The multilevel CC2 and CCSD methods with correlated natural transition orbitals

Sarai Dery Folkestad[†] and Henrik Koch[*,‡]

[†]*Department of Chemistry, Norwegian University of Science and Technology, N-7491 Trondheim, Norway*

[‡]*Scuola Normale Superiore, Piazza dei Cavaleri 7, 56126 Pisa, Italy*

E-mail: henrik.koch@sns.it

Abstract

In the multilevel coupled cluster approach, an active orbital space is treated at a higher level of coupled cluster theory than the remaining inactive orbitals. We introduce the multilevel CC2 method where CC2 is used for the active orbital space. Furthermore, we present a simplified formulation of the multilevel CCSD method where CCSD is used for the active space. The simplification lies in the evaluation of the CC2-amplitudes in the inactive space; these CC2-amplitudes have previously been determined iteratively. We use correlated natural transition orbitals to determine the active orbital spaces. The convergence of the multilevel CC2 and multilevel CCSD valence excitation energies is established with proof-of-concept calculations. The methods are also applied to two larger systems: paranitroaniline in water and amoxicillin. The calculations on the paranitroaniline-water system illustrate the usefulness of multilevel coupled cluster methods for molecules in solution and for charge transfer excitations.

1 Introduction

The steep scaling of coupled cluster theory[1,2] renders it impractical for the treatment of large molecular systems, such as molecules in solvent, polymers, and macromolecules. Motivated by its accuracy for small and medium sized molecules, great efforts have been devoted to the development of reduced scaling coupled cluster methods.

One strategy to reduce the cost of an electronic structure calculation is to introduce approximations of the electron repulsion integrals. The resolution of identity method[3–5] using prefitted auxiliary basis sets,[6–8] and the Cholesky decomposition of the integral matrix[9–11] are well-established methods. In both approaches, a higher rank tensor—the electron repulsion integrals—is approximated by contractions of lower rank tensors. In this way reduced scaling and storage requirements may be obtained.

Another strategy to reduce computational costs is to approximate the wave function rather than (or in addition to) introducing approximations of the Hamiltonian. The scaling of correlated methods may be reduced by taking advantage of the localizability of dynamical correlation.[12] In the projected atomic orbital (PAO) method, introduced by Pulay and Sæbø,[12,13] localized occupied orbitals are used together with PAOs that span the virtual space. The approach was successfully applied to Møller-Plesset perturbation theory,[14,15] and configuration interaction.[16] Hampel and Werner,[17] and Werner and Schütz[18] extended the use of PAOs to coupled cluster theory. Computational savings are obtained by defining a truncated set of virtual orbitals for each occupied orbital; only PAOs spatially close to the localized occupied orbitals are considered. Furthermore, double (and higher) excitations are neglected when the occupied orbitals are distant.

Following these developments, a plethora of local coupled cluster methods have been proposed, a selection of which may be found in Refs. 19–30. These local coupled cluster methods aim to reduce scaling—toward linear scaling—while retaining a high quality description of extensive properties such as the correlation energy.

In chemical application, intensive properties are usually of greater interest than extensive

properties. For intensive properties, such as excitation energies and transition moments, reduced space, multilevel, or embedding methods may be sufficient.

Arguably, the most common reduced space method is the frozen-core approximation; the approximation is applied by default in some quantum chemistry software packages. Truncation of the virtual space has also been explored,[31] however, not with the same success.

Reduced space coupled cluster methods for excitation energies based on the local coupled cluster PAO method, have also been developed.[32,33] The truncation of the PAOs can not be based solely on locality for the description of excited states. Information from the excitation vectors of cheaper methods, such as CIS, is used to determine which virtual orbitals to include for a given of local MO. Similar reduced space methods have also been developed using pair natural orbitals, rather than PAOs.[34]

In the LoFEx method developed by Baudin *et al.*,[35,36] coupled cluster singles and perturbative doubles[37] (CC2) and coupled cluster singles and doubles[38] (CCSD) excitation energies are approximated in a reduced space. A mixed orbital basis of natural transition orbitals (NTOs)[39,40] and local molecular orbitals are used. The number of orbitals included in the calculation is increased until the excitation energies are converged to within a preset threshold. The CorNFLEx method[41] employs correlated natural transition orbitals (CNTOs)[42] rather than NTOs. The CNTOs are constructed using excitation vectors from a (correlated) coupled cluster calculation and offer a compact description of the corresponding excitations.

In multilevel and embedding methods, different regions of the molecular system are treated at different levels of theory. These methods may use a combination of electronic structure methods, classical atomistic methods, or even continuum methods. This approach was made popular with the QM/MM method.[43–45] Other notable contributions include the ONIOM[46] methods and the polarizable continuum models.[47]

The multilevel coupled cluster method was introduced by Myhre, Sánchez de Merás, and Koch with the extended CC2 method, which we refer to as multilevel CCSD (ML-CCSD),[48,49] and later with multilevel coupled cluster singles, doubles and perturbative triples

(MLCC3).[50] In multilevel coupled cluster methods the molecular orbitals, both occupied and virtual, are partitioned into sets of so-called active and inactive orbitals. These sets are treated at different level of coupled cluster theory. Excitations with respect to the Hartree-Fock reference are characterized according to whether they are excitations *internally* within the active orbital space, *externally* within the inactive orbital space, or *semi-externally* between the active and inactive orbital spaces. Each excitation level of the cluster operator may then be restricted to include only some types of excitations. For instance, in MLCC3 the cluster operator includes internal, semi-external and external single and double excitations, but only internal triple excitations. Moreover, these triple excitations are treated perturbatively. Multilevel coupled cluster methods can be used to describe intensive properties with an accuracy approaching that of the higher level coupled cluster method, but at the cost of the lower level method. This has been demonstrated for both core and valence excitation energies.[42,49,50]

The success of a multilevel coupled cluster calculation relies heavily on the choice of orbitals to partition. Myhre *et al.* used Cholesky orbitals[51] for MLCCSD[48,49] and MLCC3.[50] These orbitals are constructed by partially decomposing the occupied and virtual Hartree-Fock densities with pivots corresponding to AOs centered on preselected active atoms. The use of Cholesky orbitals requires knowledge of the active site of the molecular property being calculated. They are particularly effective for core excitations and UV/VIS excitations on small and medium sized molecules in solvent. By using CNTOs in multilevel coupled cluster theory we avoid the need for *a priori* knowledge of the active site. The CNTOs are tailored to describe the excitation of interest and may also be used for excitations that are delocalized. However, they rely on the accuracy of the lower level method. Høyvik *et al.* showed that MLCC3 with CNTOs yield good results for core excitations.[42]

In this work, we formulate the multilevel CC2 (MLCC2) method, where CC2 is used for the active space and the CCS is used for the inactive space. A reformulation of the MLCCSD method is also presented, where CCSD is used for the active space and CC2 or

CCS, or both, is used for the inactive space. In our new formulation of MLCCSD, the CC2 double excitation amplitudes are determined from an analytical expression (similar to the approach of standard CC2 with canonical orbitals) and not iteratively as was previously done by Myhre *et al.*[48] To determine our active spaces, we use CNTOs generated from the lower level calculation. We present valence excitation energies for cytosine, guanine, amoxicillin, and formaldehyde-water and paranitroaniline-water systems, demonstrating the smooth convergence of the multilevel coupled cluster excitation energies to that of the higher level method. We also show that multilevel coupled cluster methods may be used for charge transfer excitations.

2 Theory

The wave function in MLCC2 and MLCCSD may be written as

$$|\mathrm{MLCC}\rangle = \exp(X_1 + X_2)|\mathrm{R}\rangle \tag{1}$$

where $|\mathrm{R}\rangle$ is the restricted Hartree-Fock reference. In the closed-shell, spin adapted formulation of coupled cluster theory, the single excitation operator X_1 is defined as

$$X_1 = \sum_{ai} x_i^a E_{ai} \tag{2}$$

where x_i^a is an excitation amplitude and E_{ai} is a singlet excitation operator. We use $a, b, ...$ to denote virtual orbitals, and $i, j, ...$ for occupied orbitals. In both MLCC2 and MLCCSD, X_1 includes internal, semi-external, and external excitations. The definition of the double excitation operator X_2 differs for MLCC2 and MLCCSD.

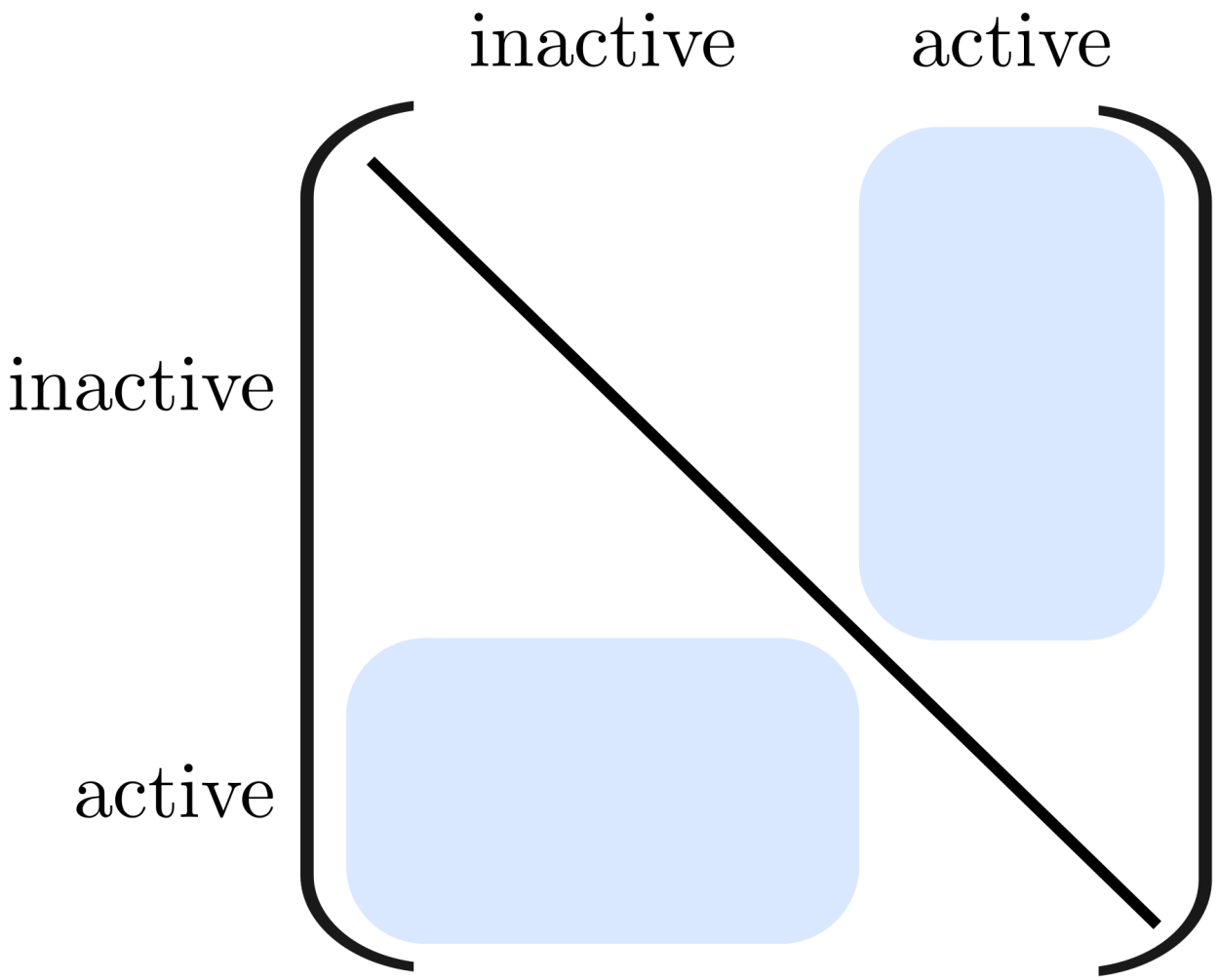

Figure 1: Structure of the occupied-occupied and virtual-virtual blocks of the Fock matrix in the basis used for an MLCC2 calculation. The diagonal (inactive-inactive and active-active) blocks are diagonal, whereas the off-diagonal (active-inactive) blocks, indicated by the blue fields, are non-zero.

2.1 Multilevel CC2

In MLCC2, we consider a set of active orbitals, denoted $\mathcal{O}^{\mathrm{CC2}}$, and a set of inactive orbitals, denoted $\mathcal{O}^{\mathrm{CCS}}$. The double excitation operator X_2 is given by

$$X_2 = S_2 = \frac{1}{2} \sum_{aibj} s_{ij}^{ab} E_{ai} E_{bj}, \quad \{a, i, b, j\} \in \mathcal{O}^{\mathrm{CC2}} \tag{3}$$

where s_{ij}^{ab} is a double excitation amplitude. The indices are restricted to $\mathcal{O}^{\mathrm{CC2}}$ and S_2 generates internal double excitations.

The Hamiltonian of the system is divided into the Fock operator F and a fluctuation potential U:

$$H = F + U \tag{4}$$

S_2 is determined to first order in U. The projected coupled cluster equations are

$$\Omega_{\mu_1} = \langle \mu_1 | \, \tilde{H} + [\tilde{H}, S_2] \, | \mathrm{R} \rangle = 0 \tag{5}$$

$$\Omega_{\mu_2^{\mathrm{S}}} = \langle \mu_2^{\mathrm{S}} | \, \tilde{H} + [F, S_2] \, | \mathrm{R} \rangle = 0 \tag{6}$$

where $\tilde{H} = \exp(-X_1) H \exp(X_1)$ is the X_1-transformed Hamiltonian. The projection space $\{\langle \mu_2^S |\}$ is associated with the S_2 operator. The equations are solved in a basis where the occupied-occupied and the virtual-virtual blocks of the Fock matrix are block diagonal, see Figure 1. Since S_2 only generates internal excitations, we obtain a closed form expression for the s-amplitudes from eq (6) in this basis:

$$s_{ij}^{ab} = -\frac{\tilde{g}_{aibj}}{\epsilon_{ij}^{ab}}, \quad \{a, i, b, j\} \in \mathcal{O}^{\mathrm{CC2}} \tag{7}$$

Here, $\epsilon_{ij}^{ab} = F_{aa} + F_{bb} - F_{ii} - F_{jj}$, and $\tilde{g}_{aibj}$ are X_1-transformed electron repulsion integrals. Except for the restriction of some summations to the active orbitals, the MLCC2 equations are equivalent to the CC2 equations.

Excitation energies are size-intensive and may therefore be well described within the multilevel coupled cluster theory. Excitation energies in the multilevel coupled cluster framework can be obtained through linear response (LR) theory—in a manner analogous to standard LR-CC.[52,53] The MLCC2 excitation energies correspond to the eigenvalues of the MLCC2

Jacobian,

$$\boldsymbol{A}^{\mathrm{MLCC2}} = \begin{pmatrix} \langle \mu_1 \,|\, [\tilde{H}, \tau_{\nu_1}] + [[\tilde{H}, \tau_{\nu_1}], S_2] \,|\, R \rangle & \langle \mu_1 \,|\, [\tilde{H}, \tau_{\nu_2^S}] \,|\, R \rangle \\ \langle \mu_2^S \,|\, [\tilde{H}, \tau_{\nu_1}] \,|\, R \rangle & \langle \mu_2^S \,|\, [F, \tau_{\nu_2^S}] \,|\, R \rangle \end{pmatrix} \tag{8}$$

and may be obtained through the use of standard algorithms such as the Davidson algorithm.[54,55] Here, τ_{ν_1} is a single excitation operator and $\tau_{\nu_2^S}$ is an internal double excitation operator. The MLCC2 excited state equations differ from the CC2 excited state equations in the restriction of some summations to the active orbitals.

The most expensive contractions in CC2 scale as v^3o^2, where v is the number of virtual orbitals and o is the number of occupied orbitals. These contractions enter the equations that determine the cluster amplitudes and the excitation vectors. In an MLCC2 calculation, the s-amplitudes are restricted to $\mathcal{O}^{\mathrm{CC2}}$ and the corresponding contractions scale as V^2O^2v, where O and V denote the number of active occupied and active virtual orbitals, respectively. Hence, if the active space is kept fixed while the system is expanded, these contractions will initially scale linearly with respect to system size. However, all these terms involve X_1-transformed integrals of the form $\tilde{g}_{abic}$ and $\tilde{g}_{baci}$ where $\{b, i, c\} \in \mathcal{O}^{\mathrm{CC2}}$ and $a \in \mathcal{O}^{\mathrm{CCS}} \cup \mathcal{O}^{\mathrm{CC2}}$ which become small when the orbital ϕ_a is localized far away from the active space. In this implementation of multilevel coupled cluster, we do not exploit the sparsity of these two-electron integrals.

2.2 Multilevel CCSD

The MLCCSD framework allows for two-level (CCS/CCSD or CC2/CCSD) and three-level (CCS/CC2/CCSD) calculations. Hence, there are at most three orbital sets to consider: CCS orbitals ($\mathcal{O}^{\mathrm{CCS}}$), CC2 orbitals ($\mathcal{O}^{\mathrm{CC2}}$), and CCSD orbitals ($\mathcal{O}^{\mathrm{CCSD}}$). In MLCCSD, the double excitation operator X_2 is given by

$$X_2 = S_2 + T_2 \tag{9}$$

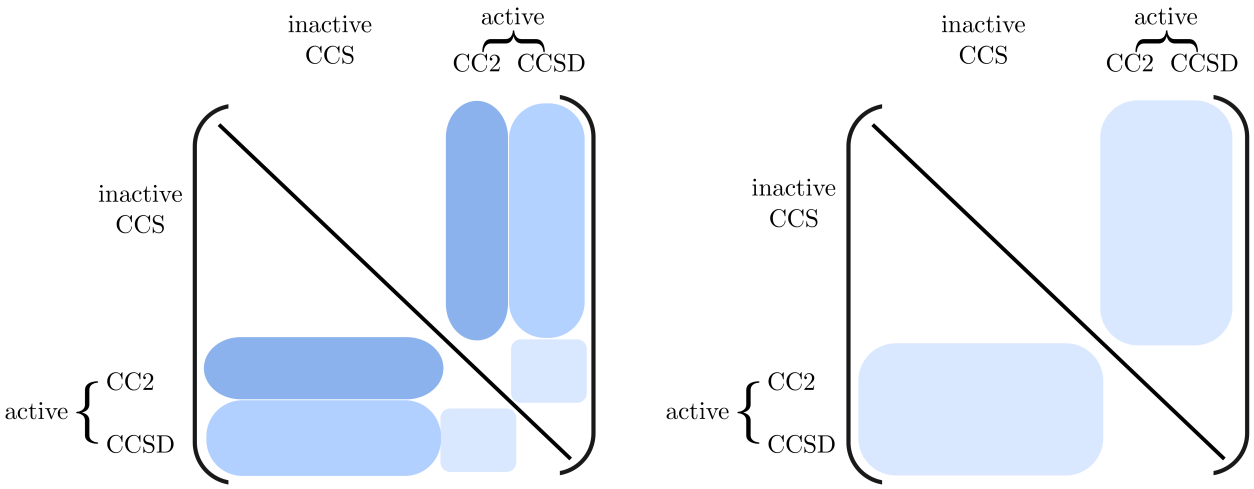

Figure 2: Structure of the occupied-occupied and virtual-virtual blocks of the Fock matrix in the bases used in an MLCCSD calculation. The basis used to construct the S_2-amplitudes is depicted on the right, and the basis used for the MLCCSD calculation is on the left. The diagonal blocks are diagonal and the blue fields indicate the active-inactive off-diagonal blocks which are non-zero.

with

$$T_2 = \frac{1}{2}\sum_{aibj} t_{ij}^{ab} E_{ai} E_{bj}, \quad \{a,i,b,j\} \in \mathcal{O}^{\mathrm{CCSD}} \tag{10}$$

In MLCCSD, we limit the indices of S_2 to the CC2 and CCSD orbital spaces:

$$S_2 = \frac{1}{2}\sum_{aibj} s_{ij}^{ab} E_{ai} E_{bj}, \quad \{a,i,b,j\} \in \mathcal{O}^{\mathrm{CC2}} \cup \mathcal{O}^{\mathrm{CCSD}} \tag{11}$$

The S_2 operator is still determined to first order in U and the t-amplitudes act as corrections to the s-amplitudes in the active CCSD orbital space.

The projected MLCCSD equations are

$$\Omega_{\mu_1} = \langle \mu_1 | \tilde{H} + [\tilde{H}, X_2] | \mathrm{R} \rangle = 0 \tag{12}$$

$$\Omega_{\mu_2^{\mathrm{S}'}} = \langle \mu_2^{\mathrm{S}'} | \tilde{H} + [F, S_2] + [\tilde{H}, T_2] | \mathrm{R} \rangle = 0 \tag{13}$$

$$\Omega_{\mu_2^{\mathrm{T}}} = \langle \mu_2^{\mathrm{T}} | \tilde{H} + [\tilde{H}, X_2] + \frac{1}{2}[[\tilde{H}, X_2], X_2] | \mathrm{R} \rangle = 0 \tag{14}$$

where the projection vectors $\{\langle\mu_2^T|\}$ are associated with T_2 and disjoint with $\{\langle\mu_2^{S'}|\}$, which is associated with the excitations of S_2 that are not internal to $\mathcal{O}^{\mathrm{CCSD}}$.

In MLCC2, we use the analytic expression for the s-amplitudes in eq (7). However, in MLCCSD we must consider eqs (13) and (14), and the s-amplitudes can not be obtained directly. We may, however, solve iteratively for the s-amplitudes. This approach was taken by Myhre et al.[48,49] In our formulation of MLCCSD, we have chosen to approximate the s-amplitudes from the expression

$$\langle\mu_2^{\mathrm{S}}|\,\tilde{H} + [F, S_2]\,|\mathrm{R}\rangle = 0 \tag{15}$$

where we use the complete doubles projection space $\{\langle\mu_2^S|\} = \{\langle\mu_2^{S'}|\} \cup \{\langle\mu_2^T|\}$ associated with S_2. We solve eq (15) in a basis where the CC2/CCSD-blocks of the occupied-occupied and virtual-virtual Fock matrices are diagonal, see Figure 2. In this basis, a closed form expression for the s-amplitudes is available:

$$s_{IJ}^{AB} = -\frac{\tilde{g}_{AIBJ}}{\epsilon_{IJ}^{AB}} \tag{16}$$

We use upper case letters to denote orbitals in the basis where the CC2-amplitudes are determined. The amplitudes s_{IJ}^{AB}, must subsequently be transformed to the proper basis for the MLCCSD calculation. The change of basis entails a mixing of CC2 and CCSD occupied orbitals and of CC2 and CCSD virtual orbitals. Thus we have:

$$s_{ij}^{ab} = \sum_{AIBJ} s_{IJ}^{AB} O_{Aa} O_{Bb} O_{Ii} O_{Jj}, \quad \{a, b, i, j\} \in \mathcal{O}^{\mathrm{CC2}} \cup \mathcal{O}^{\mathrm{CCSD}} \tag{17}$$

The orthogonal matrix $\boldsymbol{O}$, connecting the two MO bases, is obtained in the following way. We consider the transformations of a matrix in the AO basis, $\boldsymbol{Y}^{\mathrm{AO}}$, to the two different MO

bases

$$\boldsymbol{Y} = \boldsymbol{C}^T \boldsymbol{Y}^{\mathrm{AO}} \boldsymbol{C} \tag{18}$$

$$\bar{\boldsymbol{Y}} = \bar{\boldsymbol{C}}^T \boldsymbol{Y}^{\mathrm{AO}} \bar{\boldsymbol{C}} \tag{19}$$

The matrices $\boldsymbol{C}$ and $\bar{\boldsymbol{C}}$ contain the orbital coefficients of the MLCCSD basis and the basis where the s-amplitudes are constructed, respectively. From the orthonormality relations

$$\boldsymbol{C}^T \boldsymbol{S} \boldsymbol{C} = \boldsymbol{I}, \quad \bar{\boldsymbol{C}}^T \boldsymbol{S} \bar{\boldsymbol{C}} = \boldsymbol{I}, \tag{20}$$

where $\boldsymbol{S}$ is the AO overlap matrix, we have

$$\boldsymbol{C}^{-1} = \boldsymbol{C}^T \boldsymbol{S}, \quad \bar{\boldsymbol{C}}^{-1} = \bar{\boldsymbol{C}}^T \boldsymbol{S} \tag{21}$$

$$\boldsymbol{C}^{-T} = \boldsymbol{S} \boldsymbol{C}, \quad \bar{\boldsymbol{C}}^{-T} = \boldsymbol{S} \bar{\boldsymbol{C}} \tag{22}$$

From eq (18) we get

$$\boldsymbol{Y}^{\mathrm{AO}} = \boldsymbol{S} \boldsymbol{C} \boldsymbol{Y} \boldsymbol{C}^T \boldsymbol{S} \tag{23}$$

which upon insertion into eq (19) yields the relation

$$\bar{\boldsymbol{Y}} = \bar{\boldsymbol{C}}^T \boldsymbol{S} \boldsymbol{C} \boldsymbol{Y} \boldsymbol{C}^T \boldsymbol{S} \bar{\boldsymbol{C}} = (\bar{\boldsymbol{C}}^T \boldsymbol{S} \boldsymbol{C}) \boldsymbol{Y} (\bar{\boldsymbol{C}}^T \boldsymbol{S} \boldsymbol{C})^T \tag{24}$$

between the matrices in the two different MO bases. The transformation matrix is thus

$$\boldsymbol{O} = \bar{\boldsymbol{C}}^T \boldsymbol{S} \boldsymbol{C} \tag{25}$$

and its orthogonality follows directly from the conditions in eq (20).

The s-amplitudes are calculated from eqs (16) and (17). The amplitudes of X_2, are then

constructed by taking the s-amplitudes and replacing the amplitudes corresponding to the excitations internal to $\mathcal{O}^{\mathrm{CCSD}}$ with the t-amplitudes obtained from the previous iteration. Equations (12) and (14) are then solved as in standard CCSD theory, but with the doubles projection space restricted to $\{\langle\mu_2^T|\}$ and the double excitation operator X_2 restricted to $\mathcal{O}^{\mathrm{CC2}} \cup \mathcal{O}^{\mathrm{CCSD}}$. The MLCCSD excitation energies are obtained as the eigenvalues of the MLCCSD Jacobian $\boldsymbol{A}^{\mathrm{MLCCSD}}$

$$\boldsymbol{A}^{\mathrm{MLCCSD}} = \begin{pmatrix} \langle\mu_1|[\tilde{H},\tau_{\nu_1}]+[[\tilde{H},\tau_{\nu_1}],X_2]|R\rangle & \langle\mu_1|[\tilde{H},\tau_{\nu_2^S}]|R\rangle & \langle\mu_1|[\tilde{H},\tau_{\nu_2^T}]|R\rangle \\ \langle\mu_2^{S'}|[\tilde{H},\tau_{\nu_1}]|R\rangle & \langle\mu_2^{S'}|[F,\tau_{\nu_2^{S'}}]|R\rangle & \langle\mu_2^{S'}|[F,\tau_{\nu_2^T}]|R\rangle \\ \langle\mu_2^T|[\tilde{H},\tau_{\nu_1}]+[[\tilde{H},\tau_{\nu_1}],X_2]|R\rangle & \langle\mu_2^T|[\tilde{H},\tau_{\nu_2^{S'}}]+[[\tilde{H},\tau_{\nu_2^{S'}}],X_2]|R\rangle & \langle\mu_2^T|[\tilde{H},\tau_{\nu_2^T}]+[[\tilde{H},\tau_{\nu_2^T}],X_2]|R\rangle \end{pmatrix} \tag{26}$$

Here, $\tau_{\nu_2^{S'}}$ denotes double excitations that are not internal to $\mathcal{O}^{\mathrm{CCSD}}$. The matrix $\boldsymbol{A}^{\mathrm{MLCCSD}}$ is derived by differentiating the projected MLCCSD equations with respect to the $\boldsymbol{x}$ amplitudes. The most expensive term in the ground and excited state equations of standard CCSD is a contraction between double amplitudes and the electron repulsion integrals with four virtual indices, e.g., $\sum_{cd} \tilde{g}_{acbd} t_{ij}^{cd}$. This contraction has an operation count $\propto v^4 o^2$. In MLCCSD, the indices of the double amplitudes are restricted to $\mathcal{O}^{\mathrm{CCSD}} \cup \mathcal{O}^{\mathrm{CC2}}$ and the indices of the resulting vector are restricted to $\mathcal{O}^{\mathrm{CCSD}}$. We have

$$\sum_{cd} \tilde{g}_{acbd} x_{ij}^{cd}, \quad \{a,i,b,j\} \in \mathcal{O}^{\mathrm{CCSD}} \text{ and } \{c,d\} \in \mathcal{O}^{\mathrm{CCSD}} \cup \mathcal{O}^{\mathrm{CC2}} \tag{27}$$

If the system size is increased, but $\mathcal{O}^{\mathrm{CCSD}}$ is fixed, this contraction will scale quadratically with the number of virtual orbitals in $\mathcal{O}^{\mathrm{CC2}}$. If only $\mathcal{O}^{\mathrm{CCS}}$ is enlarged, this term will not scale with the system. Moreover, the X_1-transformed integral $\tilde{g}_{acbd}$ for $\{a,b\} \in \mathcal{O}^{\mathrm{CCSD}}$ and $\{c,d\} \in \mathcal{O}^{\mathrm{CCSD}} \cup \mathcal{O}^{\mathrm{CC2}}$ becomes small when ϕ_c or ϕ_d are localized far away from the active CCSD orbitals ϕ_a and ϕ_b, respectively.

2.3 Correlated natural transition orbitals

Multilevel coupled cluster methods rely on the use of orbitals that may be sensibly partitioned depending on the property of interest. Correlated natural transition orbitals (CNTOs)[42] are constructed using the excitation vectors from a (correlated) coupled cluster calculation. They are tailored to compactly describe the corresponding excited states. In the context of a multilevel coupled cluster calculation, the excitation vectors are obtained from the lower level method and the CNTOs are used to determine the active and inactive orbital spaces.

The CNTOs are generated by constructing and diagonalizing two matrices, denoted $\boldsymbol{M}$ and $\boldsymbol{N}$, which are defined below. The eigenvectors of $\boldsymbol{M}$ and $\boldsymbol{N}$ form the CNTO transformation matrices for the occupied and virtual orbitals, respectively. The $\boldsymbol{M}$ and $\boldsymbol{N}$ matrices are defined in terms of excitation vectors $\boldsymbol{R}$ from CC2 or CCSD calculations:

$$\begin{aligned} M_{ij}^{\mathrm{CNTO}} = & \sum_{a} R_{ai} R_{aj} \\ & + \frac{1}{2} \sum_{abk} (1 + \delta_{ai,bk}\delta_{ij}) R_{aibk} R_{ajbk} \end{aligned} \tag{28}$$

$$\begin{aligned} N_{ab}^{\mathrm{CNTO}} = & \sum_{i} R_{ai} R_{bi} \\ & + \frac{1}{2} \sum_{ijc} (1 + \delta_{ai,cj}\delta_{ab}) R_{aicj} R_{bicj} \end{aligned} \tag{29}$$

In MLCC2, the lower level method is CCS. With excitation vectors from a CCS calculation, the definition of the $\boldsymbol{M}$ and $\boldsymbol{N}$ matrices reduces to

$$M_{ij}^{\mathrm{NTO}} = \sum_{a} R_{ai} R_{aj} \tag{30}$$

$$N_{ab}^{\mathrm{NTO}} = \sum_{i} R_{ai} R_{bi} \tag{31}$$

and we obtain the natural transition orbitals (NTOs).[39] NTOs are suitable to describe excitations with strong single excitation character. However, as the rank of the matrix $\boldsymbol{N}^{\mathrm{NTO}}$ cannot exceed the number of occupied orbitals, NTOs can not be used to select an active

virtual orbital space.

Baudin and Kristensen[41] introduced a method to generate CNTOs from a CCS calculation by constructing approximate double excitation vectors. They define

$$R^{\mathrm{CCS}}_{aibj} = -\frac{1}{1+\delta_{ai,bj}} \frac{\hat{g}_{aibj}}{\epsilon^{ab}_{ij} - \omega^{\mathrm{CCS}}} \tag{32}$$

where ω^{CCS} is the CCS excitation energy, and $\epsilon^{ab}_{ij} = \epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j$ where the ϵ_q are orbital energies. The integrals $\hat{g}_{aibj}$ are defined as

$$\hat{g}_{aibj} = P^{ab}_{ij}\Big(\sum_c R_{ci} g_{bjac} - \sum_k R_{bk} g_{kjai} \Big) \tag{33}$$

where g_{pqrs} are the electronic repulsion integrals in the molecular orbital basis, and $P^{ab}_{ij} I_{ai,bj} = I_{ai,bj} + I_{bj,ai}$. These approximate double excitation vectors may be used to generate $\boldsymbol{M}$ and $\boldsymbol{N}$ according to eqs (28) and (29). The construction of $\boldsymbol{R}^{\mathrm{CCS}}$ entails a v^3o^2 operation and may become a bottleneck for very large systems. However, they may be generated in a reduced space, as was also proposed by Baudin and Kristensen[41]. The construction of the CNTOs from eqs (29) and (28) is also a v^3o^2 operation.

In order to describe several excitation energies, CNTOs are constructed by diagonalizing modified $\boldsymbol{M}$ and $\boldsymbol{N}$ matrices:

$$\boldsymbol{M} = \sum_k \boldsymbol{M}_k \tag{34}$$

$$\boldsymbol{N} = \sum_k \boldsymbol{N}_k \tag{35}$$

where $\boldsymbol{M}_k$ is constructed from eqs (30) or (28) using the k'th excitation vector, and similarly for $\boldsymbol{N}_k$. We may alternatively use a separate basis for each excitation. This may result in a smaller active space being necessary to treat the targeted excitation. Note, however, that separate bases will complicate the calculation of transition moments between states because

of non-orthogonality.

The active occupied and active virtual orbitals are those that correspond to the largest eigenvalues of $\boldsymbol{M}$ and $\boldsymbol{N}$, respectively. In this work, we only specify the number of active occupied orbitals n_o^{active}. The number of active virtual orbitals n_v^{active} is obtained from the ratio of the number of virtual orbitals n_v to occupied orbitals n_o in the full set:

$$n_v^{\text{active}} = \left(\frac{n_v}{n_o}\right) n_o^{\text{active}} \tag{36}$$

We may then control the size of the active space as we compare multilevel coupled cluster to the standard coupled cluster methods. Another way to select the active space is to consider the magnitude of the eigenvalues of $\boldsymbol{M}$ and $\boldsymbol{N}$: the sum of the eigenvalues of $\boldsymbol{M}$ and $\boldsymbol{N}$ is by construction equal to the number k of excitation vectors used to construct $\boldsymbol{M}$ and $\boldsymbol{N}$, see eqs (34) and (35). The active space may therefore be selected such that

$$1 - \frac{1}{k} \sum_o \lambda_o^M < \tau_M \tag{37}$$

$$1 - \frac{1}{k} \sum_v \lambda_v^N < \tau_N \tag{38}$$

where λ_o^M and λ_v^N are eigenvalues of $\boldsymbol{M}$ and $\boldsymbol{N}$, corresponding to the active orbitals, and τ_M and τ_N thresholds, which may be specified by the user.

The CNTOs are used to select the active orbitals, however, the multilevel calculation is not carried out in this CNTO basis. The calculation is performed in a basis where the Fock matrix is block diagonal, see Figures 1 and 2. After the orbitals are partitioned, we construct and block-diagonalize the occupied-occupied part of the Fock matrix. This entails a mixing among the active occupied orbitals and a mixing among the inactive occupied orbitals, respectively. The same procedure is taken with the virtual-virtual part of the Fock matrix.

2.4 The computational procedure

An MLCC2 or MLCCSD calculation with CNTOs can be summarized in the following steps:

1. A lower level calculation is performed to obtain the excitation vectors $\boldsymbol{R}$.

2. The CNTO transformation matrices are constructed by diagonalizing the $\boldsymbol{M}$ and $\boldsymbol{N}$ matrices defined in eqs (28) and (29). If $\boldsymbol{R}$ is obtained from CCS, the approximated doubles excitation vectors are first generated from eqs (32) and (33), and the resulting total excitation vector is normalized.

3. The canonical MOs are transformed to the CNTO basis and the orbitals are partitioned into active and inactive sets.

4. The occupied-occupied and virtual-virtual parts of the Fock matrix are constructed in the CNTO basis and block-diagonalized (see Figures 1 and 2). The MOs are transformed accordingly.

5. The multilevel coupled cluster equations are solved.

Steps 1 and 5 are the expensive steps of a multilevel coupled cluster calculation with CNTOs. Step 1 scales as the lower level method, i.e., N^4 for a CCS calculation and N^5 for a CC2 calculation, and step 3 has the reduced multilevel coupled cluster scaling discussed in sections 2.1 and 2.2. Additionally, step 2 contains one-shot N^5 operations, however we have not observed this step to be limiting.

3 Results and discussion

The MLCC2 and MLCCSD codes are implemented in eT—an electronic structure program developed by the authors and collaborators.[56] In the following sections, we show the convergence of the MLCC2 and MLCCSD excitation energies and illustrate the usefulness of of the methods for larger systems.

3.1 Convergence of MLCC2 and MLCCSD excitation energies

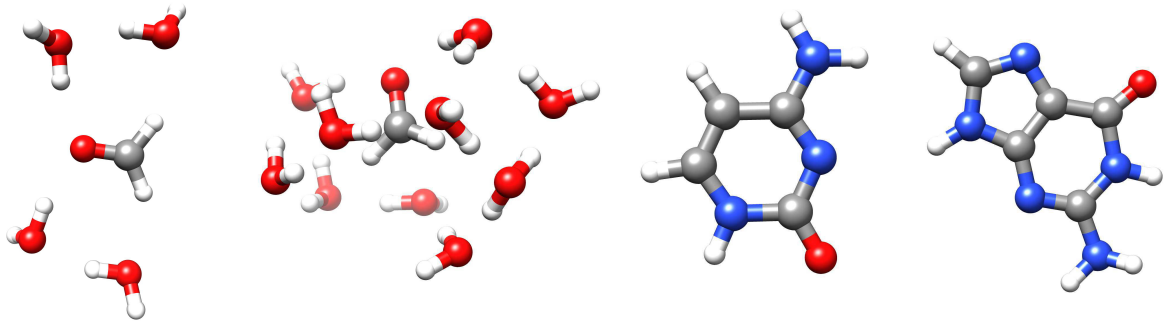

Figure 3: From left to right: formaldehyde with four water molecules, formaldehyde with ten water molecules, cytosine and guanine.

To demonstrate the convergence of the MLCC2 and MLCCSD excitation energies we consider four test systems (see Figure 3): formaldehyde together with four and ten water molecules, cytosine, and guanine. These are proof-of-concept calculations and the systems are chosen such that the reference CCSD excitation energies are readily available. The aug-cc-pVDZ basis is used in all calculations.

The convergence of the MLCC2 excitation energies for the test systems may be seen in Figures 4-7. For all the test systems, MLCC2 converges smoothly towards the CC2 excitation energies as the active space is enlarged. For formaldehyde with four and ten water molecules there are seven active virtual orbitals to every active occupied orbital. For cytosine and guanine there are six active virtual orbitals to every active occupied orbital. Convergence is faster for the formaldehyde-water systems than for cytosine and guanine. This may be because the multilevel coupled cluster methods are especially well suited for solvated systems and because a common CNTO basis (eqs. (34) and (35)) is used to describe two excited states in cytosine and guanine.

For an MLCCSD calculation, there are two options to generate the CNTOs: either from CCS or from CC2 excitation vectors. In CCS/CCSD or CCS/CC2/CCSD calculations we always use CNTOs from CCS excitation vectors. In CC2/CCSD calculations, we may choose which lower level method to use. Two-level MLCCSD results for the test systems

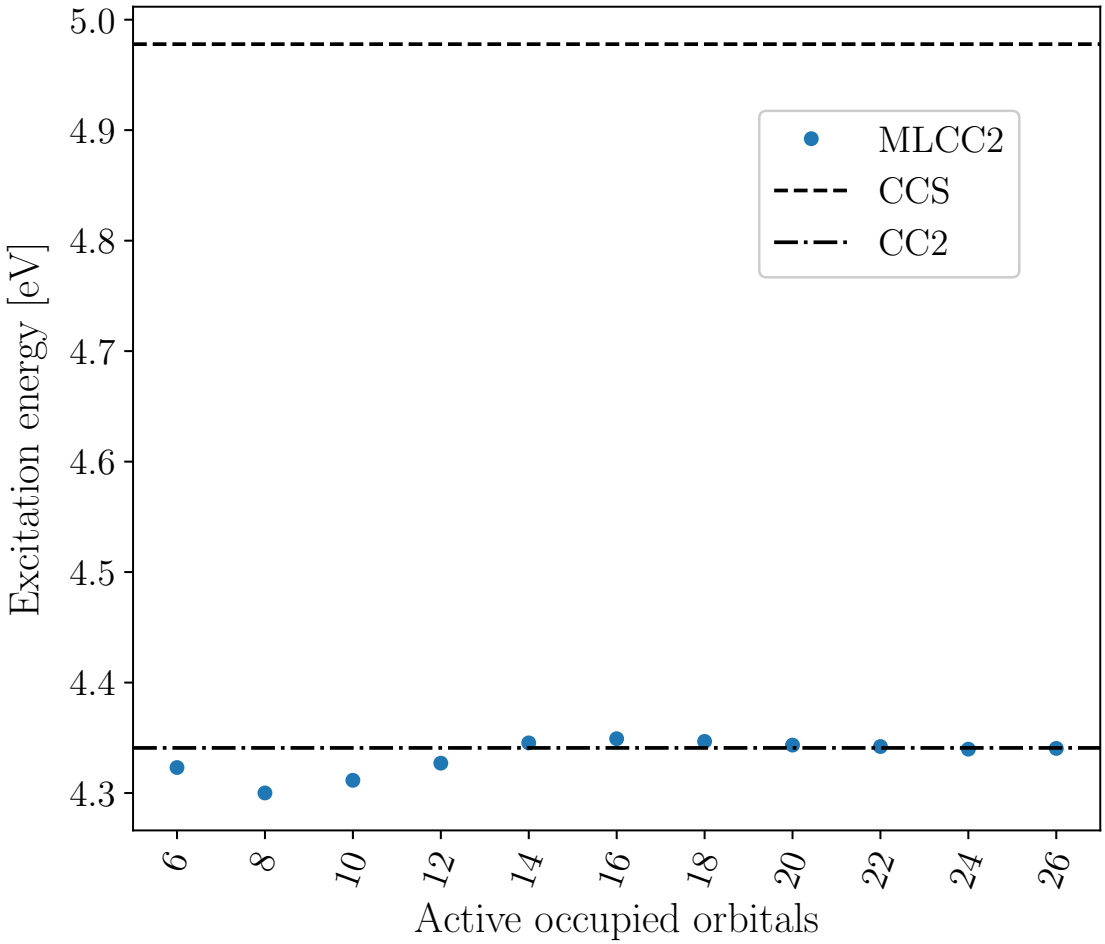

Figure 4: The lowest MLCC2 (aug-cc-pVDZ) excitation energy of formaldehyde with four water molecules. There are 7 active virtual orbitals for every active occupied orbital.

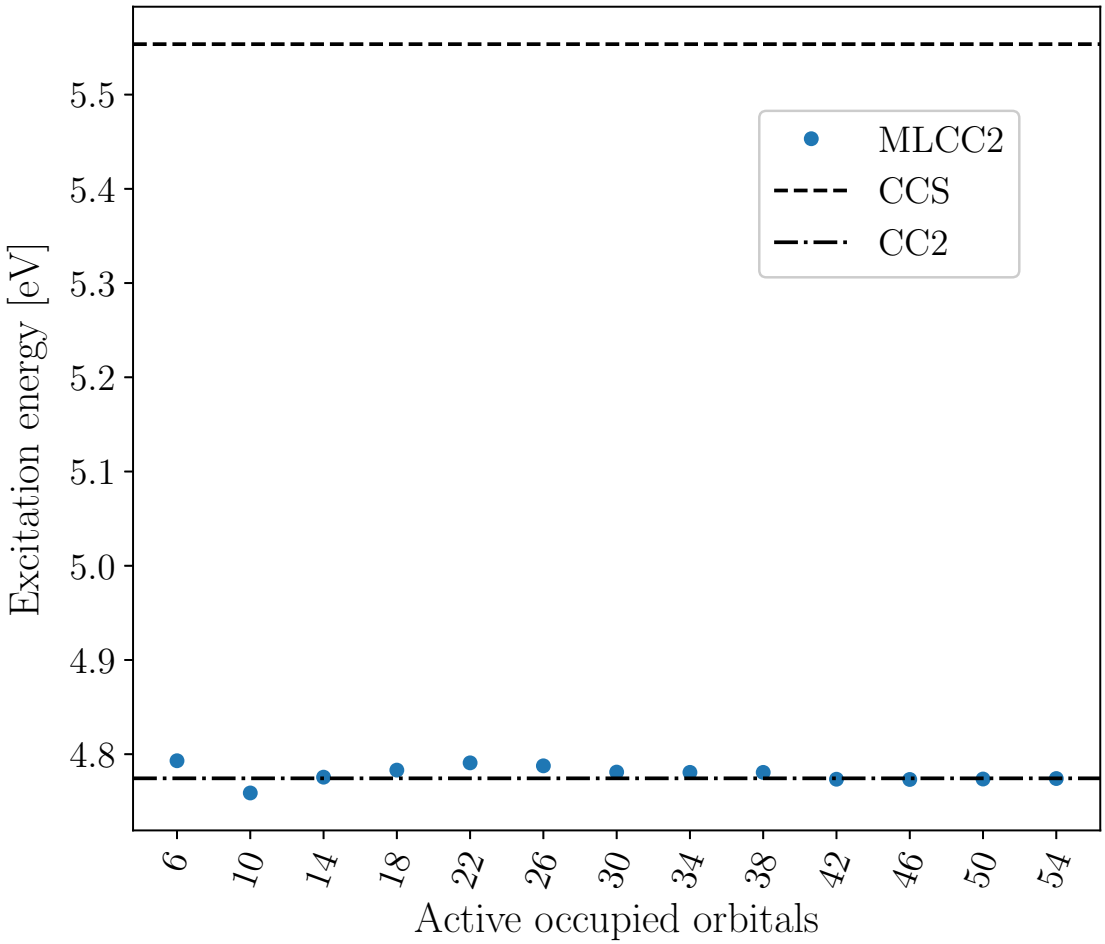

Figure 5: The lowest MLCC2 (aug-cc-pVDZ) excitation energy of formaldehyde with ten water molecules. There are 7 active virtual orbitals for every active occupied orbital.

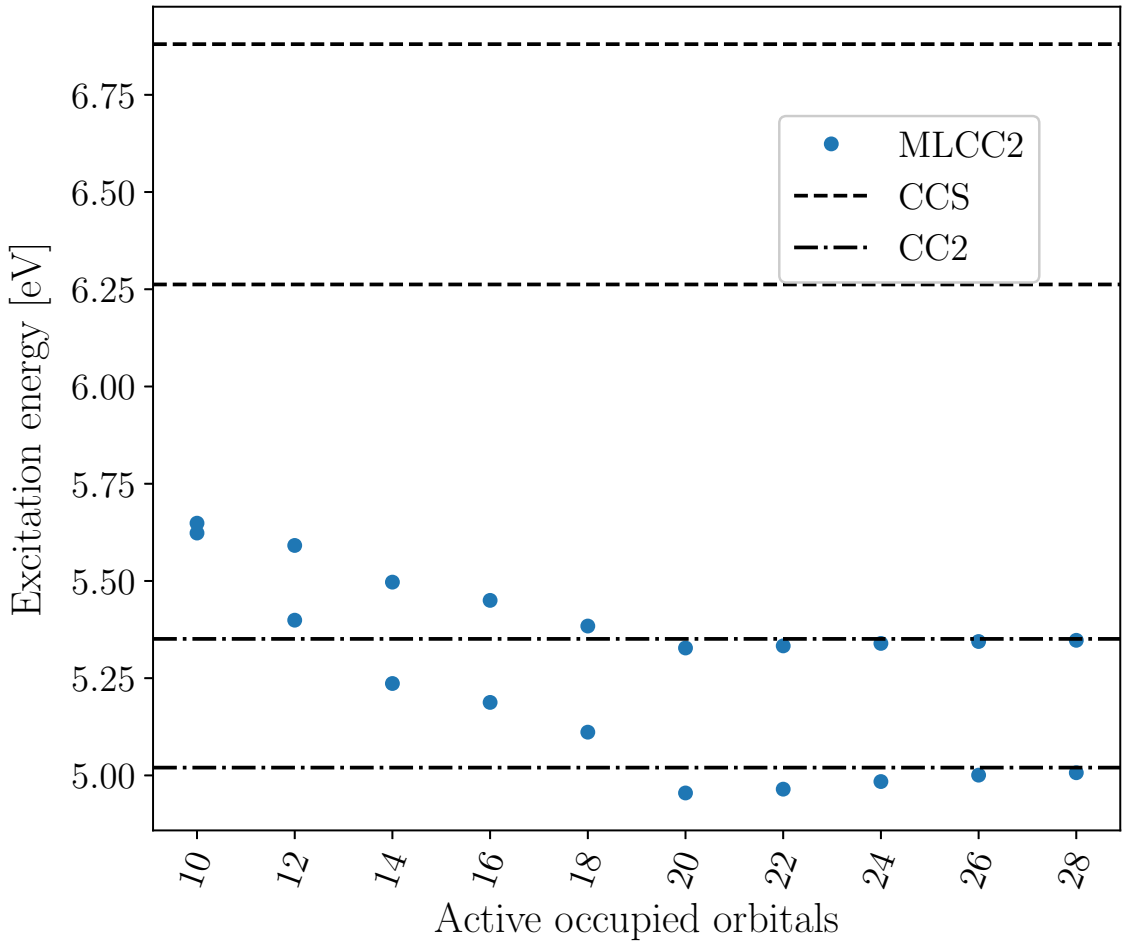

Figure 6: The two lowest MLCC2 (aug-cc-pVDZ) excitation energies of cytosine. There are 6 active virtual orbitals for every active occupied orbital.

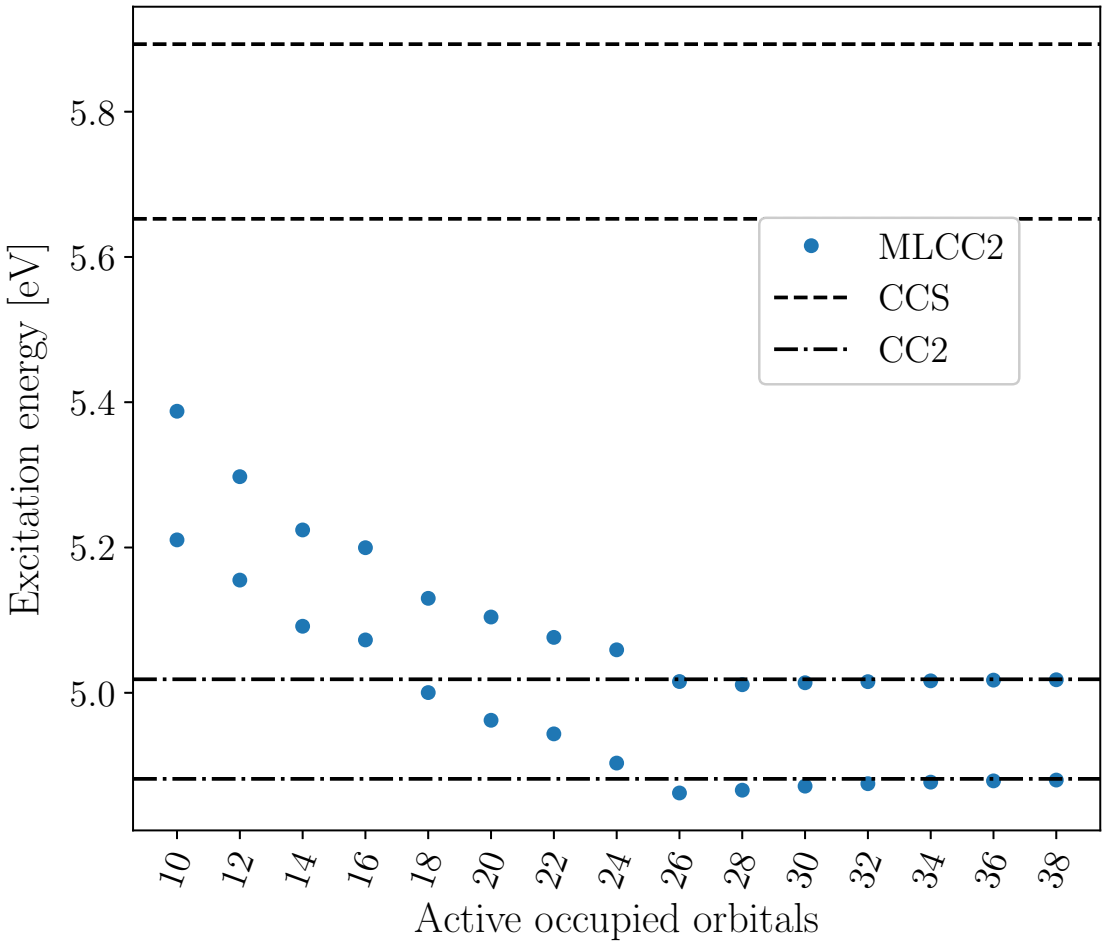

Figure 7: The two lowest MLCC2 (aug-cc-pVDZ) excitation energies of guanine. There are 6 active virtual orbitals for every active occupied orbital.

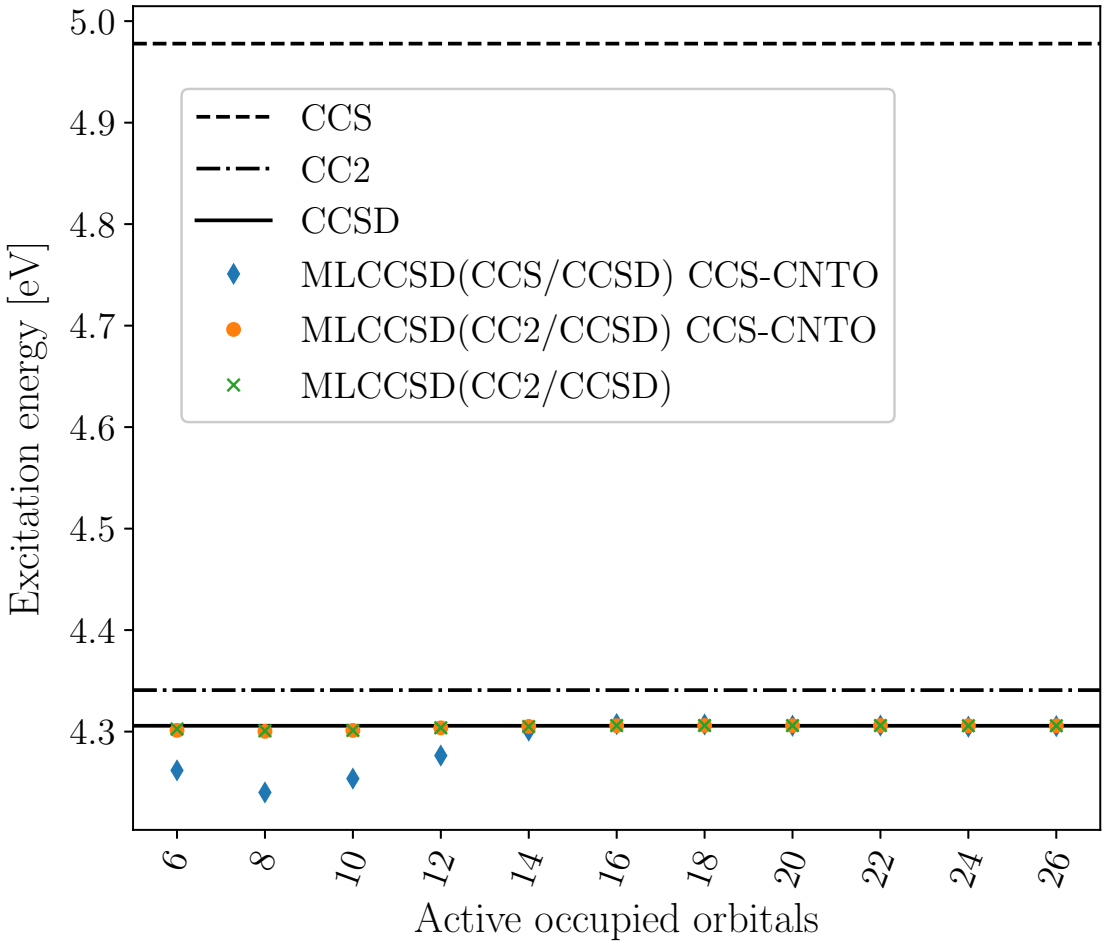

Figure 8: The lowest MLCCSD (aug-cc-pVDZ) excitation energy of formaldehyde with four water molecules. Approximated CNTOs (CCS-CNTOs) are used for both CCS/CCSD and CC2/CCSD calculations and compared to the CC2/CCSD calculations using CNTOs from CC2. There are 7 active virtual orbitals for every active occupied orbital.

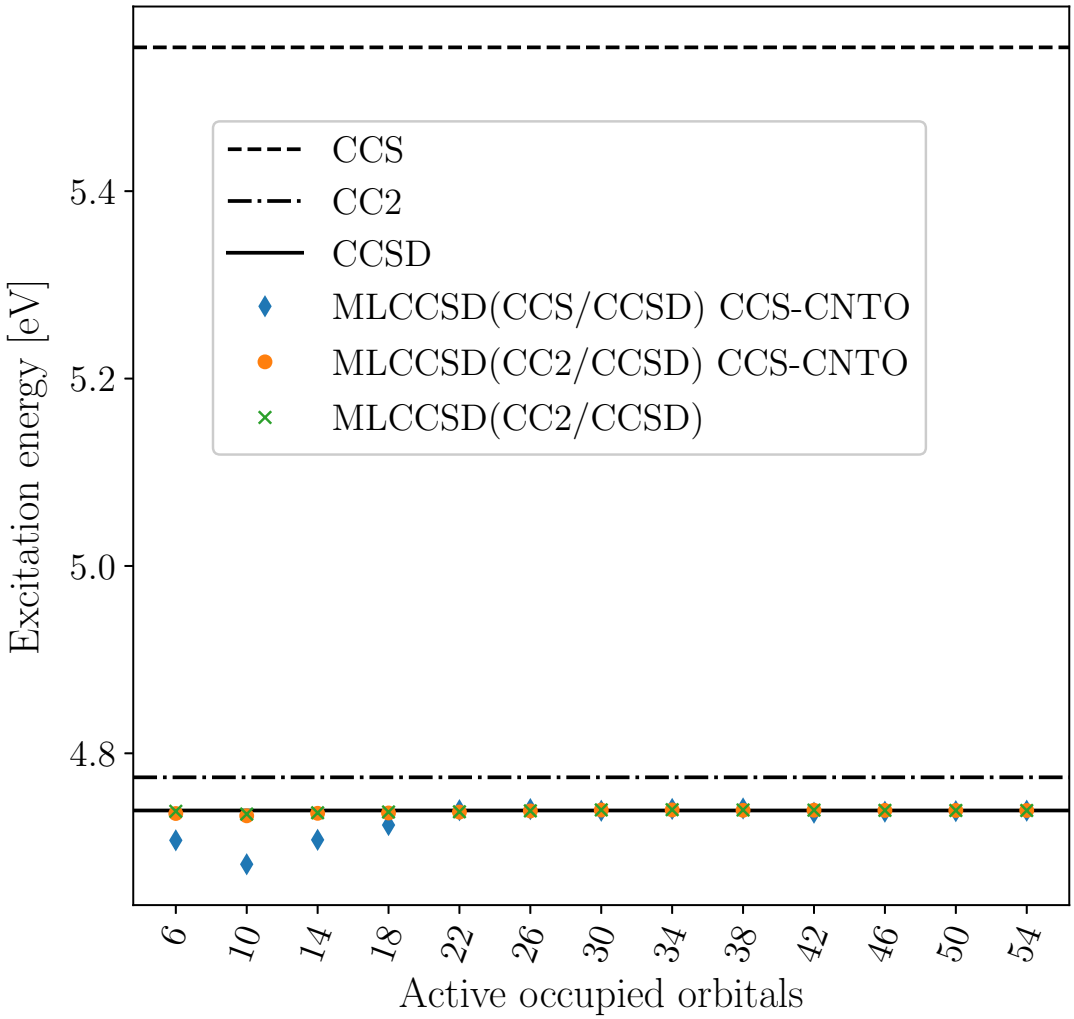

Figure 9: The lowest MLCCSD (aug-cc-pVDZ) excitation energy of formaldehyde with ten water molecules. Approximated CNTOs (CCS-CNTOs) are used for both CCS/CCSD and CC2/CCSD calculations. There are 7 active virtual orbitals for every active occupied orbital.

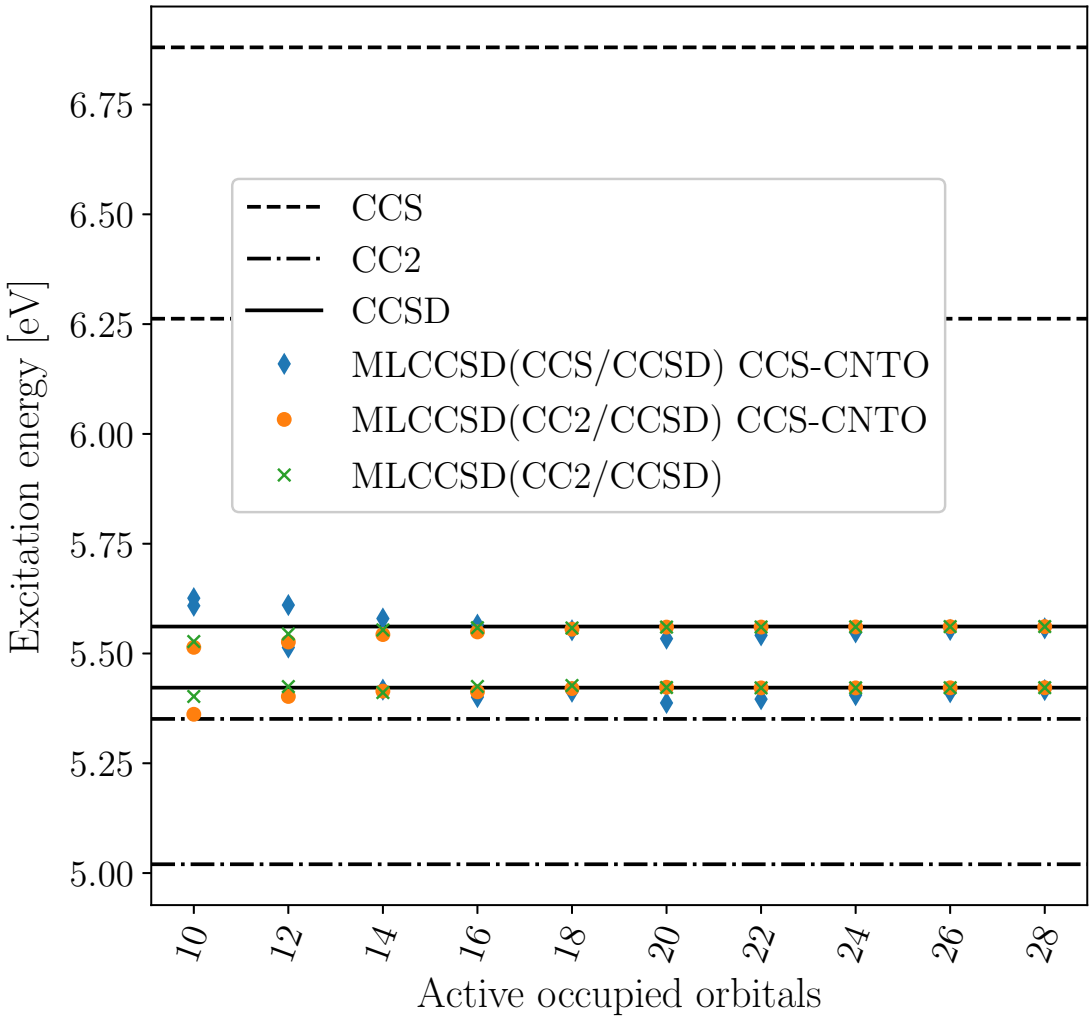

Figure 10: The two lowest MLCCSD (aug-cc-pVDZ) excitation energies of cytosine. Approximated CNTOs (CCS-CNTOs) are used for both CCS/CCSD and CC2/CCSD calculations and compared to the CC2/CCSD calculations using CNTOs from CC2. There are 6 active virtual orbitals for every active occupied orbital.

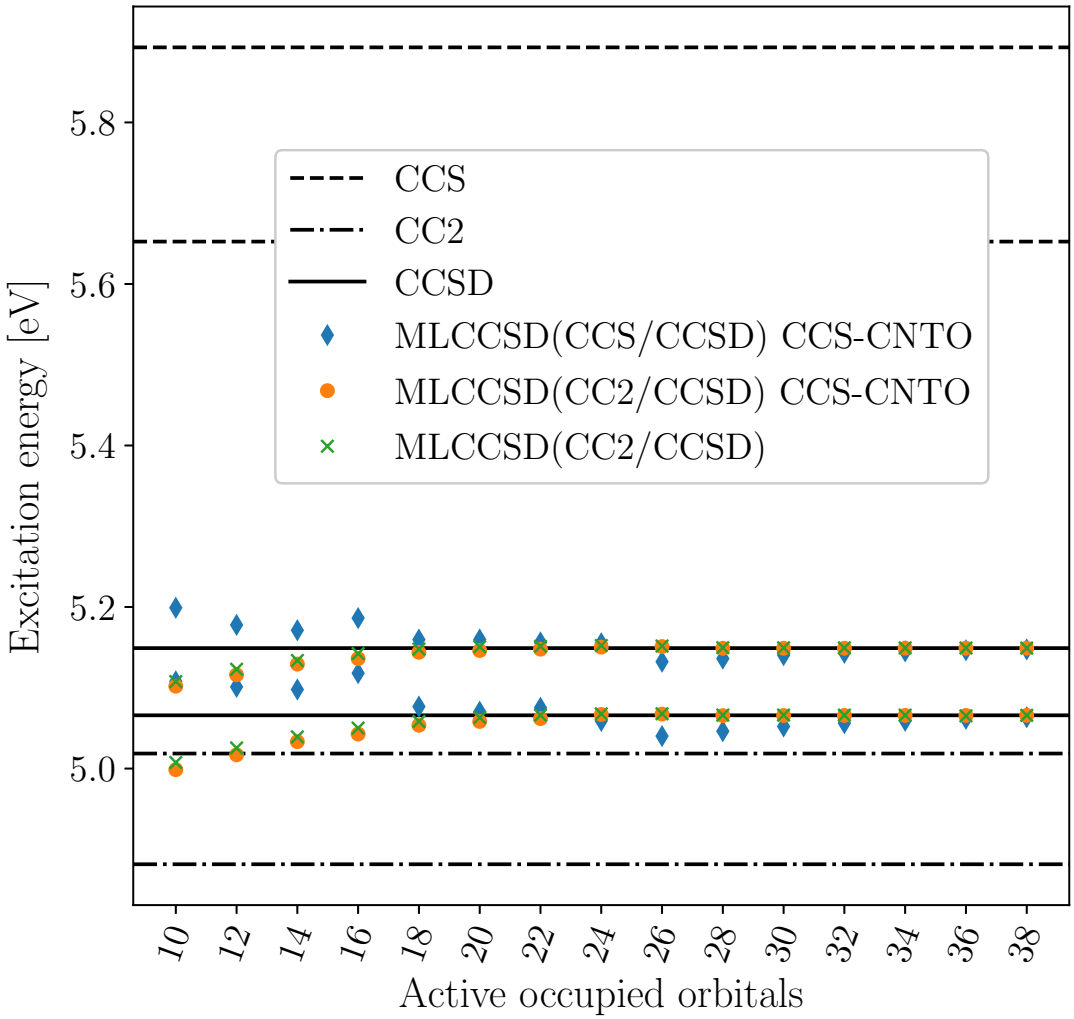

Figure 11: The two lowest MLCCSD (aug-cc-pVDZ) excitation energies of guanine. Approximated CNTOs (CCS-CNTOs) are used for both CCS/CCSD and CC2/CCSD calculations and compared to the CC2/CCSD calculations using CNTOs from CC2. There are 6 active virtual orbitals for every active occupied orbital.

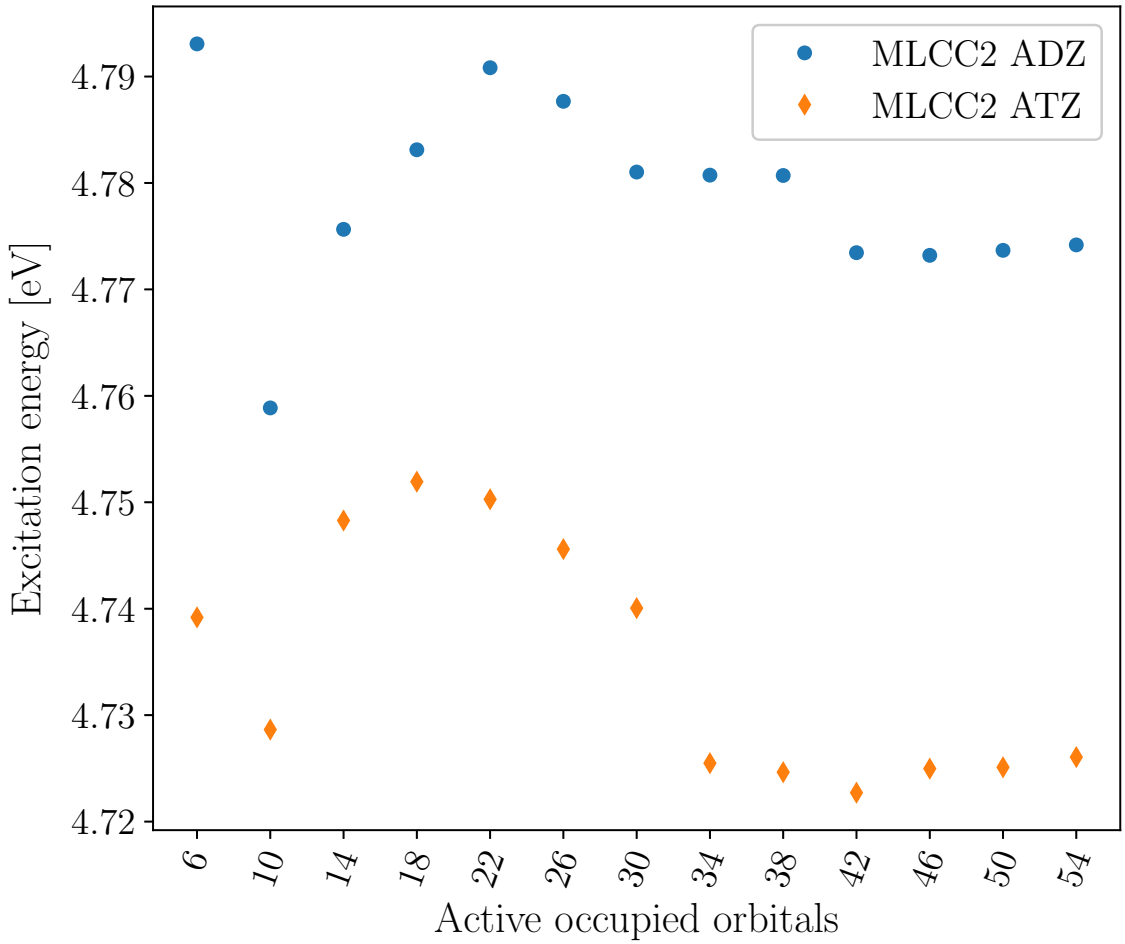

Figure 12: The lowest MLCC2 excitation energy of formaldehyde with ten water molecules using the aug-cc-pVDZ and the aug-cc-pVTZ bases. There are 7 and 17 virtual orbitals per occupied orbitals in the active spaces for aug-cc-pVDZ and aug-cc-pVTZ respectively.

are shown in Figures 8-11. The results indicate that a smaller active space might be sufficient in CC2/CCSD calculations, compared to the CCS/CCSD calculations. However, the CCS/CCSD calculation is cheaper than the CC2/CCSD calculation. Furthermore, the CNTOs generated from CCS are preferable, considering the cost of the lower level method (N^4 versus N^5 scaling) and the resulting MLCCSD excitation energies. Regardless, a limitation to the use of CNTOs for MLCCSD (either from CCS or CC2) is that the target excited state must be dominated by single excitations.

In Figure 12 we present the MLCC2 convergence for formaldehyde and ten water molecules using both the aug-cc-pVDZ and aug-cc-pVTZ basis sets. The convergence of MLCC2 is seen to carry over to the larger basis set.

3.2 Performance of MLCC2

We have performed timing comparisons of MLCC2 and CC2 for formaldehyde and ten water molecules. All calculations are performed with the eT program. The eT program is based on Cholesky decomposition of the electron repulsion integrals,[11] however, in these calculations we have not introduced any approximations of the integrals: a Cholesky decomposition threshold of 10^{-8} is used. The electron repulsion integrals are constructed directly from the Cholesky vectors during the coupled cluster calculation. The Cholesky decomposition threshold may be loosened in order to reduce the cost of both MLCC2 and CC2 calculations.

The results are given in Table 1. With MLCC2, excitation energies of CC2 quality may be obtained at a significantly reduced cost. The additional cost of orbital construction in the MLCC2 calculation amounts to less than 10 min wall time. We have also included a comparison to frozen-core CC2 (CC2-FC). By using CNTOs for multilevel coupled cluster valence excitations, the core orbitals automatically become part of the inactive orbital space. Additionally, in the multilevel coupled cluster framework, extra savings—without loss of accuracy—are obtained by restricting the virtual space. Note that for all MLCC2 calculations

Table 1: Timing comparisons for the ground and excited state calculations of formaldehyde and ten water molecules using MLCC2, CC2 and CC2-FC. The wall times t^{gs} and t^{es} to converge the ground and excited state equations (to within a 10^{-6} residual threshold) are given. Also, the average wall time per Jacobian matrix transformation $t^{\mathrm{A,es}}$ for the excited state calculation is reported. Timings were made on an Intel Xeon CPU E5-2699 v4 with 1.5TB shared memory using 22 threads.

Method	$n_o^{\mathrm{active}}/n_v^{\mathrm{active}}$	t^{gs}	$t^{\mathrm{A,es}}$	t^{es}	ω [eV]
MLCC2	10/70	9.9 min	23.7 s	9.5 min	4.7589
	18/126	11.2 min	28.5 s	12.4 min	4.7831
	26/182	13.4 min	38.0 s	19.0 min	4.7877
	34/238	15.4 min	52.7 s	26.1 min	4.7807
	42/294	18.7 min	1.3 min	46.6 min	4.7735
	50/350	22.4 min	1.8 min	1.5 h	4.7737
CC2-FC	58/416	17.3 min	2.1 min	2.1 h	4.7782
CC2	58/416	22.4 min	2.7 min	2.9 h	4.7744

the deviation from full CC2 is of the order 10^{-2}eV or less, which is well within the expected error of CC2.

3.3 Two larger systems

Finally, we consider two larger systems: paranitroaniline (PNA)—a system known for a large solvation shift of its lowest lying charge transfer excitation[57–60]—and the penicillin amoxycillin (see Figure 13).

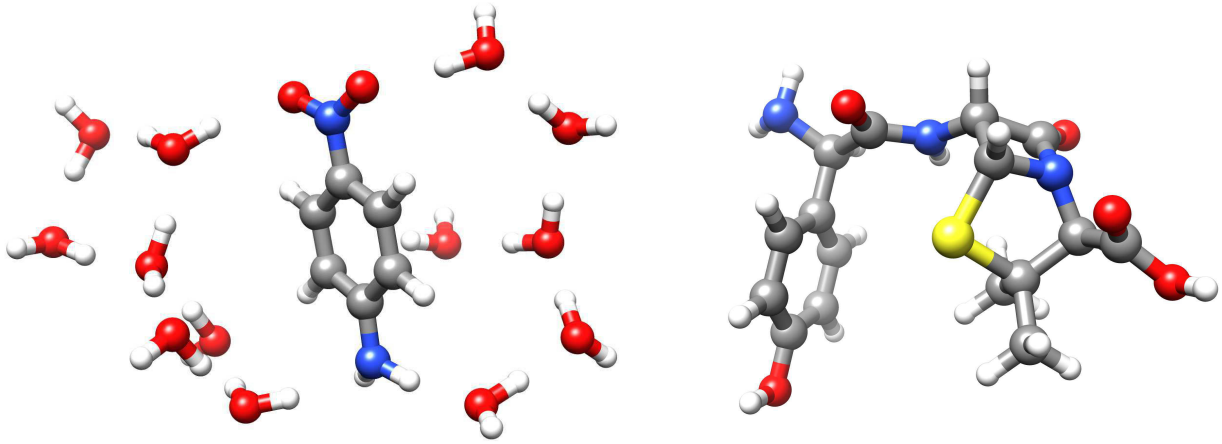

Figure 13: On the left: paranitroaniline with water. On the right: amoxicillin.

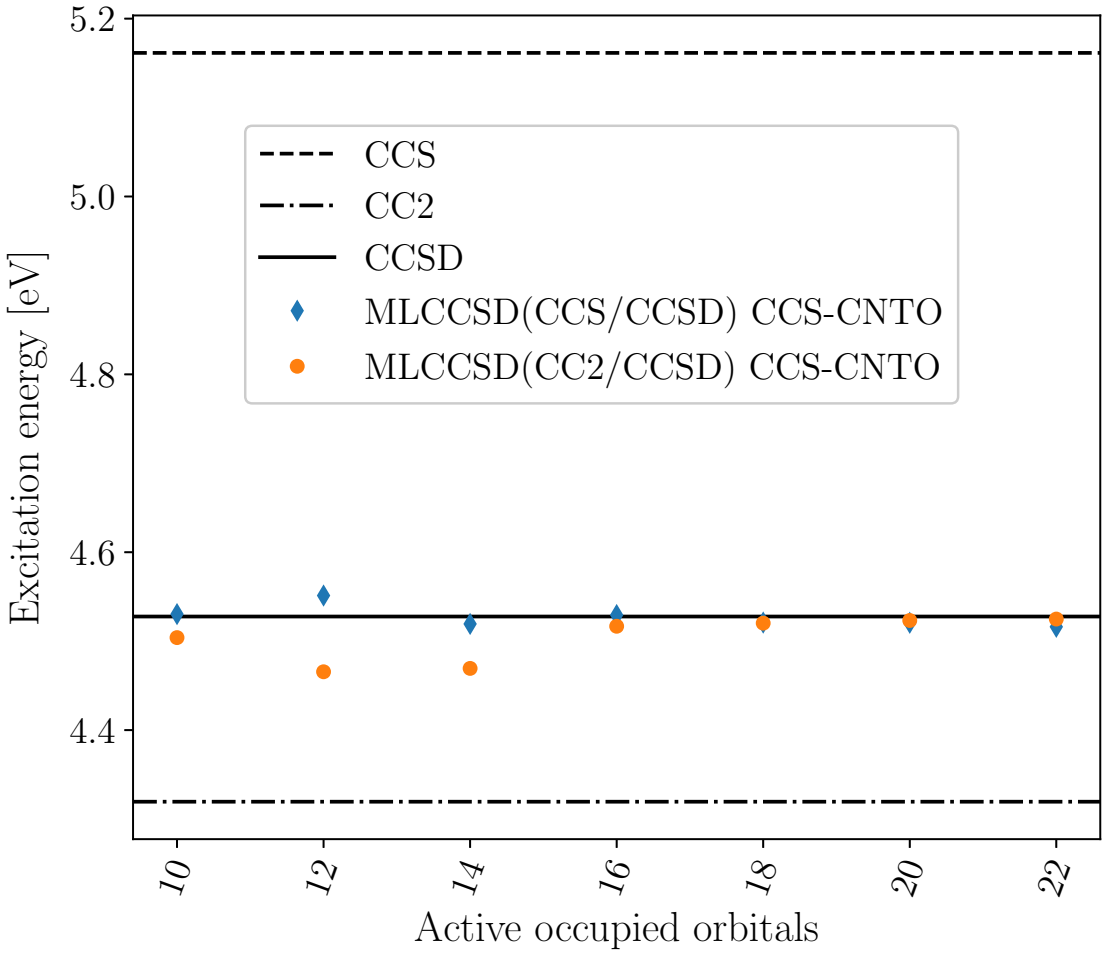

Figure 14: The second MLCCSD (aug-cc-pVDZ) excitation energy of PNA. Approximated CNTOs (CCS-CNTOs) are used for both CCS/CCSD and CC2/CCSD calculations. There are 6 active virtual orbitals for every active occupied orbital.

We have performed MLCCSD calculations on PNA and on PNA with 13 water molecules. The geometry of PNA and 13 water molecules was optimized at B3LYP/6-31+G* level using the NWChem software.[61] For the isolated PNA, we have removed the water molecules without any relaxation of the PNA geometry.

The convergence of the two-level MLCCSD (CCS/CCSD and CC2/CCSD) excitation energies corresponding to the charge-transfer excitation in PNA can be seen from Figure 14. There are six active virtual orbitals for every active occupied orbital. The charge transfer excitation is the second excited state for the given geometry and basis set (aug-cc-pVDZ).

To treat PNA together with 13 water molecules, we have performed a two-level (CCS/CCSD) and a three-level MLCCSD (CCS/CC2/CCSD) calculation. For the isolated PNA, an active space of $n_o^{\text{active}} = 16$ yields an excitation energy with an error of 10^{-2}eV with respect to full CCSD. In the MLCCSD calculations on PNA with water, we have therefore selected to use 16 occupied CCSD orbitals. In the three-level calculation, we use 20 occupied CC2 orbitals (there are 36 occupied orbitals in PNA), leaving 65 occupied orbitals in the CCS orbital space. In the two-level calculation, there are 85 occupied orbitals in the CCS orbital space. We use the cc-pVDZ basis for the water molecules, but keep the aug-cc-pVDZ basis on PNA. The full system has $n_o = 101$ and $n_v = 495$. However, in $\mathcal{O}^{\text{CC2}}$ and $\mathcal{O}^{\text{CCSD}}$ we continue to use six active virtual orbitals per active occupied orbital (the fraction which was used in the calculations on the isolated PNA molecule).

For PNA in water, the excitation energy of the charge-transfer excitation is lowered and becomes the first excited state. In Table 2 we present the MLCCSD, CCS, CC2, and CCSD excitation energies of PNA and PNA with 13 water molecules. We also report the solvation shifts. The CCS/CC2/CCSD solvation shift is calculated using the two-level (CC2/CCSD) calculation on the isolated PNA with 16 occupied CCSD orbitals and 20 occupied CC2 orbitals. Both MLCCSD calculations effectively captures the solvation shift with an error of less than 0.03eV. The three-level calculation performs marginally better than the two-level calculation.

Table 2: MLCCSD, CCSD, CC2 and CCS calculations of the lowest charge transfer excitation of PNA (ω) and for PNA with 13 water molecules (ω_{w}). The corresponding solvation shift $\Delta\omega = \omega_{\mathrm{w}} - \omega$ is also given. In the CCS/CCSD results 16 active occupied orbitals were used, the remaining occupied orbitals are in the CCS orbital space. In the CCS/CC2/CCSD calculation there are 16 occupied CCSD orbitals, 20 occupied CC2 orbitals, and the remaining occupied orbitals are in the CCS orbital space.

Method	ω [eV]	ω_{w} [eV]	$\Delta\omega$ [eV]
CCS	5.162	4.757	-0.405
CC2	4.320	3.595	-0.725
CCSD	4.528	3.849	-0.679
CCS/CCSD	4.529	3.877	-0.652
CCS/CC2/CCSD	4.517	3.830	-0.687

The penicillin amoxicillin is depicted in Figure 13. In the aug-cc-pVDZ basis the system has 96 occupied orbitals and 654 virtual orbitals. We have performed MLCC2 and MLCCSD (CCS/CCSD) calculations. The results are given in Table 3 and are compared to CC2, CCSD and CCSD-FC. We have used a set of different thresholds in the Cholesky decomposition of the integrals to illustrate how multilevel coupled cluster calculations may be combined with integral approximations. The error of MLCC2 and MLCCSD with respect to the standard coupled cluster models—CC2 and CCSD (CCSD-FC)—is on the order 10^{-2}eV, which is within the errors expected of both CCSD and CC2.

Table 3: Lowest excitation energy of amoxicillin with MLCC2, MLCCSD (CCS/CCSD), CC2, CCSD-FC and CCSD. Different thresholds for the Cholesky decomposition two-electron integrals τ are used.

Method	n_o^{active}	$\tau = 10^{-2}$	$\tau = 10^{-3}$	$\tau = 10^{-4}$
	25	4.913 eV	4.918 eV	4.915 eV
MLCC2	30	4.857 eV	4.861 eV	4.858 eV
	35	4.847 eV	4.851 eV	4.848 eV
	25	4.893 eV	4.901 eV	4.899 eV
MLCCSD	30	4.869 eV	4.876 eV	4.873 eV
	35	4.867 eV	4.877 eV	4.872 eV
CC2	-	4.833 eV	4.836 eV	4.834 eV
CCSD-FC	-	4.877 eV	4.885 eV	4.883 eV
CCSD	-	4.877 eV	-	-

4 Concluding remarks

With the multilevel coupled cluster framework, intensive molecular properties may be calculated at an accuracy approaching that of the higher level coupled cluster methodused. The scaling of multilevel coupled cluster theory will approach that of the lower level method when the inactive space is sufficiently large. However, one must also consider the cost of determining the active orbital spaces. We introduce the MLCC2 method, which may be applied to large systems in order to approximate CC2 excitation energies. We have also presented a new formulation of MLCCSD where the CC2 double excitation amplitudes are determined from a closed-form expression. Both methods show smooth convergence of excitation energies towards the value of the higher level coupled cluster method as the active space is extended. Calculations on paranitroaniline with water molecules demonstrates the efficacy of multilevel coupled cluster methods to describe solvation effects. These calculations also show that multilevel methods may be used for charge transfer excitations. The correlated natural transition orbitals are well suited to determine the active spaces for multilevel valence excitations. The CNTOs generated from CCS excitation vectors are preferable for MLCCSD when considering both cost and accuracy.

Acknowledgement

We acknowledge funding from the Marie Skłodowska-Curie European Training Network through grant agreement no. 765739 (COSINE) and the Norwegian Reasearch Council grants CCGPU (no. 263110) and TheoLight (no. 275506). H. K. acknowledges the Otto Mønsted Fond and S. D. F. acknowledges Fondet til professor Leif Tronstads minne. We acknowledge computer resources from NOTUR through project no. nn2962k.

Supporting Information Available

Cartesian coordinates for the geometries.

References

(1) Helgaker, T.; Jorgensen, P.; Olsen, J. *Molecular electronic-structure theory*; John Wiley & Sons, 2014.

(2) Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291.

(3) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. *Chem. Phys. Lett.* **1993**, *208*, 359–363.

(4) Vahtras, O.; Almlöf, J.; Feyereisen, M. W. Integral approximations for LCAO-SCF calculations. *Chem. Phys. Lett.* **1993**, *213*, 514–518.

(5) Kendall, R. A.; Früchtl, H. A. The impact of the resolution of the identity approximate integral method on modern ab initio algorithm development. *Theor. Chem. Acc.* **1997**, *97*, 158–163.

(6) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett.* **1995**, *240*, 283 – 290.

(7) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. *Theor. Chem. Acc.* **1997**, *97*, 119–124.

(8) Weigend, F. A fully direct RI-HF algorithm: Implementation, optimised auxiliary basis sets, demonstration of accuracy and efficiency. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285–4291.

(9) Beebe, N. H. F.; Linderberg, J. Simplifications in the generation and transformation of two-electron integrals in molecular calculations. *Int. J. Quantum Chem.* **1977**, *12*, 683–705.

(10) Koch, H.; Sánchez de Merás, A.; Pedersen, T. B. Reduced scaling in electronic structure calculations using Cholesky decompositions. *J. Chem. Phys.* **2003**, *118*, 9481–9484.

(11) Folkestad, S. D.; Kjønstad, E. F.; Koch, H. An efficient algorithm for Cholesky decomposition of electron repulsion integrals. *J. Chem. Phys.* **2019**, *150*, 194112.

(12) Pulay, P. Localizability of dynamic electron correlation. *Chem. Phys. Lett.* **1983**, *100*, 151–154.

(13) Sæbø, S.; Pulay, P. Local treatment of electron correlation. *Annu. Rev. Phys. Chem.* **1993**, *44*, 213–236.

(14) Pulay, P.; Sæbø, S. Orbital-invariant formulation and second-order gradient evaluation in Møller-Plesset perturbation theory. *Theor. Chim. Acta.* **1986**, *69*, 357–368.

(15) Sæbø, S.; Pulay, P. Fourth-order Møller–Plessett perturbation theory in the local correlation treatment. I. Method. *J. Chem. Phys.* **1987**, *86*, 914–922.

(16) Sæbø, S.; Pulay, P. Local configuration interaction: An efficient approach for larger molecules. *Chem. Phys. Lett.* **1984**, *113*, 13–18.

(17) Hampel, C.; Werner, H.-J. Local treatment of electron correlation in coupled cluster theory. *J. Chem. Phys.* **1996**, *104*, 6286–6297.

(18) Schütz, M.; Werner, H.-J. Low-order scaling local electron correlation methods. IV. Linear scaling local coupled-cluster (LCCSD). *J. Chem. Phys.* **2001**, *114*, 661–681.

(19) 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.

(20) 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.

(21) Flocke, N.; Bartlett, R. J. A natural linear scaling coupled-cluster method. *J. Chem. Phys.* **2004**, *121*, 10935–10944.

(22) Fedorov, D. G.; Kitaura, K. Coupled-cluster theory based upon the fragment molecular-orbital method. *J. Chem. Phys.* **2005**, *123*, 134103.

(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) Auer, A. A.; Nooijen, M. Dynamically screened local correlation method using enveloping localized orbitals. *J. Chem. Phys.* **2006**, *125*, 024104.

(25) 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.

(26) Neese, F.; Hansen, A.; Liakos, D. G. Efficient and accurate approximations to the local coupled cluster singles doubles method using a truncated pair natural orbital basis. *J. Chem. Phys.* **2009**, *131*, 064103.

(27) Ziółkowski, M.; Jansik, B.; Kjærgaard, T.; Jørgensen, P. Linear scaling coupled cluster method with correlation energy based error control. *J. Chem. Phys.* **2010**, *133*, 014107.

(28) Rolik, Z.; Kállay, M. A general-order local coupled-cluster method based on the cluster-in-molecule approach. *J. Chem. Phys.* **2011**, *135*, 104111.

(29) Yang, J.; Chan, G. K.-L.; Manby, F. R.; Schütz, M.; Werner, H.-J. The orbital-specific-virtual local coupled cluster singles and doubles method. *J. Chem. Phys.* **2012**, *136*, 144105.

(30) Riplinger, C.; Neese, F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.* **2013**, *138*, 034106.

(31) Kumar, A.; Crawford, T. D. Frozen virtual natural orbitals for coupled-cluster linear-response theory. *J. Phys. Chem. A* **2017**, *121*, 708–716.

(32) Korona, T.; Werner, H.-J. Local treatment of electron excitations in the EOM-CCSD method. *J. Chem. Phys.* **2003**, *118*, 3006–3019.

(33) Kats, D.; Korona, T.; Schütz, M. Local CC2 electronic excitation energies for large molecules with density fitting. *J. Chem. Phys.* **2006**, *125*, 104106.

(34) Helmich, B.; Hättig, C. Local pair natural orbitals for excited states. *J. Chem. Phys.* **2011**, *135*, 214106.

(35) 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.

(36) 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.

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

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

(39) Luzanov, A.; Sukhorukov, A.; Umanskii, V. Application of transition density matrix for analysis of excited states. *Theor. Exp. Chem.* **1976**, *10*, 354–361.

(40) Martin, R. L. Natural transition orbitals. *J. Chem. Phys.* **2003**, *118*, 4775–4777.

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

(42) 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.

(43) 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.

(44) Levitt, M.; Warshel, A. Computer simulation of protein folding. *Nature* **1975**, *253*, 694.

(45) Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.

(46) Vreven, T.; Byun, K. S.; Komáromi, I.; Dapprich, S.; Montgomery Jr, J. A.; Morokuma, K.; Frisch, M. J. Combining quantum mechanics methods with molecular mechanics methods in ONIOM. *J. Chem. Theory Comput.* **2006**, *2*, 815–826.

(47) Tomasi, J.; Cammi, R.; Mennucci, B.; Cappelli, C.; Corni, S. Molecular properties in solution described with a continuum solvation model. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5697–5712.

(48) Myhre, R. H.; Sánchez de Merás, A. M. J.; Koch, H. The extended CC2 model ECC2. *Mol. Phys.* **2013**, *111*, 1109–1118.

(49) Myhre, R. H.; Sánchez de Merás, A. M. J.; Koch, H. Multi-level coupled cluster theory. *J. Chem. Phys.* **2014**, *141*, 224105.

(50) Myhre, R. H.; Koch, H. The multilevel CC3 coupled cluster model. *J. Chem. Phys.* **2016**, *145*, 44111.

(51) Aquilante, F.; Bondo Pedersen, T.; Sánchez de Merás, A.; Koch, H. Fast noniterative orbital localization for large molecules. *J. Chem. Phys.* **2006**, *125*, 174101.

(52) Koch, H.; Jørgensen, P. Coupled cluster response functions. *J. Chem. Phys.* **1990**, *93*, 3333–3344.

(53) Pedersen, T. B.; Koch, H. Coupled cluster response functions revisited. *J. Chem. Phys.* **1997**, *106*, 8059–8072.

(54) Davidson, E. R. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J. Comput. Phys.* **1975**, *17*, 87–94.

(55) Hirao, K.; Nakatsuji, H. A generalization of the Davidson's method to large nonsymmetric eigenvalue problems. *J. Comput. Phys.* **1982**, *45*, 246–254.

(56) *The eT program will be described in a forthcoming publication*

(57) Kosenkov, D.; Slipchenko, L. V. Solvent effects on the electronic transitions of p-nitroaniline: A QM/EFP study. *J. Phys. Chem. A* **2010**, *115*, 392–401.

(58) Slipchenko, L. V. Solvation of the excited states of chromophores in polarizable environment: orbital relaxation versus polarization. *J. Phys. Chem. A* **2010**, *114*, 8824–8830.

(59) Sneskov, K.; Schwabe, T.; Christiansen, O.; Kongsted, J. Scrutinizing the effects of polarization in QM/MM excited state calculations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 18551–18560.

(60) Eriksen, J. J.; Sauer, S. P.; Mikkelsen, K. V.; Christiansen, O.; Jensen, H. J. A.; Kongsted, J. Failures of TDDFT in describing the lowest intramolecular charge-transfer excitation in para-nitroaniline. *Mol. Phys.* **2013**, *111*, 1235–1248.

(61) Valiev, M.; Bylaska, E.; Govind, N.; Kowalski, K.; Straatsma, T.; Dam, H. V.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T.; de Jong, W. NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun.* **2010**, *181*, 1477 – 1489.

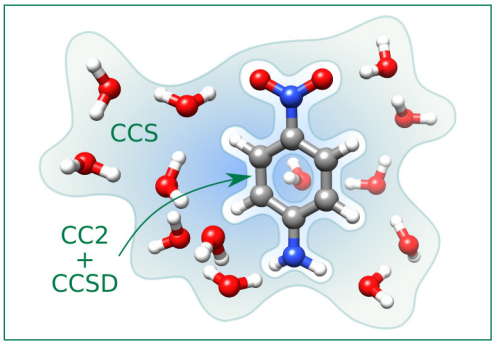

Figure 15: For Table of Contents Only