different cis- and trans-geometries, that is, 5SA-CASSCF (6,6)/6-31G ground state optimized structures from Reference 32 (cis[a] and trans[a]) and MP2/aug-cc-pVDZ ground state optimized structures (cis[b] and trans[b]). A notable difference between the cis-structures is the CNNC dihedral angles, which is 9.5° for cis[a] and 6.0° for cis[b]. For the trans structures a notable difference is the NNC bond angles, which are 117.9° and 118.2° for trans[a] and 113.3° for both angles for trans[b]. For simplicity, we only discuss CC2/aug-cc-pVTZ results, but we note that the trends observed for CC2/aug-cc-pVTZ is seen for all basis sets and also for CC3 results. For cis[a] and cis[b], the differences in excitation energies are 0.38 eV for $S_0 \rightarrow S_1$ and 0.10 eV for $S_0 \rightarrow S_2$. For the trans[a] and trans[b] structures, the differences in excitation energies are 0.20 eV for $S_0 \rightarrow S_1$ and 0.36 eV for $S_0 \rightarrow S_2$. The differences between the sets of cis- and trans-structures are, therefore, significant. When we proceed to discuss and compare against experimental results we will only consider the MP2/aug-cc-pVDZ structures (cis[b] and trans[b]). We expect MP2/aug-cc-pVDZ optimized structures to be more representative gas phase ground state structures than the ones generated using 5SA-CASSCF(6,6)/6-31G (cis[a] and trans[a]) due to the inclusion of dynamical correlation and the use of a larger basis set.

Here, we compare computed results against absorption maxima from UV–VIS spectroscopy in methanol taken from Reference 60, which are also presented in Table 3. When comparing the computed vertical excitation energies to absorption maxima, we keep in mind that an absorption maximum does not necessarily represent a vertical transition, depending on how fast nuclear rearrangement occurs. We make a rough estimate at what can be considered good agreement with experiment by considering available literature. A study conducted by Bai et al.[68] showed an average deviation of 0.11 ± 0.08 eV between CC2 vertical excitations and absorption maxima computed using a nuclear ensemble approach at the CC2 level for a set of 28 molecules (60 transitions) to account for nuclear rearrangement. Based on this, as well as errors related to the methods used,[62] we discuss the results based on the assumption that a deviation of 0.10 eV, or less, relative to the absorption maximum constitutes reasonable agreement with experiments. However, we further note that the experimental results are obtained in solution and hence both direct solvent effects and geometry will impact the results. The comparison with absorption maxima will be made for the theoretically best results, that is, CC2/aug-cc-pVTZ and CC3/aug-cc-pVDZ calculations for the cis[b] and trans[b] structures. The CC2/aug-cc-pVTZ results show good agreement for both transition of trans[b], overestimating the $S_0 \rightarrow S_1$ transition by 0.04 eV and overestimating the $S_0 \rightarrow S_2$ transition by 0.02 eV. The agreement is worse for the cis[b] $S_0 \rightarrow S_1$ transition, which is overestimated by 0.20 eV, while the cis[b] $S_0 \rightarrow S_2$ transition is in good agreement, being underestimated by 0.02 eV. It should be noted that there are differences observed when comparing our results to the CC2/aug-cc-pVTZ results presented in Reference 39. For example, their $S_0 \rightarrow S_1$ excitation energy for cis-azobenzene overestimates the absorption maximum by 0.12 eV. These difference can be explained by structural differences, as the CNNC dihedral angle of their cis-azobenzene structure deviates by 1.3° compared to the cis[b]. This further illustrates the effect structure has on vertical excitation energies.

Comparing CC3/aug-cc-pVDZ results to absorption maxima, the $S_0 \rightarrow S_1$ transition of cis[b] is overestimated by 0.23 eV, while the $S_0 \rightarrow S_2$ transition of cis[b] is underestimated by 0.02 eV. For trans[b], the $S_0 \rightarrow S_1$ transition is overestimated by 0.11 eV, while the $S_0 \rightarrow S_2$ is overestimated by 0.15 eV. Thus, CC3/aug-cc-pVDZ shows a similar agreement with absorption maxima as CC2/aug-cc-pVTZ for cis[b], while the agreement for trans[b] is worse. The worse agreement for the $S_0 \rightarrow S_1$ of trans[b] appears to be mainly due to the basis set effect, since CC2/aug-cc-pVDZ and CC3/aug-cc-pVDZ results only differ by 0.02 eV. However, the $S_0 \rightarrow S_2$ transition of trans[b] can only be explained by inherent differences between CC2 and CC3, since CC2/aug-cc-pVDZ and CC3/aug-cc-pVDZ results differ by 0.09 eV.

Another factor which must be considered when comparing vertical excitation energies with UV–VIS absorption maxima is the solvent effect. In Table 5, we present results for CC2/cc-pVDZ, CC2/aug-cc-pVDZ, and CC3/cc-pVDZ where methanol solvent effects are incorporated through PCM (see Section 2.1). The effect of including PCM

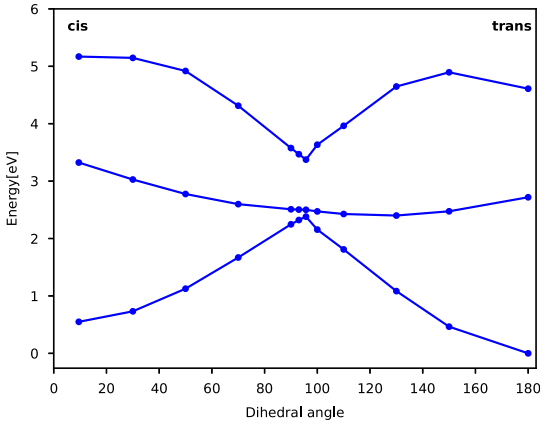

FIGURE 9 Ground state (S_0) and the two lowest excited singlet states (S_1, S_2) calculated with CC3 using cc-pVDZ, for a selected set of azobenzene structures along the rotational path from Reference 32

for these methods and basis set, is that the $S_0 \rightarrow S_1$ energies increase by 0.04–0.08 eV, while the $S_0 \rightarrow S_2$ transitions are lowered by 0.04–0.09 eV. The inclusion of solvent effects through PCM, therefore, does not appreciably improve the agreement between the computed vertical excitation energies and absorption maxima.

2.4 | Coupled cluster results for a rotational path of azobenzene including an excited state with doubly excited character

In this section, we discuss CC3 results for a cis to trans CNNC-rotational path obtained from Zhu et al.[32] The path consists of the cis and trans ground state structures, a twisted minimum energy structure optimized for the second excited state (rot-S2), all optimized with 5SA-CASSCF(6,6)/6-31G, and structures generated by linear interpolation of internal coordinates (LIIC) with respect to these three structures (for details see Reference 32). In Figure 9, CC3 results using the cc-pVDZ basis set are presented for selected structures along the described rotational path for the ground state and the states corresponding to the two lowest singlet vertical excitation energies. The maximum of S_0 corresponds to the aforementioned rot-S2 structure. To describe these states accurately, either cc-pVTZ, aug-cc-pVDZ or a bigger basis set should ideally be used, as can be seen from the basis set study of cis- and trans-azobenzene presented in Section 2.3. However, cc-pVDZ is sufficient to evaluate whether CC3 can describe these states.

Comparing to the results of Reference 32, it is seen that CC3 and 5SA-CASSCF(6,6) are in qualitative agreement for the ground and lowest excited state. However, for S_2, we see from Figure 9 that the CC3 results yield an energy barrier when going from the trans-isomer (180° dihedral angle) toward rot-S2. The barrier is absent in the CASSCF results of Zhu et al. Further, CC3 results for S_2 are reasonably flat in the region from a dihedral angle of 9.5° (cis-isomer) to between 30 and 50° before it rapidly decreases toward the minimum at rot-S2. The 5SA-CASSCF(6,6) results of Reference 32 predict an even decrease of the S_2 energy from cis to rot-S_2. Investigating the amplitudes of the CC3 calculations shows that in the region between 30–50° and 130–150° dihedral angle, the $\pi\pi^*$ state and the $n^2\pi^{*2}$ state switch ordering. At equilibrium structures the second excited state is the $\pi\pi^*$ state and in a broad region (approximately from 50 to 130° dihedral angle) around rot-S2, the $n^2\pi^{*2}$ state is the second excited state. The significant doubly excited character of the second excited state in this region is seen from the relative importance of singles amplitudes in the CC3 calculation. In Table 4 the fraction of singles amplitudes ($|\mathbf{R}_1|/|\mathbf{R}|$) is shown for the first and second excited state throughout the rotational path. As $\pi\pi^*$ and $n^2\pi^{*2}$ switch ordering, Table 4 shows a clear drop in fraction of singles amplitudes for the second excited state. Hence, for a qualitatively correct result for the second excited state in this region, triple excitations must be accounted for. In regions around the ground state equilibrium structures (cis and trans regions), the fraction of singles amplitudes is large, and lower levels of theory such as the CC2 model are sufficient as seen in Section 2.3. These results are consistent with the results in Reference 30 where $n\pi^*$, $\pi\pi^*$ and $n^2\pi^{*2}$ are computed at the multistate CASPT2 level along a rotational path for azobenzene. The results in Reference 30 show a crossing of $\pi\pi^*$ and $n^2\pi^{*2}$ at approximately 35° and 130° dihedral angle.

3 | CONCLUSIONS

The results for PSB3 showed that vertical excitation energies computed with CC2 and CCSD deviated from MRCISD+Q results by 0.02–0.27 eV and 0.13–0.28 eV, respectively. The nonparallelity values also showed that CCSD generally gave deviation in the shape of the surfaces relative to MRCISD+Q, and while the nonparallelity values were small for the CC2 ground states, larger nonparallelity values were found for the CC2 excited states. Contrary, CC3 provided excellent agreement with MRCISD+Q, with vertical excitation energies differing by 0.01–0.04 eV. As a consequence, the nonparallelity values for all ground and excited states were low. The performance of CC3 was similar to EOM-SF-CCSD (dT) presented in Reference 54, where it was seen that triples corrections are necessary. However, it appears that the use of a well-behaved triplet reference state was not required for the PSB3 surfaces explored here.

For equilibrium geometries of cis- and trans-azobenzene, a study of $S_0 \rightarrow S_1$ and $S_0 \rightarrow S_2$ vertical excitation energies computed by coupled cluster models was conducted. It was shown that the molecular geometry had the largest impact, yielding significant variations for vertical excitation energies and shifting the energy of S_1 and S_2 relative to each other. Further, a basis set of sufficient quality was required to describe these states. Both cc-pVTZ and aug-cc-pVDZ were shown to significantly change the vertical excitation energies relative to cc-pVDZ. However, aug-cc-pVDZ was shown to be the better choice of the two, since the inclusion of diffuse functions was found to be important to correctly describe $S_0 \rightarrow S_2$ transitions.

Conversely, the differences between CC2 and CC3 results were shown to be small. By comparing computed vertical excitation energies to UV–VIS absorption maxima, CC2/aug-cc-pVTZ results were found to agree well for all excitations, except the $S_0 \rightarrow S_1$ transition of cis-azobenzene. The computed vertical excitation energies were seen to be sensitive to molecular geometry, and this could in part explain the discrepancy between experimental and computational results. In the case of the CNNC rotational path from Reference 32, we presented the S_0, S_1, and S_2 surfaces using CC3. The CC3 S_0 and S_1 surfaces were in qualitative agreement with the CASSCF surfaces, while the S_2 surface differed significantly since $\pi\pi^*$ and $n^2\pi^{*2}$ switched ordering. This was seen from the onset of significant doubly excited character of the S_2 state in regions of the path. Hence, CC2 and CCSD would not be able to describe the region of the S_2 surface surrounding rot-S2, where $n^2\pi^{*2}$ was the second excited state.

Thus, the article has demonstrated that the coupled cluster models, despite being predominantly single-reference, constitute valuable tools for preliminary and simple black box studies of photoswitchable systems.

ACKNOWLEDGMENTS

The authors acknowledge computing resources through UNINETT Sigma2 the National Infrastructure for High Performance Computing and Data Storage in Norway through project numbers nn9409k. H.K, I-M.H. and R.H.M acknowledge funding from the Research Council of Norway through FRINATEK projects 263110 and 275506. A.C.P and H.K acknowledge funding from the Marie Skłodowska-Curie European Training Network "COSINE – COmputational Spectroscopy In Natural sciences and Engineering," Grant Agreement No. 765739.

DATA AVAILABILITY STATEMENT

xyz-coordinates for cis[b] and trans[b] azobenzene structures of Section 2.3 are given in Supporting Information.

ORCID

Ida-Marie Høyvik https://orcid.org/0000-0002-1239-7776

REFERENCES

[1] M. Quick, A. L. Dobryakov, M. Gerecke, C. Richter, F. Berndt, I. N. Ioffe, A. A. Granovsky, R. Mahrwald, N. P. Ernsting, S. A. Kovalenko, *J. Phys. Chem. B* **2014**, *118*, 8756.
[2] E. Tan, S. Amirjalayer, S. Smolarek, A. Vdovin, F. Zerbetto, W. Buma, *Nat. Commun.* **2015**, *6*, 5860.
[3] A. Nenov et al., *J. Phys. Chem. Lett.* **2018**, *9*, 1534.
[4] M. Robb, F. Bernardi, M. Olivucci, *Pure Appl. Chem.* **1995**, *67*, 783.
[5] B. O. Roos, P. R. Taylor, P. E. Siegbahn, *Chem. Phys.* **1980**, *48*, 157.
[6] R. J. Buenker, S. D. Peyerimhoff, W. Butscher, *Mol. Phys.* **1978**, *35*, 771.
[7] S. R. Langhoff, E. R. Davidson, *Int. J. Quantum Chem.* **1974**, *8*, 61.
[8] W. Duch, G. H. F. Diercksen, *J. Chem. Phys.* **1994**, *101*, 3018.
[9] K. Andersson, P. Malmqvist, B. O. Roos, *J. Chem. Phys.* **1992**, *96*, 1218.
[10] J. Finley, P. Åke Malmqvist, B. O. Roos, L. Serrano-Andrés, *Chem. Phys. Lett.* **1998**, *288*, 299.
[11] G. Ghigo, B. O. Roos, P. Åke Malmqvist, *Chem. Phys. Lett.* **2004**, *396*, 142.
[12] P. Å. Malmqvist, K. Pierloot, A. R. M. Shahi, C. J. Cramer, L. Gagliardi, *J. Chem. Phys.* **2008**, *128*, 204109.
[13] C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, J.-P. Malrieu, *J. Chem. Phys.* **2001**, *114*, 10252.
[14] C. Angeli, S. Borini, M. Cestari, R. Cimiraglia, *J. Chem. Phys.* **2004**, *121*, 4043.
[15] C. Angeli, M. Pastore, R. Cimiraglia, *Theor. Chem. Acc.* **2007**, *117*, 743.
[16] A. A. Granovsky, *J. Chem. Phys.* **2011**, *134*, 214113.
[17] M. Garavelli, T. Vreven, P. Celani, F. Bernardi, M. A. Robb, M. Olivucci, *J. Am. Chem. Soc.* **1998**, *120*, 1285.
[18] T. Ishikawa, T. Noro, T. Shoda, *J. Chem. Phys.* **2001**, *115*, 7503.
[19] L. De Vico, C. S. Page, M. Garavelli, F. Bernardi, R. Basosi, M. Olivucci, *J. Am. Chem. Soc.* **2002**, *124*, 4124.
[20] C. S. Page, M. Olivucci, *J. Comput. Chem.* **2003**, *24*, 298.
[21] E. Wei-Guang Diau, *J. Phys. Chem. A* **2004**, *108*, 950.
[22] A. Cembran, F. Bernardi, M. Garavelli, L. Gagliardi, G. Orlandi, *J. Am. Chem. Soc.* **2004**, *126*, 3234.
[23] L. Gagliardi, G. Orlandi, F. Bernardi, A. Cembran, M. Garavelli, *Theor. Chem. Acc.* **2004**, *111*, 363.
[24] M. Wanko, M. Hoffmann, P. Strodel, A. Koslowski, W. Thiel, F. Neese, T. Frauenheim, M. Elstner, *J. Phys. Chem. B* **2005**, *109*, 3606.
[25] I. Conti, M. Garavelli, G. Orlandi, *J. Am. Chem. Soc.* **2008**, *130*, 5216.
[26] G. Tiberio, L. Muccioli, R. Berardi, C. Zannoni, *ChemPhysChem* **2010**, *11*, 1018.
[27] S. Gozem, M. Huntress, I. Schapiro, R. Lindh, A. A. Granovsky, C. Angeli, M. Olivucci, *J. Chem. Theory Comput.* **2012**, *8*, 4069.
[28] Y. Harabuchi, M. Ishii, A. Nakayama, T. Noro, T. Taketsugu, *J. Chem. Phys.* **2013**, *138*, 064305.
[29] E. Walczak, B. Szefczyk, T. Andruniów, *J. Chem. Theory Comput.* **2013**, *9*, 4915.
[30] J. Casellas, M. J. Bearpark, M. Reguero, *ChemPhysChem* **2016**, *17*, 3068.
[31] L. Yu, C. Xu, C. Zhu, *Phys. Chem. Chem. Phys.* **2015**, *17*, 17646.
[32] C. Xu, L. Yu, F. L. Gu, C. Zhu, *Phys. Chem. Chem. Phys.* **2018**, *20*, 23885.
[33] F. Aleotti, L. Soprani, A. Nenov, R. Berardi, A. Arcioni, C. Zannoni, M. Garavelli, *J. Chem. Theory Comput.* **2019**, *15*, 6813.
[34] E. Runge, E. K. U. Gross, *Phys. Rev. Lett.* **1984**, *52*, 997.
[35] C. Møller, M. S. Plesset, *Phys. Rev.* **1934**, *46*, 618.
[36] P.-O. Åstrand, P. S. Ramanujam, S. Hvilsted, K. L. Bak, S. P. A. Sauer, *J. Am. Chem. Soc.* **2000**, *122*, 3482.
[37] N. Kurita, T. Ikegami, Y. Ishikawa, *Chem. Phys. Lett.* **2002**, *360*, 349.
[38] C. Hättig, K. Hald, *Phys. Chem. Chem. Phys.* **2002**, *4*, 2111.
[39] H. Fliegl, A. Köhn, C. Hättig, R. Ahlrichs, *J. Am. Chem. Soc.* **2003**, *125*, 9821.
[40] C. R. Crecca, A. E. Roitberg, *J. Phys. Chem. A* **2006**, *110*, 8188.
[41] R. Send, D. Sundholm, *J. Phys. Chem. A* **2007**, *111*, 27.
[42] R. Send, D. Sundholm, *J. Phys. Chem. A* **2007**, *111*, 8766.
[43] S. Yuan, Y. Dou, W. Wu, Y. Hu, J. Zhao, *J. Phys. Chem. A* **2008**, *112*, 13326.
[44] R. Send, D. Sundholm, M. P. Johansson, F. Pawłowski, *J. Chem. Theory Comput.* **2009**, *5*, 2401.
[45] H. Koch, O. Christiansen, P. Jørgensen, A. Merás, T. Helgaker, *J. Chem. Phys.* **1997**, *106*, 1808.
[46] P. Piecuch, M. Włoch, *J. Chem. Phys.* **2005**, *123*, 224105.
[47] P. U. Manohar, A. I. Krylov, *J. Chem. Phys* **2008**, *129*, 194105.
[48] P. U. Manohar, J. F. Stanton, A. I. Krylov, *J. Chem. Phys* **2009**, *131*, 114112.
[49] D. A. Matthews, J. F. Stanton, *J. Chem. Phys.* **2016**, *145*, 124102.
[50] J. Noga, R. J. Bartlett, *J. Chem. Phys.* **1987**, *86*, 7041.
[51] G. Purvis III., R. Bartlett, *J. Chem. Phys.* **1982**, *76*, 1910.
[52] O. Christiansen, H. Koch, P. Jørgensen, *Chem. Phys. Lett.* **1995**, *243*, 409.
[53] E. F. Kjønstad, R. H. Myhre, T. J. Martínez, H. Koch, *J. Chem. Phys* **2017**, *147*, 164105.
[54] S. Gozem, A. I. Krylov, M. Olivucci, *J. Chem. Theory Comput.* **2013**, *9*, 284.

[55] M. Huix-Rotllant, M. Filatov, S. Gozem, I. Schapiro, M. Olivucci, N. Ferré, *J. Chem. Theory Comput.* **2013**, *9*, 3917.
[56] S. Gozem, F. Melaccio, R. Lindh, A. I. Krylov, A. A. Granovsky, C. Angeli, M. Olivucci, *J. Chem. Theory Comput.* **2013**, *9*, 4495.
[57] R. Ditchfield, W. J. Hehre, J. A. Pople, *J. Chem. Phys.* **1971**, *54*, 724.
[58] B. Helmich-Paris, *J. Chem. Theory Comput.* **2019**, *15*, 4170.
[59] L. Wang, W. Xu, C. Yi, X. Wang, *J. Mol. Graphics Modell.* **2009**, *27*, 792.
[60] L. Vetráková, V. Ladányi, J. Al Anshori, P. Dvořák, J. Wirz, D. Heger, *Photochem. Photobiol. Sci.* **2017**, *16*, 1749.
[61] S. D. Folkestad, E. F. Kjønstad, R. H. Myhre, J. H. Andersen, A. Balbi, S. Coriani, T. Giovannini, L. Goletto, T. S. Haugland, A. Hutcheson, et al., *J. Chem. Phys.* **2020**, *152*, 184103.
[62] M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer, W. Thiel, *J. Chem. Phys.* **2008**, *128*, 134110.
[63] A. C. Paul, R. H. Myhre, H. Koch, *J. Chem. Theory Comput.* **2021**, *17*, 117.
[64] M. R. Silva-Junior, S. P. Sauer, M. Schreiber, W. Thiel, *Mol. Phys.* **2010**, *108*, 453.
[65] S. D. Folkestad, E. F. Kjønstad, H. Koch, *J. Chem. Phys.* **2019**, *150*, 194112.
[66] E. Cancès, B. Mennucci, *J. Math. Chem.* **1998**, *23*, 309.
[67] T. Helgaker, P. Jørgensen, J. Olsen, *Molecular Electronic Structure Theory*, 1st ed., Vol. 0, Wiley, Chichester **2000**.
[68] S. Bai, R. Mansour, L. Stojanović, J. Toldo, M. Barbatti, *J. Mol. Model.* **2020**, *26*, 107.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: A. Hutcheson, A. C. Paul, R. H. Myhre, H. Koch, I.-M. Høyvik, *J Comput Chem* **2021**, *42*(20), 1419. https://doi.org/10.1002/jcc.26553

Paper B

Convergence of the electronic density for a target region in cluster models of a NH_3 molecular crystal

A. Hutcheson, I-M. Høyvik,

J. Math. Chem. **2022**;
doi: 10.1007/s10910-022-01351-w

Journal of Mathematical Chemistry
https://doi.org/10.1007/s10910-022-01351-w

ORIGINAL PAPER

Convergence of the electronic density for a target region in cluster models of a NH_3 molecular crystal

Anders Hutcheson[1] · **Ida-Marie Høyvik**[1]

Received: 21 December 2021 / Accepted: 18 March 2022

Abstract

In this paper we illustrate the advantage of addressing size-intensive properties of target regions without first converging the ground-state energy of that region. We use local occupied and virtual orbitals to separate the orbital space of NH_3 clusters into an orbital space for the target region (a central NH_3 molecule) and for the remaining cluster. Convergence characteristics of the Hartree–Fock (HF) energy and, indirectly, the electronic density of the target region are shown. The calculations illustrate that although the energy of the target region will not converge with cluster size, the electronic density will. The convergence of the electronic density of the target region is subsequently exploited to obtain HF dipole moments and CC2-in-HF vertical excitation energies. For these properties convergence is seen upon the inclusion of approximately three shells beyond the target region. This shows that local size-intensive properties of a target region can be investigated without converging the energy. We further show that a minimal basis description of the outer shells are sufficient to capture the correct interaction with the target region. The possibility of computing size-intensive properties for a target region using a converged electronic density, without requiring convergence in the energy itself, is currently an underexploited feature.

Keywords Electronic density · Hartree–Fock · Coupled cluster · Cluster model · Molecular crystal · Local properties · Local orbitals

✉ Ida-Marie Høyvik
ida-marie.hoyvik@ntnu.no

Anders Hutcheson
anders.hutcheson@ntnu.com

[1] Department of Chemistry, The Norwegian University of Science and Technology, Høgskoleringen 5, 7491 Trondheim, Norway

Published online: 29 April 2022

1 Introduction

In 1996 Kohn [1] introduced the term "nearsightedness" of electrons in many-atom systems. This feature describes that local electronic properties, such as the electronic density, depend significantly on the effective external potential only at nearby points. The convergence of the electronic density for a local region has also been discussed by others [2–5]. The feature has been exploited to develop numerous linear-scaling electronic structure models and embedding schemes. However, unlike the electronic density, the ground-state electronic energy is not nearsighted. Hence, for systems such as molecular crystals, the energy of a target region requires the infinite surroundings to be taken into account, thus motivating the use of periodic boundary conditions. A target region may be a unit cell, single molecule or group of molecules in the crystal. Since the constituents of molecular crystals are molecules, wave function models originally developed for single molecules have therefore been extended to periodic codes. Notable examples are the plane wave periodic MP2 method [6] implemented in VASP, the periodic divide-expand-consolidate MP2 method developed by Pedersen et al. [7], the periodic MP2 and CCSD developed by McClain et al. [8], and the local MP2 method [9] available in the Cryscor program.

To achieve linear-scaling wavefunctions for molecular systems, local molecular orbital (MO) spaces are often used. Explicit localization of occupied MOs has been a popular topic for many decades, with seminal contributions such as the widely used Edmiston-Ruedenberg [10], Foster-Boys [11], and Pipek-Mezey [12] localization functions. In addition to explicit localization, several approaches for generating local virtual spaces exist such as projected atomic orbitals (PAOs) [13], pair-natural orbitals [14–16] and correlated natural transition orbitals [17]. However, for the infinite molecular crystals orbital space locality faces two challenges (i) orthogonalization tails and (ii) near-linear dependencies.

Orthogonalization tails compromise compactness of the description, since for a given region in space, local MOs outside the region will be required to have components inside the region due to the orthogonality requirement. Using high-quality atomic orbital (AO) basis sets, there is a large number of MOs centered far outside the given region which must have components inside the region. Parts of the orbital space in a given region is therefore spanned by MOs centered outside the region. These components may cause the correlation energy to converge slowly in local correlation approaches relative to when non-orthogonal orbitals such as the PAOs are used. This is seen in the results presented by Werner and collaborators [18] and is explicitly shown by Hansen et al. [19]. The results of Hansen et al. indicate that some issues concerning orthogonalization tails may be circumvented by using PAOs rather than local virtual MOs (or Wannier orbitals).

With respect to near-linear dependencies, AOs are per construction designed to describe atoms in free space, and for large or dense systems the concerted effect of AOs on different atomic centers greatly enhances near-linear dependencies.

Near-linear dependencies will not appreciably affect the locality of the occupied space, but the virtual space will be severely affected [20]. Dealing with near-linear dependencies in a periodic framework may compromise the basis set quality, since removing AOs from the unit cell will remove AOs from all unit cells. While near-linear

dependencies can be avoided by using plane wave basis sets, there are other problems with this approach, such as the need for a high energy cut-off to achieve accurate results.

When targeting local size-intensive properties of molecular crystals, the nearsightedness of electrons may allow the use of cluster models. Cluster models enables different basis sets to be used in different regions of the cluster. Hence one may use smaller basis sets outside a target region. The problems of orthogonalization tails and near-linear dependencies may therefore be alleviated by using cluster models. Further, cluster models allows for any molecular wave function based scheme to be used for its description. Due to the nearsightedness of the electronic density, local size-intensive properties in a target region may be converged with cluster size. The traditional approach would be to first compute the converged energy, however, the energy of the target region will not be converged for reasonably sized clusters. Therefore, if one does not aim to converge the energy with cluster size, any existing fragment or orbital based correlated wave function models [21–42] can be used to compute size-intensive properties of the target region.

In this paper, we illustrate that one may obtain converged size-intensive properties of a target region without converging the energy of the target region. Cluster models of various sizes gives us the opportunity to show how electronic properties converge with increasing cluster size. Furthermore, the cluster models allows us to explore how choice of basis set outside the target region affects the computed properties. One possibility being to choose a high-quality basis set in and around the region of interest, while low-quality basis sets (such as minimal basis) may be used further out in the cluster model. A minimal basis does not exhibit flexibility to describe accurate molecular properties, but it may adequately describe long-range effects between the electronic density of the targeted region and the electronic density far away. The aim of this paper is therefore not to obtain accurately computed electronic properties, but rather to illustrate fundamental concepts related to the convergent nature of the electronic density and how it may be exploited.

The paper is organized as follows. In Sect. 2 we provide a theoretical background on how to partition the energy in terms of a target region and the remainder of the cluster, as well as an argument from an optimization vantage point for why the electronic density should converge with cluster size. In Sect. 3 we present computational details and describe the cluster models of an NH_3 molecular crystal, and in Sect. 4 we present numerical illustrations using these cluster models. A summary and concluding remarks is given in Sect. 5.

2 Theoretical background

In this section we present equations and background for the partitioning of the Hartree–Fock electronic energy in terms of a partitioning of a fully optimized Hartree–Fock density matrix for the cluster. The density and energy partitioning used was introduced by Høyvik et al. [43] for multilevel Hartree–Fock. We note that the partitioning is used to illustrate the convergence behaviour of the energy. The partitioning used here is not a requirement for computing size-intensive properties of a target-region, as one may

use any existing fragmentation or orbital space partitioning approach for this. The only requirement is that the approach allows for the effective interaction of the target region with the rest of the cluster. After presenting the partioning of the energy, we discuss why it is reasonable to assume that the electronic density of a target region will converge with cluster size by analyzing the optimization procedure from a local perspective rather than a canonical (diagonalization based) perspective.

2.1 Energy partitioning

We consider a cluster model of a molecular crystal where the total electronic density for the cluster is given by $\mathbf{D} = \mathbf{D}_t + \mathbf{D}_r$. $\mathbf{D}_t$ is the density of the target region, and $\mathbf{D}_r$ is the remainder of the density. The total density $\mathbf{D}$ is the Hartree–Fock density matrix for the full cluster. All three densities, $\mathbf{D}$, $\mathbf{D}_t$ and $\mathbf{D}_r$ satisfy the trace, symmetry and idempotence criteria of density matrices representing Slater determinants, i.e. $\mathbf{D}_t$ and $\mathbf{D}_r$ can be viewed as constructed from separate subsets of orthogonal MOs describing the full density matrix $\mathbf{D}$. The electronic energy (excluding nuclear repulsion) for the target region described by $\mathbf{D}_t$, interacting with the rest of the cluster described by $\mathbf{D}_r$ is given by,

$$\begin{aligned} E_t &= \mathrm{Tr}[\mathbf{h}\mathbf{D}_t] + \frac{1}{4}\mathrm{Tr}[\mathbf{D}_t\mathbf{G}(\mathbf{D}_t)] + \frac{1}{2}\mathrm{Tr}[\mathbf{D}_t\mathbf{G}(\mathbf{D}_r)] \\ &\equiv E_t^{\text{1-el}} + E_t^{\text{2-el}} + E_t^{\text{int}}, \end{aligned} \tag{1}$$

where we have defined $E_t^{\text{1-el}}$, $E_t^{\text{2-el}}$ and E_t^{int} as the one-electron, two-electron and interaction contributions, respectively. Two-electron and interaction terms are defined through the $\mathbf{G}$ matrix which in the MO basis is defined through elements

$$G(\mathbf{M})_{ij} = \sum_{kl}(2g_{ijkl} - g_{ilkj})M_{kl}, \tag{2}$$

where we have introduced two-electron integrals in the Mulliken notation

$$g_{ijkl} = \int\int \phi_i^*(\mathbf{r}_1)\phi_j(\mathbf{r}_1)\frac{1}{r_{12}}\phi_k^*(\mathbf{r}_2)\phi_l(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2. \tag{3}$$

The electronic part of the Hartree–Fock energy (nuclear repulsion excluded) of the full cluster is given by

$$E = E_t + \mathrm{Tr}[\mathbf{h}\mathbf{D}_r] + \frac{1}{4}\mathrm{Tr}[\mathbf{D}_r\mathbf{G}(\mathbf{D}_r)], \tag{4}$$

i.e., the energy for the target region, E_t, plus the contributions which only depend on $\mathbf{D}_r$. Note that all two-electron interactions between the target region and the rest of the cluster is included in E_t.

As is well-known, the energy of a target region in an infinite (or similarly, large) system, is not a local quantity even if the target density matrix $\mathbf{D}_t$ is localized in space.

The reason for this is long-range interactions between the local target density and nuclei and between the local target density and the electron density of the rest of the system. Hence, the energy of the target region E_t will not converge appreciably for finite cluster sizes. This is well recognized and part of the motivation behind periodic treatments of infinite systems. However, an important point here is that the lack of convergence of the energy with system size does not necessarily reflect a lack of convergence for the electronic density in the target region. In the next section we present an optimization viewpoint of why the electronic density of a target region in a finite cluster of a molecular crystal converges with cluster size.

2.2 Density convergence

For the optimization of a Hartree–Fock state, diagonalization (Roothaan–Hall) based schemes are widely used. However, the diagonalization based schemes impose extra restrictions on the MOs to generate a diagonal Fock matrix (through canonical orbitals) whereas the optimization condition only requires a block-diagonal Fock matrix. To investigate the effect on the target electronic density by the increasing cluster size, it is instructive to consider an optimization based scheme which does not enforce a canonical basis. Hence, we look at a scheme based on an exponential parametrization of MO coefficients [44, 45], where in each iteration new orbitals are generated by a unitary transformation,

$$\tilde{\mathbf{C}} = \mathbf{C}\exp(\boldsymbol{\kappa}). \tag{5}$$

The anti-symmetric parameter matrix $\boldsymbol{\kappa}$ contains only non-redundant parameters, i.e., only the occupied-virtual blocks of $\boldsymbol{\kappa}$ are non-zero for a closed-shell state. A quadratic model of the total energy for the cluster can then be constructed, giving a linear-equation Newton based optimization scheme where in each Hartree–Fock iteration, we solve the linear equation

$$\mathbf{H}\boldsymbol{\kappa} = -\mathbf{G}. \tag{6}$$

Therefore, the solution in each iteration formally is given by,

$$\boldsymbol{\kappa} = -\mathbf{H}^{-1}\mathbf{G}, \tag{7}$$

although the equations are usually solved in an reduced space (iterative) manner. $\mathbf{H}$ and $\mathbf{G}$ are the electronic Hessian and gradient, respectively. In the (non-canonical) orbital basis the electronic Hessian is [46]

$$H_{ai,bj} = 4(\delta_{ij}F_{ab} - \delta_{ab}F_{ij} + 4g_{aibj} - g_{abij} - g_{ajib}), \tag{8}$$

where F_{pq} are elements of the MO Fock matrix, and the electronic gradient is [46]

$$G_{ai} = -4F_{ai}. \tag{9}$$

Exploiting that the electronic Hessian is diagonally dominant, we may write the orbital rotation parameters approximately as

$$\kappa_{ai} \approx -H_{ai,ai}^{-1} G_{ai} = 4 H_{ai,ai}^{-1} F_{ai}. \tag{10}$$

The diagonal Hessian elements are $H_{ai,ai} = 4(F_{aa} - F_{ii} + 4g_{aiai} - g_{aaii} - g_{aiia})$ and their magnitude is therefore dominated by the difference $F_{aa} - F_{ii}$. Hence, $H_{ai,ai}^{-1}$ will not be a divergent term for systems with non-vanishing HOMO–LUMO gaps, since $F_{aa} - F_{ii}$ will for such systems always be of a reasonable size. Close to convergence of the Hartree–Fock state, the occupied-virtual Fock matrix element F_{ai} will be small since these are gradient terms, but far from convergence (in the initial Hartree–Fock iterations), F_{ai} will generally have a significant size.

For illustrative purposes, we now assume that in each Hartree–Fock iteration we have an orbital basis of local occupied and virtual MOs. For molecular systems (including cluster models of molecular crystals) with non-vanishing HOMO–LUMO gaps we know that such a basis exists, and in principle we can for each iteration use redundant orbital rotations to generate such a basis. In each iteration the resulting new occupied orbital i can be written as

$$\tilde{C}_{\mu i} = C_{\mu i} + \sum_a C_{\mu a} \kappa_{ai} + \mathcal{O}(\kappa^2) \tag{11}$$

Hence, for each iteration in the energy optimization, an orbital i in the target region will get an amount of virtual MO a mixed in, weighted (to first order) by the magnitude of κ_{ai}. Hence, if an occupied MO i (and hence, the electronic density) is to be changed by a virtual orbital a centered far away, κ_{ai} must be of a significant size. If we look at Eq. (10), we see that the size of κ_{ai} will be determined by the size of F_{ai}. For local orbitals i and a centered far away from each other F_{ai} exhibits a rapid decay [47] and therefore the MOs of the target region (and thus the density of the target region) will at some point be unchanged when increasing the cluster size. The use of local MOs in this analysis is convenient from a conceptual point of view since local MOs also enable a partitioning into a target density which is local. However, the density is invariant with respect to redundant rotations, and the convergence properties of a local part of the density with cluster size is indifferent to choice of basis.

3 Methodology

In this section we present computational details, description of the NH_3 molecular clusters used and how the active excitation space for CC2-in-HF calculations are chosen.

3.1 Computational details

LSDalton [48] was used for the Hartree–Fock calculations and subsequent localization of the occupied orbitals. For the Hartree–Fock calculation, the screening threshold for

the integral evaluation was set to 10^{-10} and convergence threshold was set to 10^{-8}. The localization of the occupied orbitals is done by using the second power of the second central moment localization function [49]. CC2-in-HF excitation energies were computed in eT [50], by using the local orbitals obtained from LSDalton. For CC2-in-HF the CC2 wave function is constructed in a subset of the full orbital space whereas interaction with the inactive space (frozen Hartree–Fock orbitals) enters through the Fock matrix. For details see Ref. [50]. The frozen core approximation was used for all CC2-in-HF calculations. The decomposition of the electron repulsion integrals [51] threshold was set to 10^{-5}. The energy and residual thresholds for the coupled cluster ground state were set to 10^{-7}, and the residual threshold for the coupled cluster excited states were set to 10^{-3}.

3.2 Description of NH_3 molecular clusters

In this paper we look at cluster models of crystalline NH_3 (crystal structure obtained from Ref. [52]). Clusters comprised of 13, 40, 143, 324, 579 and 953 molecules are used and the models represent a target region (central NH_3 molecule) with a certain number of shells. The clusters are depicted in Fig. 1, as well as how the cluster is

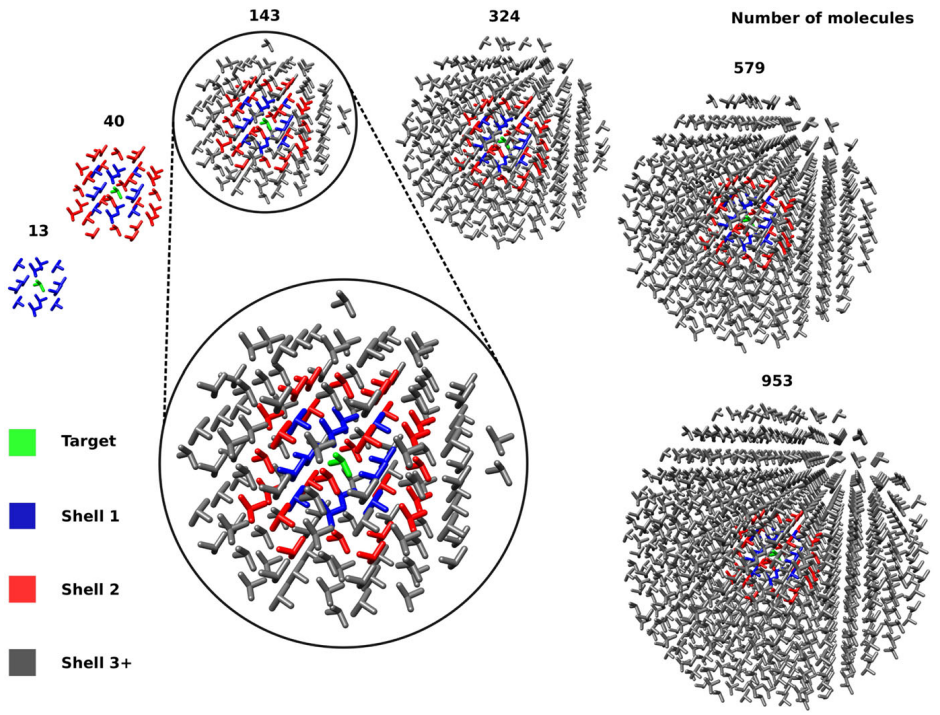

Fig. 1 The ammonia clusters used, with the regions where different basis set can be used colored. The target region (green) is the central NH_3 molecule, shell 1 (blue) is the first shell of 12 molecules, shell 2 (red) is the second shell of 27 molecules, while shell 3+ (grey) is any shells beyond the second shell. In the target region aug-cc-pVTZ is always used, the basis sets of shell 1 and shell 2 are variable, and STO-3G is always used for shell 3+ (Color figure online)

divided into regions where different basis sets may be used. In the target region we always use the aug-cc-pVTZ basis set [53], while for shell 1 and shell 2 we vary the choice of basis set. For shells beyond shell 2, hereby termed shell 3+, STO-3G [54] is always used. When augmented basis sets are used, augmentation is only included for nitrogen atoms, and the non-augmented version of the basis set is used on hydrogen atoms.

3.3 Orbital space partitioning

We use local occupied MOs generated by a trust-region minimization [55] of the second power of the second central moment [49] localization function. The occupied orbital space is partitioned based on the centers, defined by the expectation value of the orbital position vector $(\langle\phi_i \mid \hat{x} \mid \phi_i\rangle, \langle\phi_i \mid \hat{y} \mid \phi_i\rangle, \langle\phi_i \mid \hat{z} \mid \phi_i\rangle)$, of the local MOs. An occupied MO ϕ_i belongs to the region (target, shell 1, etc.) to which it is situated closest to, as defined by the l_2 norm of the difference between the orbital position vector and the atomic positions. For the virtual space, we use the projected atomic orbitals (PAOs) generated for the virtual active region of the cluster (see Sect. 3.4).

3.4 The excitation space for CC2-in-HF calculations

The vertical $S_0 \rightarrow S_1$ excitation energies presented here are computed using a fixed occupied CC2 active space, where only occupied orbitals centered in the target region is included. To be able to investigate the convergence of the excitation energies with an increasing virtual space, in addition to cluster size, we define two different CC2 active virtual regions; target + shell 1 and target + shell 1-2. The inclusion of a larger occupied space would allow for relaxation effects in the occupied space yielding lower excitation energies. However, the purpose of this study is to explore the effect of cluster size and basis set for the excitation energies. We note that the target region approach is only useful for size-intensive (local) properties.

4 Numerical illustrations using NH_3 clusters

In this section we use the NH_3 clusters described in Sect. 3.2 to illustrate effects of cluster size on the energy and the electronic density in the target region (the central NH_3 molecule). The convergence of the electronic density will indirectly be illustrated through presenting results for the Hartree–Fock electric dipole moment and CC2 $S_0 \rightarrow S_1$ vertical excitation energies for the target region. Further, we will explore how these properties depends on the quality of the basis set outside the target region (shell 1 and shell 2).

4.1 Effect of cluster size

4.1.1 Target region energy versus target region electronic density

In this section we present Hartree–Fock calculations on clusters comprised of 13, 40, 143, 324, 579 and 953 NH_3 molecules where an orbital space partitioning (see Sect. 3.3) of the Hartree–Fock orbital space is used to divide the system into a density for the target region and the remainder density. The basis sets used are aug-cc-pVTZ in the target region (central NH_3 molecule), aug-cc-pVDZ for shell 1, STO-3G for shell 2 and STO-3G for shell 3+, see Sect. 3.2 for a description of the clusters.

We present the differences in energy contributions (see Sect. 2.1) to the target region energy for increasing cluster size. I.e., we present results for

$$\Delta E(m,n) = E^m - E^n \tag{12}$$

where E^m is an energy ($E_t^{1\text{-el}}$, $E_t^{2\text{-el}}$, E_t^{int} or E_t from Eq. (1)) for a cluster containing m molecules and E^n is an energy for a cluster containing n molecules. The results are presented in Table 1. We first consider results for the one-electron contribution, $E_t^{1\text{-el}}$, to the target region. We see that the one-electron energy contribution to the target region energy is increasing in magnitude (but is negative), due to the long-range interaction between the electrons in the target region and all nuclei of the cluster. For the two-electron interaction between the electrons in the target region and the electrons outside the target region, E_t^{int}, we see the same long-range effects, except that here the energy contributions are increasing in magnitude and are of a positive sign. Hence, these energy contributions do not converge with cluster size. However, we see that $\Delta E_t^{1\text{-el}}$ and ΔE_t^{int} have similar values, except with opposite signs. Hence, they nearly cancel each other out. This is seen from the total energy of the target region E_t, where we see that ΔE_t is small, but not converged. On the other hand, we see from Table 1 that $E_t^{2\text{-el}}$ is converging with cluster size. We see that the sign of $\Delta E(m,n)$ for $\Delta E_t^{2\text{-el}}$ switches e.g., between $\Delta E(40, 13)$ and $\Delta E(143, 40)$ and that they are of similar order of magnitude. The same is seen for $\Delta E_t^{2\text{-el}}$ of $\Delta E(324, 142)$ and $\Delta E(579, 324)$. Furthermore, the absolute value $\Delta E(579, 324)$ for $\Delta E_t^{2\text{-el}}$ is seen to be larger than that of $\Delta E(324, 142)$ (0.00037 a.u. versus 0.00019 a.u.). However, considering the increase in number of molecules, the energy differences in terms of change per molecule is of the same size. Looking at $\Delta E_t^{2\text{-el}}$ for the cluster sizes of 953 and 579 molecules, we see that the difference is down to 0.000008 a.u. This implies that, beyond a given cluster size, $\mathbf{D}_t$ is not significantly affected by extending the cluster further. This is a numerical illustration of the theoretical discussion in Sect. 2.2.

4.1.2 Effect of cluster size on dipole moment

In this section we investigate how the cluster size affects the Hartree–Fock electric dipole moment of the target region (central NH_3 molecule). The dipole moment of the target region is computed using the trace of $\mathbf{D}_t$ and the dipole operator. The aug-cc-pVTZ basis set is used in the target region, aug-cc-pVDZ for shell 1, STO-3G for

Table 1 Energy differences Eq. (12) of the energy contributions, $E_t^{1\text{-el}}$, $E_t^{2\text{-el}}$ and E_t^{int} from Eq. (1) as well as their sum, E_t, for ammonia clusters of different sizes

Cluster sizes	$\Delta E_t^{1\text{-el}}$	$\Delta E_t^{2\text{-el}}$	ΔE_t^{int}	ΔE_t
$\Delta E(40, 13)$	− 243.29175	− 0.00252	243.73391	0.43964
$\Delta E(143, 40)$	− 632.16812	0.00242	632.30466	0.13896
$\Delta E(324, 142)$	− 798.07264	− 0.00019	798.07633	0.00349
$\Delta E(579, 324)$	− 896.23328	0.00037	896.20599	− 0.02692
$\Delta E(953, 579)$	− 1102.70560	0.000008	1102.70167	− 0.00392

The clusters are described using aug-cc-pVTZ in the target region (central NH_3 molecule), aug-cc-pVDZ for shell 1, STO-3G for shell 2 and STO-3G for shell 3+, see Sect. 3.2 for a description of the clusters. All values are in a.u

Table 2 Hartree–Fock electric dipole moments (given in Debye) for the target region (central NH_3 molecule), computed for cluster sizes from 13 to 953 ammonia molecules

Cluster size	Dipole moment [D]
13	2.25
40	2.06
143	2.14
324	2.12
579	2.14
953	2.14

The clusters are described using aug-cc-pVTZ in the target region (central NH_3 molecule), aug-cc-pVDZ for shell 1, STO-3G for shell 2 and STO-3G for shell 3+, see Sect. 3.2 for a description of the clusters

shell 2 and STO-3G for shell 3+. The computed dipole moments are given in Table 2. Increasing the cluster size from 13 molecules to 40 molecules reduces the dipole moment by 0.19 D, and increasing the cluster from 40 to 143 molecules increases it by 0.08 D. Increasing the cluster size from 143 to 324 molecules reduces the dipole moment by 0.02 D, and further increases of cluster size gives the same value for the dipole moment as the cluster with 143 molecules. Considering that the mean absolute error in Hartree–Fock dipole moments for molecules is found to be 0.16 D [56], these variations are negligible. Hence, the electric dipole moment of the target region is converged at 143 molecules (target region + 3 shells). This convergence is expected as the results in Table 1 indicate a converged target electronic density matrix.

4.1.3 Effect of cluster size on CC2-in-HF vertical excitation energies

In Table 3 we present CC2-in-HF excitation energies for NH_3 clusters containing 143, 324 and 579 molecules, with the occupied active space restricted to the target region (central NH_3 molecule) and the virtual active restricted to the target region + shell 1 (see Sect. 3.4). The basis set for the target region is aug-cc-pVTZ, whereas the basis set for shell 2 and beyond is STO-3G. For shell 1 we present results using both STO-3G and aug-cc-pVDZ, to see whether the choice of basis in the active virtual region yield different convergence characteristics with cluster size.

Table 3 CC2-in-HF $S_0 \rightarrow S_1$ excitation energies (eV) obtained for clusters of 143, 324 and 579 molecules

Basis (shell 1)	143	324	579
STO-3G	9.26	9.24	9.26
aDZ	9.03	9.01	9.03

All calculations are performed with the target region as the active occupied region and the virtual active region set as the target region + shell 1. The basis sets used for the target region is aug-cc-pVTZ and the basis set for shell 2 and beyond is STO-3G

We first consider results where STO-3G was used for shell 1. From Table 3 we see that increasing the cluster size from 143 to 324 molecules the excitation energy is decreased from 9.26 to 9.24 eV. Increasing the cluster size further to 579 molecules changes it back up to 9.26 eV. Using aug-cc-pVDZ for shell 1 we see that the excitation energy is lowered as expected, since more flexibility is added in the active virtual region. However, we see the same convergence behavior as for the calculations where only STO-3G was employed beyond the target region. Hence, for cluster sizes beyond 143 molecules we see that excitation energies oscillates somewhat, but the change is of only 0.02 eV. Considering the intrinsic errors in the CC2 model, the effect on the excitation energy by increasing the cluster size beyond 143 molecule can be considered negligible. In particular, the change in excitation energy of 0.02 eV (= 0.0007 a.u.) should be considered in contrast to the non-converged local region Hartree–Fock energy E_t of the ground state (see Table 1). The absolute differences in E_t for cluster sizes of 143, 324 and 579 molecules are 0.00349 a.u. and 0.02692 a.u. Hence, we see that excitation energies for local excitations may converge with cluster size, even if the ground state energy for the target region does not converge. It is important to note that this does not mean that the presented excitation energies are converged with respect to the chosen active occupied and virtual space. In Sect. 4.2 a study on choice of basis sets is carried out.

4.2 Effect of basis set

4.2.1 Effect of basis set on the Hartree–Fock electric dipole moment

In this section we present how the choice of basis set for shell 1 and 2 affects the dipole moment of the target region (central NH_3 molecule), for a cluster containing 143 molecules. The basis sets of shell 1 and 2 are varied, while for the target region and shell 3+ are kept to aug-cc-pVTZ and STO-3G, respectively. The results are presented in Table 4. If we first consider the results where STO-3G is used for shell 2, Table 4 contain results for using STO-3G, cc-pVDZ, cc-pVTZ, aug-cc-pVDZ and aug-cc-pVTZ. All computed dipole moment values for these basis sets in shell 1 are 2.14–2.15 D. Hence, the target electronic density is not appreciably affected by the choice of basis in shell 1. We next consider the calculations where the basis set of shell 2 is cc-pVDZ, and the basis sets of shell 1 are cc-pVDZ, cc-pVTZ and aug-cc-pVDZ. All calculations result in a dipole moment of 2.17 D, which is 0.02–0.03 D higher than that of the STO-3G in shell 2 calculations. We further see that changing the basis

Table 4 Dipole moment for the target region (central NH_3 molecule) computed for a cluster containing 143 molecules, using different basis sets for shell 1 and 2

Basis (shell 1)	Basis (shell 2)	Dipole moment [D]
STO-3G	STO-3G	2.14
cc-pVDZ	STO-3G	2.15
cc-pVTZ	STO-3G	2.15
aug-cc-pVDZ	STO-3G	2.14
aug-cc-pVTZ	STO-3G	2.14
cc-pVDZ	cc-pVDZ	2.17
cc-pVTZ	cc-pVDZ	2.17
aug-cc-pVDZ	cc-pVDZ	2.17
cc-pVTZ	cc-pVTZ	2.17
aug-cc-pVDZ	aug-cc-pVDZ	2.16

The basis sets used for the target region and shell 3+ are aug-cc-pVTZ and STO-3G, respectively

set in shell 2 to cc-pVTZ and aug-cc-pVDZ, the computed dipole moment stays at 2.16–2.17 D. Considering the mean absolute error in Hartree–Fock dipole moments for molecules is found to be 0.16 D [56], these variations are negligible. Hence, it can be concluded that the the basis sets used in shell 1 and 2 has a negligible effect on the dipole moment of the target region. This should be seen in context with the reasonably large basis set (aug-cc-pVTZ) used for the target region. The large basis set of the target region ensures that we do not rely on a large basis set in shell 1 to improve the description of the target region.

4.2.2 Effect of basis set on CC2-in-HF vertical excitation energies

In this section we explore how changing the basis set of shell 1 and 2 affects CC2-in-HF excitation energies, as well as the effect of increasing the active virtual space to also include shell 2. In contrast to the Hartree–Fock dipole moment which only relies on the electronic density in the target region, the excitation energies further requires an active virtual space which extends beyond the target region. Hence, basis set effects is expected to be important here.

Results for how the excitation energies are affected by varying the basis set of shell 1 and 2 is presented in Table 5. The excitation energies are computed for the cluster containing 143 molecules, with aug-cc-pVTZ being used for the target region (central NH_3 molecule) and STO-3G being used for shell 3+ (see Sect. 3.2). The occupied active region is set to the target region, while the virtual active region is set to either the target region + shell 1 or the target region + shell 1–2 (see Sect. 3.4).

We first consider the results for which STO-3G is used for shell 2. We see that calculations using target + shell 1 and target + shell 1-2 as active virtual region yields identical results with respect to varying the basis set of shell 1. Increasing the basis set of shell 1 from STO-3G to cc-pVDZ reduces the excitation energy by 0.20 eV. Increasing the basis set of shell 1 from cc-pVDZ to cc-pVTZ, however, does not change the excitation energies. Increasing the basis set of shell 1 to aug-cc-pVDZ and

Table 5 CC2 $S_0 \rightarrow S_1$ excitation energies (eV) showing the effect of varying the basis set of shell 1 and 2, using either the target region + shell 1 or the target region + shell 1–2 as the virtual active region

Basis (shell 1)	Basis (shell 2)	Virtual active region	
		Target + Shell 1	Target + Shell 1–2
STO-3G	STO-3G	9.26	9.26
DZ	STO-3G	9.06	9.06
TZ	STO-3G	9.06	9.06
aDZ	STO-3G	9.03	9.03
aTZ	STO-3G	9.03	9.03
DZ	DZ	9.07	9.06
TZ	DZ	9.07	9.05
aDZ	DZ	9.05	9.04
TZ	TZ	9.07	9.05
aDZ	aDZ	9.05	9.03

All calculations are performed for the cluster containing 143 molecules, with the target region set as the occupied active region and aug-cc-pVTZ and STO-3G being used for the target region and shell 3+, respectively

aug-cc-pVTZ yields an excitation energy 0.03 eV lower than when using cc-pVDZ to cc-pVTZ in shell 1. We therefore see that the basis set of shell 1 one must be of sufficient size, but that increasing the basis set of shell 1 beyond cc-pVDZ gives only modest changes in the excitation energies.

We next consider the results which go beyond STO-3G in shell 2. As seen from Table 5 the excitation energies are increased by between 0.00 and 0.02 eV relative to the results where STO-3G is used in shell 2. This is the case both when the active virtual region is target + shell 1 as well as when the active virtual region is target + shell 1–2. From this we see two main points; (1) the basis set quality in shell 1 is of greater importance than the quality of the basis set in shell 2, and (2) increasing the virtual region beyond shell 1 is of little importance even when using reasonably large (aug-cc-pVDZ) basis sets in shell 2. It therefore appears that the smaller virtual active region (target and shell 1) is sufficient to obtain converged excitation energies with respect to the virtual space. Note that since only the occupied space for the target region is included, no occupied relaxation effects are taken into account.

5 Conclusion

In this paper we use local occupied and virtual orbital spaces for cluster models of a NH_3 crystal to show the convergence characteristics of the Hartree–Fock energy and electronic density in a small target region (single NH_3 molecule). The size of the cluster models ranges from 13 to 953 NH_3 molecules. The calculations illustrate that although the energy of a target region will not converge with cluster size, the energy contribution which only depends on the electronic density of the target region will. Based on this it can be concluded that the electron density of the target region

converges and as a consequence local size-intensive properties may be computed. The convergence of the target electronic density with respect to cluster size yields the possibility to evaluate how the density (and hence properties) is affected by basis set choices in shells around the target region. The properties used for these numerical illustrations are the Hartree–Fock electric dipole moment of the target region, which only relies on the electronic density matrix and local CC2-in-HF vertical excitation energies, which also relies on the virtual space.

Since the electronic density converges with cluster size, so do the target region Hartree–Fock dipole moments. Converged values are obtained using approximately three shells around the central NH_3 molecule. We further show that the effect of quality of the basis set used in the shells outside the target region is negligible for the dipole moment calculations. Using a minimal basis seems to be sufficient to capture the long range effects of the shells on the electronic density of the target region. The lowest singlet CC2-in-HF vertical excitation energies are also seen to converge with cluster size and basis set, despite the fact that the ground state Hartree–Fock energy does not converge with cluster size. Excitation energies using occupied and virtual spaces spanning the target region and the target region + shell 1, respectively, show convergence with cluster size with only small variations beyond three shells. Increasing the virtual space beyond shell 1 is seen to effect the excitation energies in the order of 0.00–0.02 eV. Unlike for the dipole moment, the excitation energies requires that the basis set of shell 1 is of sufficient size, but the quality of the basis set for shells beyond has little effect. For the vertical excitation energies presented, it is important to note that the occupied space only comprise of the occupied space of the target region (the central NH_3 molecule). Including a larger occupied space would result in relaxation effects which would lead to lower excitation energies. However, the intention of this study is not to obtain quantitative accuracy in the lowest local excitation of the NH_3 crystal, but rather to illustrate the cluster and basis set dependencies of the excitation energies.

The paper thus demonstrates fundamental aspects relating to the convergence of electronic density of a target region, and hence local size-intensive properties, with respect to cluster size. It is therefore not necessary to require an infinite system and converged energy to compute size-intensive properties for a target region of a molecular crystal using the electronic density. We further see that although the cluster needs to be of a given size to obtain a converged electronic density in the target region, a minimal basis description of the outer shells are sufficient to capture the correct interaction with the target region. The concepts illustrated in this paper are attainable using any fragmentation or orbital space partitioned based approach, as long as it contain the effective interaction between the target region and the rest of the cluster.

Acknowledgements The authors acknowledge computing resources through UNINETT Sigma2 the National Infrastructure for High Performance Computing and Data Storage in Norway through Project no. nn9409k, and use of the IDUN HPC cluster at NTNU [57]. I-MH acknowledge funding from the Research Council of Norway through FRINATEK Project 275506.

Funding Open access funding provided by NTNU Norwegian University of Science and Technology (incl St. Olavs Hospital–Trondheim University Hospital)

Supplementary information Geometries of the clusters for presented results are available at https://dataverse.no/dataverse/ntnu. with DOI https://doi.org/10.18710/UOTXTK.

References

1. W. Kohn, Density functional and density matrix method scaling linearly with the number of atoms. Phys. Rev. Lett. **76**, 3168–3171 (1996). https://doi.org/10.1103/PhysRevLett.76.3168
2. W. Yang, Direct calculation of electron density in density-functional theory. Phys. Rev. Lett. **66**, 1438–1441 (1991). https://doi.org/10.1103/PhysRevLett.66.1438
3. E. Prodan, W. Kohn, Nearsightedness of electronic matter. Proc. Natl. Acad. Sci. **102**, 11635–11638 (2005). https://doi.org/10.1073/pnas.0505436102
4. I. Zhang, J. Jiang, B. Gao, X. Xu, Y. Luo, RRS-PBC: a molecular approach for periodic systems. Sci. China Chem. **57**, 1–6 (2014). https://doi.org/10.1007/s11426-014-5183-y
5. C. Huang, Embedded cluster density approximation for exchange-correlation energy: a natural extension of the local density approximation. J. Chem. Theory Comput. **14**, 6211–6225 (2018). https://doi.org/10.1021/acs.jctc.8b00471
6. M. Marsman, A. Grüneis, J. Paier, G. Kresse, Second-order Møller-Plesset perturbation theory applied to extended systems. I. Within the projector-augmented-wave formalism using a plane wave basis set. J. Chem. Phys. **130**, 184103 (2009). https://doi.org/10.1063/1.3126249
7. E. Rebolini, G. Baardsen, A.S. Hansen, K.R. Leikanger, T.B. Pedersen, Divide-expand-consolidate second-order Møller-Plesset theory with periodic boundary conditions. J. Chem. Theory Comput. **14**, 2427–2438 (2018). https://doi.org/10.1021/acs.jctc.8b00021
8. J. McClain, Q. Sun, G.K.L. Chan, T.C. Berkelbach, Gaussian-based coupled-cluster theory for the ground-state and band structure of solids. J. Chem. Theory Comput. **13**, 1209–1218 (2017). https://doi.org/10.1021/acs.jctc.7b00049
9. C. Pisani, M. Busso, G. Capecchi, S. Casassa, R. Dovesi, L. Maschio, C. Zicovich-Wilson, M. Schütz, Local-MP2 electron correlation method for nonconducting crystals. J. Chem. Phys. **122**, 094113 (2005). https://doi.org/10.1063/1.1857479
10. C. Edmiston, K. Ruedenberg, Localized atomic and molecular orbitals. Rev. Mod. Phys. **35**, 457–464 (1963). https://doi.org/10.1103/RevModPhys.35.457
11. S.F. Boys, Construction of some molecular orbitals to be approximately invariant for changes from one molecule to another. Rev. Mod. Phys. **32**, 296–299 (1960). https://doi.org/10.1103/RevModPhys.32.296
12. J. Pipek, P.G. Mezey, A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. J. Chem. Phys. **90**, 4916–4926 (1989). https://doi.org/10.1063/1.456588
13. P. Pulay, Localizability of dynamic electron correlation. Chem. Phys. Lett. **100**, 151–154 (1983). https://doi.org/10.1016/0009-2614(83)80703-9
14. C. Edmiston, M. Krauss, Pseudonatural orbitals as a basis for the superposition of configurations. I. He2+. J. Chem. Phys. **45**, 1833–1839 (1966). https://doi.org/10.1063/1.1727841
15. W. Meyer, Ionization energies of water from PNO-CI calculations. Int. J. Quantum Chem. **5**, 341–348 (1971). https://doi.org/10.1002/qua.560050839
16. W. Meyer, PNO-CI studies of electron correlation effects. I. Configuration expansion by means of nonorthogonal orbitals, and application to the ground state and ionized states of methane. J. Chem. Phys. **58**, 1017–1035 (1973). https://doi.org/10.1063/1.1679283
17. I.M. Høyvik, R.H. Myhre, H. Koch, Correlated natural transition orbitals for core excitation energies in multilevel coupled cluster models. J. Chem. Phys. **146**, 144109 (2017). https://doi.org/10.1063/1.4979908
18. C. Krause, H.J. Werner, Comparison of explicitly correlated local coupled-cluster methods with various choices of virtual orbitals. Phys. Chem. Chem. Phys. **14**, 7591–7604 (2012). https://doi.org/10.1039/C2CP40231A
19. A.S. Hansen, G. Baardsen, E. Rebolini, L. Maschio, T.B. Pedersen, Representation of the virtual space in extended systems—a correlation energy convergence study. Mol. Phys. **118**, e1733118 (2020). https://doi.org/10.1080/00268976.2020.1733118

20. I.M. Høyvik, The spectrum of the atomic orbital overlap matrix and the locality of the virtual electronic density matrix. Mol. Phys. **118**, e1765034 (2020). https://doi.org/10.1080/00268976.2020.1765034
21. H. Stoll, Correlation energy of diamond. Phys. Rev. B **46**, 6700–6704 (1992). https://doi.org/10.1103/PhysRevB.46.6700
22. S. Saebo, P. Pulay, Local treatment of electron correlation. Annu. Rev. Phys. Chem. **44**, 213–236 (1993). https://doi.org/10.1146/annurev.pc.44.100193.001241
23. S.R. Gadre, R.N. Shirsat, A.C. Limaye, Molecular tailoring approach for simulation of electrostatic properties. J. Phys. Chem. **98**, 9165–9169 (1994). https://doi.org/10.1021/j100088a013
24. C. Hampel, H. Werner, Local treatment of electron correlation in coupled cluster theory. J. Chem. Phys. **104**, 6286–6297 (1996). https://doi.org/10.1063/1.471289
25. P. Maslen, M. Head-Gordon, Non-iterative local second order Møller-Plesset theory. Chem. Phys. Lett. **283**, 102–108 (1998). https://doi.org/10.1016/S0009-2614(97)01333-X
26. M. Schütz, G. Hetzer, H.J. Werner, Low-order scaling local electron correlation methods. I. Linear scaling local MP2. J. Chem. Phys. **111**, 5691–5705 (1999). https://doi.org/10.1063/1.479957
27. K. Kitaura, E. Ikeo, T. Asada, T. Nakano, M. Uebayasi, Fragment molecular orbital method: an approximate computational method for large molecules. Chem. Phys. Lett. **313**, 701–706 (1999). https://doi.org/10.1016/S0009-2614(99)00874-X
28. G.E. Scuseria, P.Y. Ayala, Linear scaling coupled cluster and perturbation theories in the atomic orbital basis. J. Chem. Phys. **111**, 8330–8343 (1999). https://doi.org/10.1063/1.480174
29. S. Li, J. Ma, Y. Jiang, Linear scaling local correlation approach for solving the coupled cluster equations of large systems. J. Comput. Chem. **23**, 237–244 (2002). https://doi.org/10.1002/jcc.10003
30. M. Schütz, A new, fast, semi-direct implementation of linear scaling local coupled cluster theory. Phys. Chem. Chem. Phys. **4**, 3941–3947 (2002). https://doi.org/10.1039/B203994J
31. T. Crawford, R.A. King, Locally correlated equation-of-motion coupled cluster theory for the excited states of large molecules. Chem. Phys. Lett. **366**, 611–622 (2002). https://doi.org/10.1016/S0009-2614(02)01639-1
32. N. Flocke, R.J. Bartlett, A natural linear scaling coupled-cluster method. J. Chem. Phys. **121**, 10935–10944 (2004). https://doi.org/10.1063/1.1811606
33. S. Hirata, M. Valiev, M. Dupuis, S.S. Xantheas, S. Sugiki, H. Sekino, Fast electron correlation methods for molecular clusters in the ground and excited states. Mol. Phys. **103**, 2255–2265 (2005). https://doi.org/10.1080/00268970500083788
34. D. Kats, T. Korona, M. Schütz, Local CC2 electronic excitation energies for large molecules with density fitting. J. Chem. Phys. **125**, 104106 (2006). https://doi.org/10.1063/1.2339021
35. O. Christiansen, P. Manninen, P. Jørgensen, J. Olsen, Coupled-cluster theory in a projected atomic orbital basis. J. Chem. Phys. **124**, 084103 (2006). https://doi.org/10.1063/1.2173249
36. J.E. Subotnik, A. Sodt, M. Head-Gordon, A near linear-scaling smooth local coupled cluster algorithm for electronic structure. J. Chem. Phys. **125**, 074116 (2006). https://doi.org/10.1063/1.2336426
37. J. Friedrich, M. Hanrath, M. Dolg, Fully automated implementation of the incremental scheme: application to CCSD energies for hydrocarbons and transition metal compounds. J. Chem. Phys. **126**, 154110 (2007). https://doi.org/10.1063/1.2721538
38. M. Kobayashi, H. Nakai, Extension of linear-scaling divide-and-conquer-based correlation method to coupled cluster theory with singles and doubles excitations. J. Chem. Phys. **129**, 044103 (2008). https://doi.org/10.1063/1.2956490
39. F. Neese, F. Wennmohs, A. Hansen, Efficient and accurate local approximations to coupled-electron pair approaches: an attempt to revive the pair natural orbital method. J. Chem. Phys. **130**, 114108 (2009). https://doi.org/10.1063/1.3086717
40. M. Ziółkowski, B. Jansík, T. Kjærgaard, P. Jørgensen, Linear scaling coupled cluster method with correlation energy based error control. J. Chem. Phys. **133**, 014107 (2010). https://doi.org/10.1063/1.3456535
41. Z. Rolik, M. Kállay, A general-order local coupled-cluster method based on the cluster-in-molecule approach. J. Chem. Phys. **135**, 104111 (2011). https://doi.org/10.1063/1.3632085
42. J. Yang, G.K.L. Chan, F.R. Manby, M. Schütz, H.J. Werner, The orbital-specific-virtual local coupled cluster singles and doubles method. J. Chem. Phys. **136**, 144105 (2012). https://doi.org/10.1063/1.3696963
43. S. Sæther, T. Kjærgaard, H. Koch, I.M. Høyvik, Density-based multilevel Hartree-Fock model. J. Chem. Theory Comput. **13**, 5282–5290 (2017). https://doi.org/10.1021/acs.jctc.7b00689

44. S. Høst, J. Olsen, B. Jansík, L. Thøgersen, P. Jørgensen, T. Helgaker, The augmented Roothaan-Hall method for optimizing Hartree-Fock and Kohn-Sham density matrices. J. Chem. Phys. **129**, 124106 (2008). https://doi.org/10.1063/1.2974099
45. I.M. Høyvik, Convergence acceleration for the multilevel Hartree-Fock model. Mol. Phys. **118**, 1626929 (2020). https://doi.org/10.1080/00268976.2019.1626929
46. T. Helgaker, P. Jørgensen, J. Olsen, *Molecular Electronic Structure Theory*, 1st edn. (Wiley, New York, 2000)
47. I.M. Høyvik, K. Kristensen, T. Kjaergaard, P. Jørgensen, A perspective on the localizability of Hartree-Fock orbitals. Theoret. Chem. Acc. (2014). https://doi.org/10.1007/s00214-013-1417-x
48. K. Aidas, C. Angeli, K.L. Bak, V. Bakken, R. Bast, L. Boman, O. Christiansen, R. Cimiraglia, S. Coriani, P. Dahle, E.K. Dalskov, U. Ekström, T. Enevoldsen, J.J. Eriksen, P. Ettenhuber, B. Fernández, L. Ferrighi, H. Fliegl, L. Frediani, K. Hald, A. Halkier, C. Hättig, H. Heiberg, T. Helgaker, A.C. Hennum, H. Hettema, E. Hjertenæs, S. Høst, I.M. Høyvik, M.F. Iozzi, B. Jansík, H.J.A. Jensen, D. Jonsson, P. Jørgensen, J. Kauczor, S. Kirpekar, T. Kjærgaard, W. Klopper, S. Knecht, R. Kobayashi, H. Koch, J. Kongsted, A. Krapp, K. Kristensen, A. Ligabue, O.B. Lutnæs, J.I. Melo, K.V. Mikkelsen, R.H. Myhre, C. Neiss, C.B. Nielsen, P. Norman, J. Olsen, J.M.H. Olsen, A. Osted, M.J. Packer, F. Pawlowski, T.B. Pedersen, P.F. Provasi, S. Reine, Z. Rinkevicius, T.A. Ruden, K. Ruud, V.V. Rybkin, P. Sałek, C.C.M. Samson, A.S. de Merás, T. Saue, S.P.A. Sauer, B. Schimmelpfennig, K. Sneskov, A.H. Steindal, K.O. Sylvester-Hvid, P.R. Taylor, A.M. Teale, E.I. Tellgren, D.P. Tew, A.J. Thorvaldsen, L. Thøgersen, O. Vahtras, M.A. Watson, D.J.D. Wilson, M. Ziolkowski, H. Ågren, The Dalton quantum chemistry program system. WIREs Comput. Mol. Sci. **4**, 269–284 (2014). https://doi.org/10.1002/wcms.1172
49. B. Jansík, S. Høst, K. Kristensen, P. Jørgensen, Local orbitals by minimizing powers of the orbital variance. J. Chem. Phys. **134**, 194104 (2011). https://doi.org/10.1063/1.3590361
50. S.D. Folkestad, E.F. Kjønstad, R.H. Myhre, J.H. Andersen, A. Balbi, S. Coriani, T. Giovannini, L. Goletto, T.S. Haugland, A. Hutcheson, I.M. Høyvik, T. Moitra, A.C. Paul, M. Scavino, A.S. Skeidsvoll, Å.H. Tveten, H. Koch, eT 1.0: an open source electronic structure program with emphasis on coupled cluster and multilevel methods. J. Chem. Phys. **152**, 184103 (2020). https://doi.org/10.1063/5.0004713
51. S.D. Folkestad, E.F. Kjønstad, H. Koch, An efficient algorithm for Cholesky decomposition of electron repulsion integrals. J. Chem. Phys. **150**, 194112 (2019). https://doi.org/10.1063/1.5083802
52. A.W. Hewat, C. Riekel, The crystal structure of deuteroammonia between 2 and 180 K by neutron powder profile refinement. Acta Crystallogr. A **35**, 569–571 (1979). https://doi.org/10.1107/S0567739479001340
53. T.H. Dunning, Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. J. Chem. Phys. **90**, 1007–1023 (1989). https://doi.org/10.1063/1.456153
54. R. Ditchfield, W.J. Hehre, J.A. Pople, Self-consistent molecular-orbital methods. IX. An extended gaussian-type basis for molecular-orbital studies of organic molecules. J. Chem. Phys. **54**, 724–728 (1971). https://doi.org/10.1063/1.1674902
55. I.M. Høyvik, B. Jansik, P. Jørgensen, Trust region minimization of orbital localization functions. J. Chem. Theory Comput. **8**, 3137–3146 (2012). https://doi.org/10.1021/ct300473g
56. K.L. Bak, J. Gauss, T. Helgaker, P. Jørgensen, J. Olsen, The accuracy of molecular dipole moments in standard electronic structure calculations. Chem. Phys. Lett. **319**, 563–568 (2000). https://doi.org/10.1016/S0009-2614(00)00198-6
57. M. Själander, M. Jahre, G. Tufte, N. Reissmann. EPIC: an energy-efficient, high-performance GPGPU computing research infrastructure (2019)

Paper C

Coupled cluster core excitation spectra of methane and carbon dioxide clathrates

A. Hutcheson, I-M. Høyvik,

In preparation

ISBN 978-82-326-5466-6 (printed ver.)
ISBN 978-82-326-6388-0 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU
Norwegian University of
Science and Technology